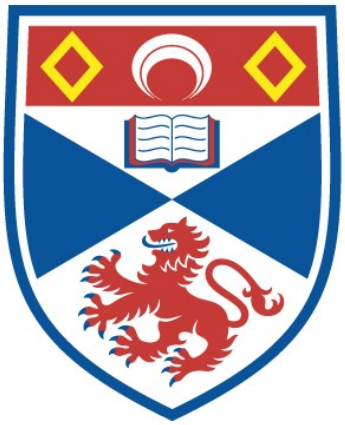
MOLECULAR TAXONOMY OF THE GENUS 'ACACIA (MILLER)' AND RELATED SPECIES IN THE MIMOSOIDEAE

Julian Robinson

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1997

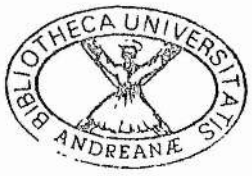
Molecular taxonomy of the genus Acacia MILLER. and related species in the Mimosoideae

by

Julian Robinson.

A thesis submitted to the University of St.Andrews for the degree of Doctor of Philosophy.

School of Biological and Preclinical Medicine,
University of St. Andrews,
September 1995.



Declaration.

I, Julian Robinson, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degrees or professional qualification

Julian Robinson.
September 1995.

Statement.

I was admitted to the faculty of science of the University of St. Andrews under Ordinance General No. 12 in October 1991 and as a candidate for the degree of Ph.D. in October 1992.

Julian Robinson.
September 1995.

Certificate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Requirements appropriate to the degree of Ph.D.

R.Ingram.
September 1995.

Acknowledgements.

Completion of this thesis would have been impossible without the help of several colleagues. These are:

Dr. Ruth Ingram, my supervisor, who has patiently guided me through the pitfalls of producing this thesis.

Dr Stephen Harris, for his tuition in the practical side of this thesis, and for giving me an understanding of the principles of molecular taxonomy.

Chris Fagg. The majority of plant material used in this thesis was collected by Chris, and provided by the Oxford Forestry Institute. Chris arranged for the provision of this material. I am also indebted to him for providing me with the opportunity to study *Acacia* in Kenya.

Dr Lourdes Rico Arce. During the final period of this thesis Lulu provided me with material from many difficult to obtain species, I am grateful to her for this.

Bruce Maslin, for sending me draft copies of papers used in this thesis, and also for providing a shortened list of 'critical' taxa from subgenus Phyllodineae.

Dr Richard Abbott, for taking over my supervision during my final year and for finalising details concerning the submission of this thesis.

The receipt of a studentship from the University of St.Andrews is gratefully acknowledged.

Abstract.

The construction of a chloroplast DNA restriction enzyme site based phylogeny for the genus *Acacia* and several related genera was completed. This investigation was initiated due to a perceived need for an independent viewpoint on the phylogeny of the genus. The results of this analysis challenge several previous attempts at classification of the genus using morphological characters. The cpDNA data suggest that subgenus *Acacia* and subgenus *Aculeiferum* are closely related. The third subgenus of *Acacia*, subgenus *Phyllodineae*, is probably unrelated to either of the other two subgenera. Instead it appears to be closely related to taxa in the Ingeae, a sister tribe to the Acacieae. This analysis also suggests that *Faidherbia albida* is basal to the Ingeae.

The interspecific relationships of taxa within each of the subgenera of *Acacia* were also partly resolved. Within subgenus *Acacia* the African accessions studied could not be resolved due to a lack of variation. The American accessions of subgenus *Acacia* were resolved, and appear to confirm groupings within these taxa suggested by morphological analyses. Within subgenus *Aculeiferum* the interspecific relationships were less clear, and little support was given to the sections proposed by Vassal (1972), with the exception of section *Filicinae*, which appears to be monophyletic. The relationships of taxa within subgenus *Aculeiferum* in regard to their geographical origin suggest that subgenus *Aculeiferum* was quite well differentiated when Gondwanaland fragmented.

The results of an investigation into the putative hybrid *A.laeta* appear to confirm earlier suggestions that it is a hybrid between *A.senegal* and *A.mellifera*. The appearance of non-additive hybrid phenotypes in the ribosomal nuclear DNA studied prevented an unequivocal determination of the parents of *A.laeta*. On the basis of the cpDNA characters it appears that *A.mellifera* is always the maternal parent.

The relationships of the subspecies of *A.tortilis* using RAPD techniques, revealed an interesting divide between north African and south African accessions. The taxa studied appeared to be grouped primarily according to their geographical location, the subspecific designation appearing to be of secondary importance. The divide between the accessions appears to be the boundary of one of the phytogeographical regions of Wickens (1976), corresponding approximately to the Kenyan-Tanzanian border. The physical or biological basis for this boundary is unknown. The investigation also proved the utility of the RAPD technique for investigations of this nature.

Contents.

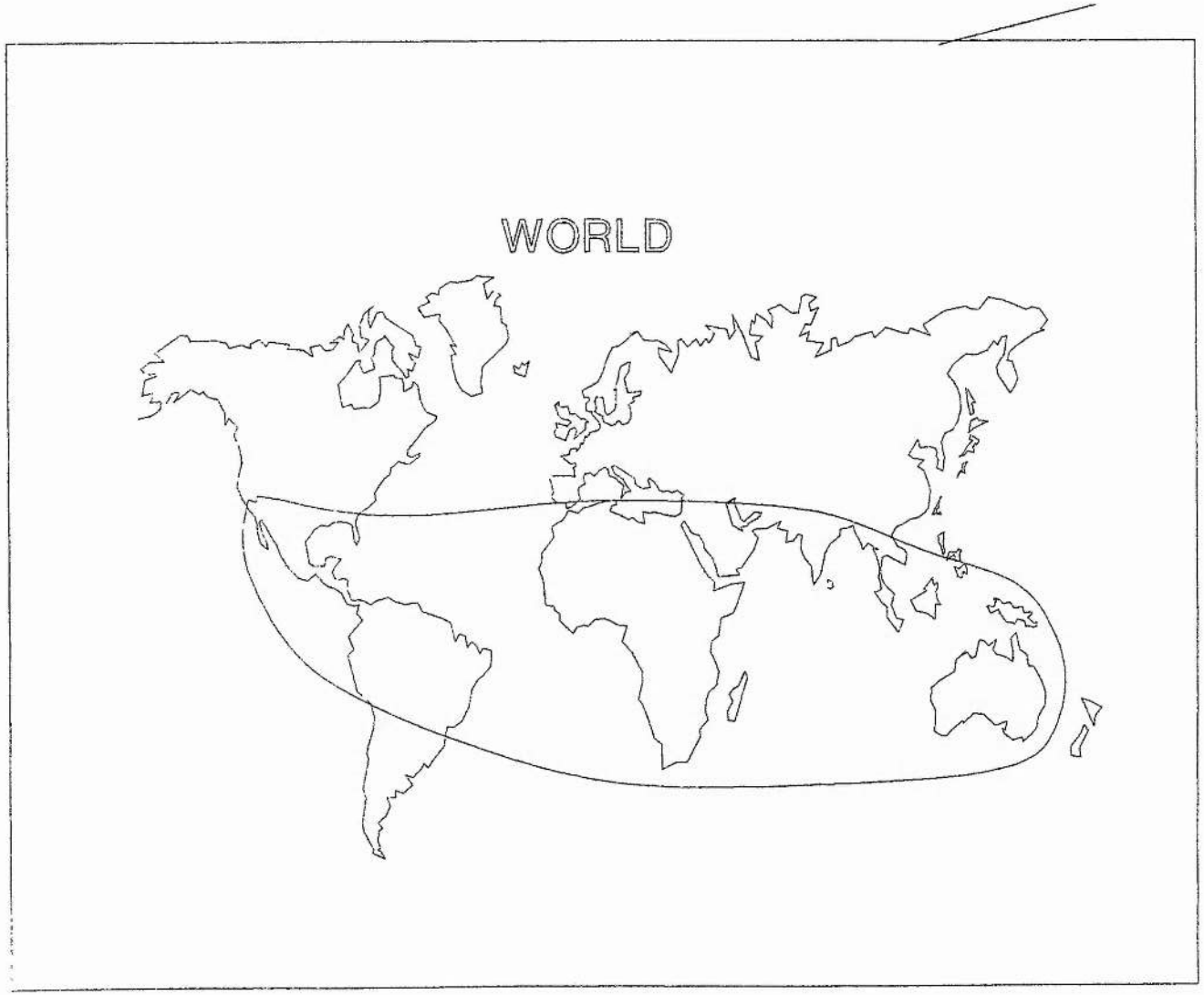
Chapter 1
Introduction to *Acacia*.

The genus *Acacia* MILLER, one of the largest within the Leguminosae, is a widespread genus of tropical-subtropical trees. It ranges from Central/South America through most of Africa, to south-east Asia and Australia (see map 1.1). There are at present approximately 1,200 described species of *Acacia*, but this number will probably increase when the Flora of Australia is revised. The majority of these species are endemic to Australia, where the genus has diversified considerably.

1.1 General Background.

Acacia species can be found in many savannah ecosystems throughout the world. Although they play an important and often dominant role in many of these ecosystems little research has been undertaken on the applied ecology of the genus, apart from those species which are agronomically important.

Acacias are predominantly trees or shrubs associated with temperate or tropical savannah regions where they have a tendency to exploit arid or semi-arid areas rather than densely forested areas. This is probably due to their low tolerance of low light intensities (Ross, 1981). Some species can grow in forested areas, e.g. *A. kraussiana*, although these appear to be obligate climbers (Ross, 1981) and so maintain an emergent position in the canopy. Acacias which occupy the more typical arid sites nevertheless occupy a wide range of habitats e.g. from *A. nubica* on the bottom of the rift valley, to the frost hardy *A. erioloba* which can grow at altitudes of up to 1900m (6250 ft.). Indeed, there appears to be a high level of specialisation in Acacias with the majority of species having a clearly defined ecological range imposed by



Map 1.1. A map of the approximate distribution of the genus *Acacia*.

edaphic and climatic factors (New, 1984). An example of this is *A. seyal*. Smith (1949) investigated the relationship between the clay content of the soil and the amount of rainfall on sites where *A. seyal* occurred. He found that *A. seyal* only occurred on high clay content soils if there was a correspondingly high rainfall. Smith (1949) concluded that runoff of rainwater was greater on clay soils than on sandy soils, so less water would be stored in a clay soil. *A. seyal*, unlike many other African Acacias which avoid drought by having long tap roots, is only shallow rooted and appears to be limited by the water-holding capabilities of the soil. Climatic conditions which limit the distribution are usually seasonal events such as the occurrence of winter or summer rains. The avoidance of desiccation appears to be important. To this end *Acacia* trees employ several different strategies. *A. willardiana* in the Sonoran desert has very small leaves, which are reduced to 1-3 pinnae (Shreve, 1951) which would appear to be an adaptation to reduce transpiration. In the same area *A. greggii* is winter deciduous (New, 1984), winter being the driest time of the year. Many other species are also facultatively deciduous.

Although *Acacia* is undoubtedly important in the savannah ecology very little research has been carried out concerning the functional ecology of such areas and little is known about how *Acacia* interacts with the whole ecosystem. One example of the interaction can be seen in the study of Adams (1967) who investigated a well defined *Acacia*-grass cycle in Sudan. In some areas *A. mellifera* forms thickets which are to some extent fireproof, while in surrounding areas the grass present is readily burnt. Adams (1967) found that in a year when a fire occurred the establishment of *A. mellifera* was enhanced by the loss of the grass. However, if rains followed the burn, or if the burn occurred more than once, then the areas were re-invaded by grass. The reasons for this were that the rains accelerated the growth rate of the grass enhancing the competitive effect. The additional burn weakened the *A.*

mellifera and minimised reproduction giving the grasses a change to re-invade.

Likewise little research has taken place concerning the interactions between animals and Acacias. In Africa Acacias provide a valuable food source for many native grazing animals as well as for domesticated animals. The influence of grazing may have a considerable effect if it is severe, and so this has the potential to alter the structure of the vegetation interactions (Werger, 1977, Versey-Fitzgerald, 1974). A well known example of this is the effect grazing elephants have on *A. xanthophloea*. The elephants push smaller trees over to access their foliage and pods. They also debark older trees causing death. Both of these processes affect the surrounding vegetation.

As with many other species, Acacias have been introduced into areas outside their natural range, many of these introductions have led to the introduced species becoming a 'weed'. The majority of these introductions have been between Africa and Australia with introductions occurring both ways. Examples of both scenarios are well documented. *A. mearnsii*, an Australian species was introduced into South Africa for use in the tan bark industry. Inevitably, it escaped from cultivation and is now a serious pest along river banks (van der Jooste, 1965). In the other direction, *A. nilotica* was introduced from Africa into Australia probably as a fodder source for sheep. When farming practices subsequently changed, and the sheep were replaced with cattle, *A. nilotica* became a pest in areas in which it was introduced. The cattle did not graze on *A. nilotica* as the sheep had done and consequently the number of *A. nilotica* trees increased (pers. comm., C. Fagg, 1994).

In other cases, *Acacia* species have been introduced expressly for the purpose of sand dune stabilisation. Although many local species of *Acacia* would be adequate, exotic species are at present often used for stabilisation.

Two species that are commonly used are *A. saligna* and *A. cyclops*, both from Australia. Sale (1948) investigated the use of *A. saligna* in Palestine. He noted that in areas planted with *A. saligna* the sand was stabilised within three years. After 6-7 years *A. saligna* had formed thickets and under these a litter layer was beginning to build up.

Although a few species are able to regenerate through root or stem suckers, e.g. *A. albida* (syn. *Faidherbia albida*), most species rely on sexual reproduction. The African Acacias produce large number of flowers, e.g. *A. tortilis* has been recorded to have ~400 inflorescences per metre of flowering twig (Fagg and Barnes 1990), but have a low pod/flower ratio (Ross, 1979). This is seen in the large number of flowers which abort, ~90%. It is thought that this is an adaptation to attract insects to individually small flowers for pollination (Ross, 1979). The pollen vector for Acacias is not specific, pollination being secured by a wide range of insects and small vertebrates, although bee pollination seems to be prevalent (Olng'otie, 1992). Details concerning the breeding systems are sketchy. Their pollination strategies do however suggest that they are strong outbreeders (Olng'otie, 1992). These strategies include protogynous dichogamy, andromonoecy and self-incompatibility.

Seed dispersal is unspecialised with seed either dropping to the ground or being ejected from drying dehiscent pods. Animals play an important role in dispersing the seed away from the parent. The seeds are eaten by many grazing animals e.g. in Africa by elephant, gazelle, etc. who ingest large numbers of pods when they are in season. The seeds pass through the animal's gut and are deposited in the faeces. Apart from the additional nutrient this provides, it appears that germination time is decreased and germination rate is increased when the seed travels through an animal (New, 1984). In addition, damage to the seed by bruchid larvae, a major pest of *Acacia*, is decreased (New, 1984).

1.2 Man and Acacia.

Many of the species within this genus have agronomic importance in rural areas of Africa, and in other countries where *Acacia* occurs. There are, however, two commercially important products derived from *Acacia*, gum arabic and tan bark. Gum arabic comes mainly from one species, *A. senegal*. This is a sap exudate from the bark of the tree. About 90% (Fagg and Barnes 1990) of the commercial gum arabic production comes from *A. senegal*, with most of the remaining 10% coming from related *Acacia* species. The gum has a large number of uses, it can be found in many confectionery products and medicines.

The high level of tannins in the bark of many species of *Acacia* has led to their world-wide use in tanning. The most commonly used species on a commercial scale is *A. mearnsii* which has been introduced from its native Australia into many other countries. However, this use of *Acacia* is in decline, due to use of synthetic chemicals in the tanning of leather and an increase in the use of artificial materials such as plastic.

The other species tend to be used on a smaller scale. The ranges of uses is large, but can be loosely divided into two broad categories - agricultural and domestic.

1.3 Agricultural Uses.

Acacia trees fit in well with the pastoral system of livestock rearing. The leaves constitute a staple browse for goats, camels and cattle. The pods, when produced, are highly nutritious and are an important supplement for livestock during the dry season. The flowers also provide a major food item for livestock, as over 90% abort and drop to the ground where they are eaten (Kayongo Male and Field, 1983). The pollen and nectar production of the flower also provides an abundant resource for bees. Many rural populations

keep bees for honey production. In some areas this is now a cash 'crop' imported into Europe.

Several species are used in agroforestry schemes e.g. *A. senegal*. In some areas crops are cultivated under young *A. senegal* for approximately five years. The area is then left fallow for the next fifteen years while the trees are tapped for gum and used as browse. At the end of this rotation the trees are coppiced and the cycle begins again. There are also extensive data which suggest that many crops, such as sorghum, millet and groundnuts, grown under *Acacia* trees have increased yields (Woods, 1989, see plate 1.1 following page 13). This is probably due to the nitrogen fixing capabilities of *Acacia* root nodules and to shading of the crop during hot and dry periods.

Many species are drought resistant and are able to grow in near-desert conditions. These species give promising prospects for land reclamation in the Sahel and other arid areas.

1.4 Domestic Uses.

In many areas of Africa firewood is in short supply, *Acacia* wood can provide an important source of quality firewood. In some rural areas the wood is burned into charcoal which is then sold on to nearby urban populations.

As well as the pods providing a livestock food, they are also suitable for human consumption. The bark and gum are used as traditional medicines, some of the pharmacologically active compounds from these medicines have been identified (e.g. Hagos et al. 1987).

We can see from the above list of uses that *Acacia* species have earned their forestry designation as multi-purpose trees. There is currently a great deal of research into such species, which provide an important resource for rural populations.

1.5 Development of the project.

In this study I initially concentrated on the African species of the genus. This was linked with an Overseas Development Administration (ODA) and Oxford Forestry Institute (OFI) research scheme (R.4348) concerned with the study and acquisition of genetic resources of African Acacias. The ODA project was initiated because of a perceived need for the improved use of the genetic resources of African Acacias. To this end seed was collected from four prospectively important species (*A. tortilis, A. nilotica, A. senegal* and *A.* (syn. *Faidherbia*) *albida*) and several other potentially important species. My initial remit was to study the genetic diversity of African Acacias using the seed collections of the OFI and ODA. Simmons (1991) believes that one of the major problems facing the improvement of multipurpose trees, such as *Acacia*, is that very little is known about the amount and structure of genetic variation of such species. It is important that details about the genetic variation are investigated as this is will give a sound basis for any attempts to improve these multipurpose trees. This was to be accomplished using cpDNA restriction site mapping primarily, and other methods, such as RAPDs, if necessary.

However, as I researched the taxonomy of the genus it became apparent that there was no consensus concerning the taxonomy of the African species, or on their inter-relationships with each other or within the wider spectrum of the genus as a whole. The relationships of the genus to other taxa within the subfamily Mimosoideae and its evolution were also in dispute. Before the genetic diversity of the African species of *Acacia* could be investigated, it would be necessary to clarify their relationships to each other. Their relationships could only be clarified as part of an investigation of the relationships between taxa in the tribes Mimoseae, Ingeae and Acacieae. The methodology which was to be used to investigate the genetic diversity of African Acacias was also applicable to the investigation of the relationships of

African Acacias and their relationships to other *Acacia* species and other taxa in the Mimoseae and Ingeae. This became the main thrust of the thesis.

1.6 Background to the morphological characters which formed the basis of previous classifications.

Historically the major characters used to classify Acacias have been the gross morphology of the inflorescence and stipules. To some extent sub-generic divisions based on these appear to be natural, and reflect the partitioning of the African Acacias into two groups. Other more cryptic characters such as seedling ontogeny, pollen morphology, phytochemicals and caryology can also be useful in classification of higher level groupings within the genus. Other morphological characters such as pod characteristics, flower colour, habit, position of extrafloral nectaries, inflorescence arrangement and bark colour are only really important in distinguishing between species.

1.6.1 Inflorescence

The inflorescence shape is a very convenient character for dividing *Acacia* into two groups. In most species it is either capitate (see plate 1.2, following page 13) or spicate (see plate 1.3 and 1.4). However, there are two exceptions, i.e. *A. dolichocephala* and *A. mellifera* subsp. *detinens* whose inflorescences are subcapitate (see plate 1.5). If one looks closely at the length of the inflorescence in different species it is apparent that there is a ranges of sizes. For example spicate inflorescences cover those of *A. lahai* where they can reach up to 7 cm., and of *A. moggii* (see plate 1.3) whose inflorescence is only 1.2-1.8 cm. long though still in a distinctly spicate form. The varying character rather than the shape of the inflorescence therefore appears to be the flower axis length. The length of the axis is variable, but can

to divided into two groups; those with a very short axis and those with a longer axis.

1.6.2 Stipules, Spines and Prickles.

The presence of spinescent or non-spinescent stipules provides another useful way of dividing *Acacia*. Spinescent stipules are often referred to as spines, they are sharp pointed, hardened, modified stipules and have a vascular supply that is continuous with that of the branch. Those species with non-spinescent stipules have developed a different form of armament. These are sharp pointed epidermal outgrowths, known as prickles, and have no vascular tissue. The two forms of defence are mutually exclusive, i.e. species with non-spinescent stipules have prickles and species with spinescent stipules (spines) never have prickles. Both spines and prickles are permanent and not deciduous, whereas in some non-spinescent species the stipules can be deciduous.

There is a wide range of forms that both spines and prickles can take. Those species with prickles can be subdivided into three groups according to the position of the prickles. The prickles can occur all over the stem (e.g. *A. ataxacantha*, see plate 1.6), or in threes limited to the nodes (e.g. *A. senegal*), or in pairs limited to the nodes (e.g. *A. mellifera*). Stipular spines always occur at the nodes but the shape and size of the spines can be important. For example the spines of *A. seyal* var. *fistula* (see plate 1.7) can reach up to 8 cm. and they are often basally fused to form an 'ant gall', *A. karroo* has been recorded with spines up to 25 cm. in length! The spines on *A. tortilis* are often recurved and look superficially like a prickle, but 'true' spinescent stipules can always found on the same plant, thus clarifying their true nature.

1.6.3 Correlations between armament type and inflorescence shape.

The two previous characters, i.e. inflorescence shape and stipule spinescence are to some degree correlated. In African species of the genus the majority of species with spicate inflorescences are armed with prickles and have non-spinescent stipules whilst the majority of species with capitate inflorescences have spinescent stipules. There are, however, a number of species which run counter to this trend and form natural groups, e.g. *A. horrida*, *A. bussei* and *A. lahai* (see plate 1.4), all have spicate inflorescences and spinescent stipules.

These two characters define a natural split in the African species and in the genus as a whole. However since there is an overlap between the two, one of these characters must be secondarily evolved, i.e. one of these characters has arisen at least twice during the evolution of *Acacia*. The evidence suggests that it is the nature of the stipules that naturally separates the two groups as this character is correlated with several other morphological features which divide the genus. For example, all the species of *Acacia* in Africa with spinescent stipules have an involucel on the peduncle regardless of their inflorescence shape. The evidence is reviewed by Ross (1979).

1.6.4 Habit.

All the species from Africa are woody and grow as shrubs, scandescent shrubs, climbers or trees. They range in size from ground-hugging species such as *A. edgeworthii* to trees up to 40m high, e.g. *A. galpinii*. Many species have distinctive growth forms which aid identification, e.g. *A. senegal* var. *leiorachis* is a tall spindly tree whereas other varieties of *A. senegal* tend to be small shrubs or trees with rounded crowns.

1.6.5 Bark.

This is another character which varies considerably amongst African Acacias and is useful in species verification. For example the unique greenish yellow bark of *A. xanthophloea* distinguishes it from all the other African species. Likewise *A. hockii* is distinguishable from *A. seyal*, a closely related species, by its papery bark which peels off.

1.6.6 Flower Colour.

Flower colour is taxonomically important (Ross, 1979). Two main groups of flower colour are found in Africa, pale-yellow to white and bright yellow. The colour of the inflorescence is constant for each species.

1.6.7 Inflorescence Form.

The inflorescence is usually either racemose or paniculate. This is also an important taxonomic character. Robbertse (1974) reviewed the evolution of the inflorescence in Acacia. He concluded that was possible to group the South African species of Acacia into 11 natural groupings according to their inflorescence. These groups differentiated species according to their inflorescence shape, inflorescence arrangement (paniculate or racemose) and the nature of the stipules. In addition, in species with non-spinescent stipules the position of the prickles was used to separate groups of species, e.g. Robbertse distinguished *A. polycantha* from *A. ataxacantha* by the occurrence of prickles all along the stem of *A. ataxacantha*, while the prickles on *A. polycantha* are limited to the nodes.

Robbertse viewed the evolution of the inflorescence as follows. The progenitor of *Acacia* had non-spinescent stipules and no prickles, with spicate flowers on a paniculate inflorescence. The progenitor gave rise to two groups in Africa with the evolution of either spinescent stipules or prickles. In both these groups changes to the inflorescence have been similar. The

inflorescence form has been reduced, first losing the secondary axes and then losing the primary axes to form racemes from the leaf axils. Also in both groups species have had the flowering axis reduced to form globose inflorescences.

1.6.8 'Ant-galls'.

In several African species the stipular spines are swollen into what are called 'ant galls' (see plate 1.8). There is no evidence that ants cause the swollen spines (Ross, 1979), it seems that the ants merely take advantage of these galls. In many species which display the galls, such as *A. seyal* var. *fistula*, no ants have been found. However in one species, *A. drepanolobium*, it appears that a mutualistic association occurs. There are also many species in the New World where associations between ants and Acacias occur, and these perhaps are the most studied mutualistic associations between insects and plants. The topic of ants and Acacias is covered by Janzen (1966, 1974).

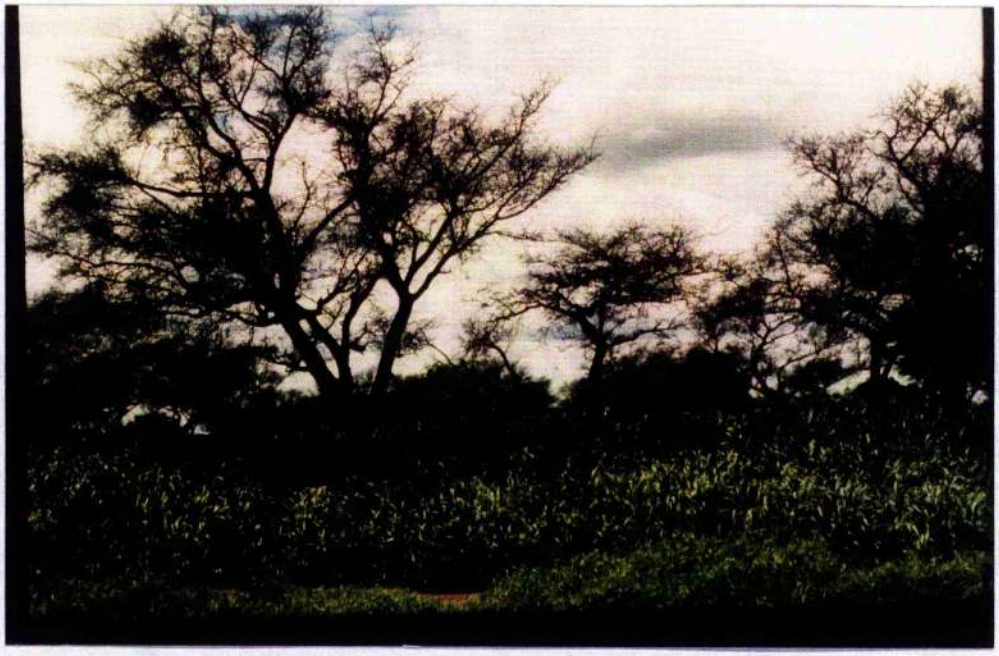


Plate 1.1 ▲. *Acacia albida* over Sorghum in Burkino Faso. Photo from C.Fagg OFI.

Plate 1.2 ▼. The capitate inflorescences of *Acacia karroo*, Zimbabwe Photo from C.Fagg OFI.

Plate 1.3 ▲. The spicate inflorescences of *Acacia moggii*, Somalia. Photo from C.Fagg OFI.

Plate 1.4 ▼. The spicate inflorescences of *Acacia lahai*, Kenya. Note the combination of spicate inflorescences and spinescent stipules.

Plate 1.5 ▲. The subglobose inflorescences of *Acacia dolichocephala*, Kenya.

Plate 1.6 ▼. The stem of *Acacia ataxacantha*, showing the distribution of prickles along its stem.

Plate 1.7 ▲. Acacia *seyal* var. *fistula*, Malawi. The bases of the stipules are fused to form 'ant galls'. Photo from C.Fagg OFI.

Plate 1.8 ▼. *Acacia drepanolobium*, Kenya. The swollen stipule in the center of the picture is an ant gall, which inside is hollow. Photo from C.Fagg OFI.

Chapter 2
An introduction to the taxonomy of the genus *Acacia*

"It hath a tubulous flower consisting of one leaf, with many stamina or threads, which are many of them collected into a kind of sphere or globe: the pointal of the flower afterwards becomes a pod in which are included several seeds, each of which is separated by transverse diaphragms, and are generally surrounded with a sweetish pulp."

2.1 Early classifications of *Acacia*.

The quotation above is part of Philip Miller's generic description for *Acacia*, based on the "Egyptian Thorn".[1] It appeared in his Gardeners Dictionary (1754). Miller was the first person, after 1st May 1753, to use the name *Acacia* in a generic sense, and is considered the author of *Acacia* (Ross, 1973). I believe before the phylogeny of *Acacia* is investigated that the historical aspects of its taxonomy should be examined. This is what I intend to accomplish in this chapter.

Following the generic description, Miller recognised 24 species of *Acacia*, although he did not attempt to divide these species into subgeneric groups. Many of the species which Miller described as *Acacia* are not now referable to *Acacia* as the generic limits imposed by Miller were very vague. The same is true for many other genera in the Mimosoideae, i.e. the generic limits first established were loose and consequently there has been a great deal of re-allocation of species.

[1] Miller did not give specific names to the taxa he included in his genus Acacia, but instead used common names (pers. com. C.Fagg OFI). Egyptian Thorn is referable to the species *Acacia nilotica* (L.) WILLD. EX DEL.

Lamarck (1783; cited in Ross, 1979) listed 58 species under *Acacia*, dividing these species into two groups. The groupings were based on armature of the species, i.e. whether the stipules were spinescent or not. He did not apply any rank to these two groupings.

Willdenow (1806; cited in Ross, 1979) listed 102 species of *Acacia* recognising seven groups based on vegetative characters. Again these groups were not given any rank.

De Candolle (1825; cited in Ross, 1979) divided the 258 species he listed into four sections, which were not named, based mainly on leaf characters with the nature of the stipules and inflorescence delimiting species within the sections.

It was not until Bentham published his series of papers (1842, 1844, 1845, 1846, and 1875) that some sort of order was imposed on the genus *Acacia* and the subfamily Mimosoideae in general. His descriptions became the 'benchmarks' for many of the genera he investigated. Bentham for the first time clearly defined the generic limits of *Acacia*, and excluded species he did not think belonged to the genus. Prior to this the limits were ill-defined, and it was a heterogeneous assemblage of plants.

Bentham (1841) adopted the tribe Acacieae REICHENBACH as one of the three tribes constituting the suborder Mimosoideae. Initially (1842a, 1842b) included with the genus *Acacia*, in the tribe Acacieae, were the genera *Albizzia*, *Calliandra*, *Lysiloma*, *Zygia*, *Enterolobium*, *Pithecellobium*, *Serianthes*, *Inga* and *Affonsea*. Bentham (1865) subsequently made substantial changes to the tribe by restricting it to just *Acacia* and creating the tribe Ingeae BENTH. to accommodate the other nine genera. This was maintained for the final revision of subfamily Mimosoideae (Bentham, 1875). This revision has been the basis for all subsequent revisions.

Bentham divided *Acacia* into six series, these being defined by several characters. Primarily the vegetative characters, foliage and spinescence

were used. Inflorescence played a less important role, and fruit characters were largely ignored in the division of the genus.

Bentham's classification met with later criticism. For example, Newman (1932) considered Bentham's classification to be "too static" in concept. However, two things must be borne in mind: the genus now includes twice as many species as considered by Bentham and many of Bentham's decisions were, by necessity, based on specimens which by modern standards would be considered inadequate (Ross, 1973). However, Bentham's classification has stood the test of time. No comprehensive attempts to re-group all species in the genus were attempted for nearly a century. Even recent classifications have supported Bentham's major subdivisions of the genus.

Britton and Rose (1928), in the North American Flora, divided the American members of the genus into many new genera, reflecting the discrete partitioning of characters amongst Bentham's series. This classification was primarily based on fruit characters, with vegetative and inflorescence characters discriminating at lower levels. Although Britton and Rose contended that the genus was too diverse and difficult to handle, their 'splitting' of the genus into twelve separate genera was not taken up by other workers.

2.2 Classifications after Bentham.

2.2.1 Guinet and Vassal

The first major challenge to Bentham's classification was that of Vassal (1972). Using pollen data from Guinet (1969) and supplementing it with his own extra characters derived mainly from seedling ontogeny, seedling and stipule characters, Vassal produced what he considered to be a phylogenetic classification. In this he recognised three subgenera; *Acacia*, *Aculeiferum* and

Heterophyllum (= *Phyllodineae fide* Ross, 1981[2]). In addition he removed *Acacia albida*, an aberrant African species, to the monotypic genus *Faidherbia* CHEV., although keeping it within the tribe Acacieae.

Shortly after this Guinet and Vassal (1978) summarised their current knowledge of *Acacia* and "attempted to define precisely the subdivisions of the genus, and determine their degree of relationship". This paper was the culmination of their previous work, and the following discussion concentrates on their conclusions in this paper.

The characters they looked at were similar to those described in previous papers, but with additional data from species not included in the former papers. The traits displayed in each of the characters were categorised as 'unspecialised', 'specialised' and 'highly specialised'. The character state assessment was then grafted on to the subgenera of Vassal (1972) and series of Bentham (1875), and trends in the specialisation of characters were compared and contrasted between the subgenera or series

Guinet and Vassal's conclusions were as follows:-

1) On the basis of pollen morphology, subgenus *Acacia* was the most specialised and subgenus *Aculeiferum* the least. Subgenus *Phyllodineae* shared a number of important characters with subgenus *Aculeiferum*, such as no columellae and simple apertures, but for most characters subgenus *Phyllodineae* was more specialised.

2) Subgenus *Acacia* could be distinguished from the other two subgenera by its highly specialised chromosome characters, such as the range of ploidy levels found in the group (diploid to 16-ploid), and the low level of karyotypic homogeneity found in some members. The other two genera are nearly homogeneous. This again suggests that subgenus *Acacia* is the most specialised.

---

[2] The subgeneric name *Phyllodineae* (D.C.) SERINGE has priority and has been adopted in place of *Heterophyllum* VASSAL. In this thesis only the subgeneric name *Phyllodineae* will be used.

3) The seeds of subgenus *Acacia* are often highly specialised. Guinet and Vassal viewed having 2-3 seeds in a row, in an indehiscent pod, as a specialised character of subgenus *Acacia*.

4) It was rare for the cotyledon and leaves to be specialised in subgenera *Acacia* and *Aculeiferum* but in subgenus *Phyllodineae* there were many specialisations, such as the development of phyllodes.

5) Subgenus *Acacia* had many unique inflorescence characters.

Guinet and Vassal noted that if the correlation between characters and subgenera were true then subgenera *Aculeiferum* and *Phyllodineae* were more closely related than either was to subgenus *Acacia*, despite the fact that subgenus *Acacia* and subgenus *Aculeiferum* share a common geographical distribution.

As to the origin of the genus; Guinet and Vassal postulated an origin in West Gondwanaland, approximately Mexico to Bolivia today (Guinet and Vassal, 1978). The reasoning for this suggestion was that the species "groups" which are represented in the American continent have a preponderance of 'unspecialised' character states. In addition characters which are absent from species in America can be found in indirectly related genera in America. The two examples they use to illustrate this point are extraporate pollen and phyllodes. The pollen character (extraporate apertures) is "fundamentally" Australian, but also exists in some genera closely related to *Piptadenia* (Guinet, 1967). Phyllodes are unique to some species in subgenus *Phyllodineae*, but they can be found in several American species of *Mimosa* (Burkart, 1952). These points convinced them that the American continent contained most of the "evolutionary potentialities" which are now found in the genus *Acacia*.

A further argument used by Guinet and Vassal to support a Central American origin for the genus was that "specialisation of characters occurs ... when one moves away from America". The examples they used to illustrate

this argument were; the development of polyploidy; an increase in the size of pollen grains; the appearance of arillate funicles and the presence of prickles.

Several authors have suggested that the development of polyploidy and a reduction in length of total chromatin often indicates distance from the centre of differentiation of a group (cf. Ehrendorfer, 1954; Moore, 1968). A tendency towards asymmetrical karyotypes accompanies both these characters in *Acacia*, and this is a trend which according to Stebbins (1971, 1974) occurs within numerous genera and indicates specialisation.

The size of the pollen (probably in relation to polyploidy) increases away from the hypothesised point of origin. This increase is quite evident on all continents with increasing distance away from the equator.

In subgenus *Aculeiferum* and subgenus *Phyllodineae* the appearance of arillate funicles is also consistent with Guinet and Vassal's hypothesis e.g. in subgenus *Aculeiferum* there is an absence of an aril in America, and a tendency towards one in section *Aculeiferum* in Africa and Asia. In subgenus *Phyllodineae*, the arillate funicle is very frequent and often the aril is complex.

The final example is the appearance of prickles. In the two sections of subgenus *Aculeiferum* found in America, none of the species in section *Filicinae* have prickles and only a few species from section *Monocanthea* have prickles. However, in species from all sections found in Africa and Asia prickles are generally present, scattered or localised on the stem.

These facts support Guinet and Vassal's hypothesis (1978) that the expansion of the genus in the American continent occurred within a relatively limited area, roughly comparable to that occupied now by section *Filicinae* (Mexico to Bolivia).

The main criticism of Guinet and Vassal's work has centred on how they decided the polarity of characters. They were also criticised for the apparently arbitrary way in which they divided a continuous range of variation into discrete classes and then assigned one of these classes as unspecialised.

An example of this, highlighted by Pedley (1986), concerned seed characters. Guinet and Vassal designated seeds <5 mm as small and unspecialised; 5-10 mm as specialised and seeds >10 mm as highly specialised. Why were the seed sizes split up into these groupings? No explanation was offered. Was there continuous variation of seed size? Or perhaps seed size was split up into more or less discrete classes. There was also no reason to presuppose that small seeds were unspecialised. It is easy to imagine a scenario where small seeds would be adaptive e.g. if seed predation by insects was high then small seeds could confer a selective advantage.

However, in spite of the shortcomings listed above, Guinet and Vassal produced a very workable phylogenetic hypothesis for the genus, and this was used as a basis for the next re-examination of the taxonomy of *Acacia* by Pedley (1986).

2.2.2 Pedley.

In the same year as Guinet and Vassal's paper, Pedley published a revision of *Acacia* species in Queensland, Australia (Pedley, 1978). He also recognised three subgenera, modelled on Bentham's series. Although this classification was not presented as a "natural scheme", it was a good attempt at incorporating the best aspects of the Bentham and Vassal classifications into a single usable scheme (Maslin, 1989). Pedley (1987) revealed that even by this time (1978), he was convinced that subgenus *Acacia* warranted recognition as a distinct genus. He did not adopt this for his 1978 classification, although Pedley (1981) informally suggested that *Acacia* should be divided into two genera, *Acacia* and the genus '*Z*'. i.e. *Zigmaloba* RAFINESQUE., which would have incorporated subgenera *Aculeiferum* and *Phyllodineae*. Pedley did, however, admit that "whether or not *Acacia* should be split is partly a philosophical question".

Re-evaluating the data from his 1978 classification as well as collecting and collating new data, Pedley (1986) published what he considered to be a phylogenetic treatment of the genus. He used a wide variety of new characters as well as data already published. The characters used included; the morphology of seedlings, leaves, flowers and inflorescence; anatomy of pod; the occurrence of extra-floral nectaries; free amino acids of the seeds; flavonoid compounds in heartwoods; cyanogenic compounds; palynological characters and susceptibility to rusts (Pedley, 1986 and refs. therein). He decided that three genera should be recognised; *Acacia* MILLER (= *Acacia* subgenus *Acacia*), *Senegalia* RAFINESQUE (= *Acacia* subgenus *Aculeiferum*) and *Racosperma* (D.C.) MARTIUS (= *Acacia* subgenus *Phyllodineae*). These new genera corresponded with the previous subgeneric divisions within *Acacia* (Vassal, 1972, Pedley, 1981).

Use of the name *Acacia* thus becomes slightly confusing. Is it being used in the sense of Vassal (1972) as a generic name for all three subgenera, or is being used in the sense of Pedley for just *Acacia* subgenus *Acacia*? To overcome this problem in the present work Pedley's use of the generic name *Acacia* will be written as *Acacia sensu stricto* (*Acacia sens. str.*)

Pedley was aware of the implications of splitting the genus, "which would affect a large number of botanists throughout the tropics and subtropics"(Pedley, 1986). He set himself two tasks; 1) to decide whether *Acacia* should be treated as three separate taxa and 2) to decide whether these taxa should be treated as genera rather than subgenera or sections.

The first of these tasks was easy. Since Vassal(1972) first distinguished three subgenera using pollen and other characters, the distinctiveness of each subgenus has been confirmed as more data have accumulated.

At what level the three taxa within *Acacia* should be recognised was a greater problem to solve. The reasoning which prompted Pedley to recognise three genera was as follows.

Firstly he considered the 'wider picture', i.e. what was the current trend in similar taxonomic situations in other genera? The recognition of narrowly circumscribed genera is a modern tendency in large and economically important families. Let us not forget that there are at present over 1000 described species in the genus *Acacia*, and in all likelihood this figure will continue to increase. Pedley briefly mentioned ongoing projects to fragment other large groups e.g. the grasses, Papilionoideae, Caesalpinioideae, *Casuarina* and *Eucalyptus*.

Secondly, in an attempt to quantify whether *Acacia* is broadly or narrowly conceived he applied the Index of Diversity (α) of Williams (1964) to the tribe Acacieae as well as several other taxa from the Leguminosae. To summarise the results of this investigation; a value of α=11 was found to be "modal" for the Mimosoideae and in the tribe Acacieae (*Acacia* with its three subgenera and the genus *Faidherbia*) a value of α=0.25 was obtained. A high value of α indicates that genera are narrowly conceived. Such a value may indicate either that the family is in fact diverse, or that taxonomists have taken a narrow view of the genera.

These results suggested to Pedley that the genus *Acacia* , as defined by Guinet and Vassal, was too broadly conceived. If, however, Pedley's proposed classification was used instead, a value of α=0.5 was obtained; double the previous figure. If Pedley's proposals were taken up then the size of his more narrowly circumscribed genera would be "in keeping with the size of genera of other tribes of low diversity in Leguminosae".

Pedley went on to consider the evolutionary trends of the genus. Like Guinet and Vassal (1978) Pedley concluded that subgenus *Aculeiferum* and subgenus *Phyllodineae* were closely related, both being clearly distinguished from subgenus *Acacia*. *Acacia sens. str.* had a distinctive pattern of non-protein amino acids in its seeds, colporate pollen, stipular spines and involucels.

Unlike Guinet and Vassal who did not specify the relationships of the tribe to other tribes and genera within the Mimosoideae, Pedley suggested taxa which he thought were related to his genera. "*Acacia sens. str.* has some affinity with *Pithecellobium* (Ingeae), and *Senegalia* and *Racosperma* have affinities with *Calliandra* (Ingeae)". "Direct derivation of *Senegalia* from *Acacia s.s.* or the reverse is unlikely; derivation of *Acacia sens. str.* and *Senegalia-Racosperma* from within different lines within the Ingeae is suggested". Although not explicitly stated, this statement means that the tribe Acacieae is polyphyletic. No data were presented to explain the reasoning behind suggesting *Pithecellobium* and *Calliandra* as relatives.

For both of these reasons, Pedley decided to recognise three genera. He concluded that the genus as defined by Guinet and Vassal(1978) is polyphyletic and too broadly conceived.

2.3 Reactions to Pedley's classification.

Pedley's proposals were greeted unenthusiastically by the majority of workers in Legume systematics. In Berlin 1987, The International Group for the Study of the Mimosoideae (I.G.S.M.) met to discuss Pedley's proposal for the recognition of three genera. The majority view of the meeting was that to recognise three genera within *Acacia* as proposed by Pedley was premature (Maslin, 1987). The meeting prompted Maslin (1988) to review Pedley's (1986) proposals and the evidence supporting them in a paper entitled "Should *Acacia* be divided?", or, "Wattle become of *Acacia*?"

Maslin first reviewed Pedley's evidence for recognising three genera within *Acacia*, specifying the characters he considered Pedley found important; i.e. pollen, free amino acids of the seeds, stipular spines, phyllodes and the flowering system. He detailed Pedley's interpretation of the data and then summarised the evidence from his own perspective.

In his discussion he formulated five questions that when answered would decide the best course of action. These were: 1) is *Acacia* polyphyletic? 2) does the available evidence permit the recognition of higher order taxa within *Acacia*, 3) if so how many taxa? 4) what is the appropriate rank and name for higher order taxa that are recognised? 5) is the name *Racosperma* validly published?

1) Is *Acacia* polyphyletic? Maslin stated that this question is fundamental to deciding whether or not *Acacia*, as currently defined, should be treated as more than one genus. This is a pertinent question. If *Acacia* is polyphyletic then more than one genus has to be recognised. If, however, *Acacia* is monophyletic then the argument for recognition of more than one genus becomes less obviously valid. As mentioned before there appear to be two phyletic assemblages, namely subgenus *Acacia* vs. subgenera *Aculeiferum* and *Phyllodineae*. The question is; do these two latter groupings share a common ancestor?

Maslin stated that "Pedley suggested that ... subgenus *Acacia* has affinities with *Calliandra* and *Pithecellobium* while the subgenera *Aculeiferum* and *Phyllodineae* have affinities with *Paraserianthes* (Ingeae)". In fact Pedley (1986) wrote that "*Acacia sens. str.* has some affinity with *Pithecellobium* , and *Senegalia* and *Racosperma* have affinities with *Calliandra* ." If we ignore Maslin's misquote the point being made is that Pedley suggested these 'affinities' without providing any evidence to support them, and therefore we cannot consider Pedley's suggested 'affinities' until evidence is presented. Guinet (1990) adds weight to this objection by Maslin by arguing that in spite of numerous similarities between *Acacia* and the Ingeae, pollen characters suggest that the genus (or genera) must instead be viewed as an early offshoot of the tribe Mimoseae.

Maslin suggested that Guinet (1990) supported the independent derivation of subgenus *Acacia* and the combined subgenera *Aculeiferum*-

Phyllodineae. While this is strictly true, it is a very narrow interpretation of Guinet's conclusion. Guinet (1990) felt that the individualisation of two main groups was supported by pollen morphology and clearly isolated subgenus *Acacia*. The absence of transitions between the two groups reflected a lack of a close relationship. Guinet's conclusions were that "the tribe Acacieae is a grade (Kanis, 1986) rather than a clade, and that accepting an origin of *Acacia* among different lines of the *Piptadenia* group (Mimoseae) would better explain the long disputed phyletic significance of many characters in *Acacia* when viewed as a natural unit with subgenus *Acacia* interpreted as either basic or derived". This is a much less dogmatic statement. Maslin also stated that Guinet described subgenus *Acacia* as sharing important pollen characters with *Calliandra sens. str*[3]. While this is also true, i.e. Guinet does list characters shared by the two taxa, Guinet goes on to say that "pollen differences between *Acacia* and *Calliandra sens. str.* are so numerous that they do not suggest (a) close relationship between the two genera" (my underlining).

Maslin (1988) also suggests that serological studies by Brain (1987) not available to Pedley (1986) further support notion of an origin for subgenus *Acacia* independent of that of subgenus *Aculeiferum* and subgenus *Phyllodineae* (Brain, 1987, quoted in Maslin, 1988).

Considering the above points and those of Pedley (1986), Maslin concluded that although "phylogenetic relationships have not been satisfactorily elucidated, present evidence suggests that *Acacia* is not monophyletic". Maslin thus concluded that there is no consensus of opinion on the phylogenetic relationships of genus *Acacia* or its subgenera save that subgenera *Aculeiferum* and *Phyllodineae* are closely related to each other and only distantly related to subgenus *Acacia*.

---

[3]The genus *Calliandra* is at present under reveiw, *Calliandra sens. str* refers to species of *Calliandra* with 8-grain polyads. That is the genus *Calliandra sens. lat.* with *Zapoteca* and Asian-Madagascan species of *Calliandra* excluded.

2) Does the available evidence permit the recognition of higher order taxa within *Acacia*? Maslin went on to discuss whether any higher order taxa should be recognised within *Acacia*. As mentioned before, two characters in particular, pollen morphology and free amino acids of the seeds have been interpreted by Pedley as "fundamental differences" between subgenus *Acacia* and the combined subgenera *Aculeiferum* and *Phyllodineae*. These characters appear to indicate that *Aculeiferum* and *Phyllodineae* are closely related. These groupings are further reinforced by data not available to Pedley (serology, Brain, 1987; cyanogenesis, Conn *et al.*, 1989). Thus, it seems that at least two higher order taxa could exist, as Pedley suggested in 1981, prior to recognising three genera in 1986.

However, Maslin (1988) went on to explain that current evidence suggested that within each of these genera/subgenera further work could well support further splitting, e.g. section *Filicinae* could be recognised as a separate genus. Pedley (1986) considered that section *Filicinae* 'could well be treated as a separate genus' but Guinet (1990) regarded it as distinct group with its affinities nearest to the tribe Mimoseae.

Because there is such uncertainty in the definition of the taxa in question and their phylogenetic relationships, Maslin believed that "informed taxonomic judgements" about the rank of taxa within genus *Acacia* could not be made.

3) What rank should be applied to any higher order groups within *Acacia*? This question has been partly answered, i.e. the uncertainty about the phylogenetic relationships of the taxa in question and the speculative suggestions that subgenera/ genera could be further split meant that suggesting ranks for these higher order groups would be premature.

4) What is the appropriate rank and name for higher order taxa that are recognised? Maslin pointed out that the answer to this question, of phylogenetic grouping and taxonomic rank, must also take into account the systematic placement of the tribe Acacieae within the Mimosoideae. Pedley

(1986) suggested that the Acacieae might be derived from different lines within the Ingeae and that the two tribes (Acacieae and Ingeae) 'should probably be united'. If Acacieae were combined with the Ingeae it is impossible to say how many genera would result. The rank criterion adopted in the Acacieae is different to that within the Ingeae, and a further complication is that the generic limits in the Ingeae are currently under review. Guinet (1990) as previously mentioned, viewed the Acacieae as an early offshoot of the Mimoseae. A consensus does not exist on the phylogeny of Acacieae, Ingeae and Mimoseae, and until such a consensus is achieved, it is "prudent to retain the current infrageneric rank for groups within *Acacia sens. lat.*" (Maslin, 1988)

Maslin's final question concerns the validity of the name *Racosperma* as used by Pedley. According to Maslin acceptance of the name *Racosperma* MARTIUS is debatable.

To summarise Maslin (1988) concluded that: 1) the evidence for dividing Acacia is inconclusive and/or incomplete. 2) Further studies are required to ascertain how many higher order taxa can be recognised within *Acacia sens. lat.* 3) The designation of rank to higher order taxa should be undertaken following resolution of the taxonomic status and affinities of the tribe Acacieae, relative to genera within the tribes Ingeae and Mimoseae. 4) Pedley's justification for recognising three genera was not convincing and 5) the validity of the name *Racosperma* is equivocal.

Maslin's conclusions are indicative of the general reaction to Pedley's paper, i.e. that his proposals were premature and that a great deal of investigation needed to be completed before any firm conclusion regarding the phylogeny and taxonomy of *Acacia* could be reached.

2.4 Additional pollen studies from Guinet.

Guinet (1990), as previously mentioned, also discussed the affinities of the genus *Acacia*. As Maslin had only a draft copy of this paper, it deserves further discussion.

Guinet surveyed the possible affinities of the genus *Acacia* using pollen characters in the course of a wider review of pollen characters in the tribes Ingeae and Mimoseae. It had previously been suggested that these taxa have affinities to *Acacia*.

Guinet's conclusions were that two main groups in *Acacia* were clearly distinguishable on the basis of pollen morphology. These two groups were *Acacia* subgenus *Acacia* and *Acacia* subgenera *Aculeiferum* and *Phyllodineae*. He did not feel that Pedley's generic distinction between *Senegalia* and *Racosperma* was borne out.

His conclusion concerning the generic position of *Acacia sens. lat.* has already been mentioned, i.e. in spite of numerous similarities with the Ingeae, pollen characters suggest that the genus (or genera) must be viewed as an early offshoot of the tribe Mimoseae.

2.5 Formulating a new approach to the classification of *Acacia*.

A shortcoming of the classifications previously discussed (Bentham, 1875; Guinet and Vassal, 1972; Pedley, 1986) is that they lack a rigorous framework supporting them. Although both phenetic and cladistic methodologies were probably available to Guinet and Vassal (1978) and definitely for Pedley (1986), neither study used these techniques, preferring to analyse the data subjectively. It is unlikely that either classification would be published today.

After his reply to Pedley (1986), Maslin became interested in the status of the tribe Acacieae, publishing a list of 'critical species' on which to build a comparative data set (Maslin and Stirton, in press). The justification for taking

this approach is the large number of taxa described for *Acacia*. It could be a lifetime's work to analyse every species. Instead Maslin and Stirton chose representative species from each of the major groups. Their aim was to list species which represented, as far as possible, the taxonomic variation within the genus. These would be the 'critical species' upon which analyses could be based.

2.6 A cladistic study of the Mimosoideae.

In 1993 Chappill and Maslin presented a paper at the 3rd International Legume Conference, Kew entitled 'A phylogenetic assessment of the tribe Acacieae' (Chappill and Maslin, 1995). In this paper they presented a cladistic analysis of the genera of the Mimosoideae including *Acacia*. Following a short discussion of previous classifications, they discussed the reasoning behind their analysis. They believed that "uncertainties exist not only regarding the tribal status of Acacieae but with respect to the classification and phylogenetic relationships of *Acacia*". The paper discussed their initiatives which were aimed at examining the composition and status of higher groups within Acacieae, and the relationship of these subgroups to others within the subfamily Mimosoideae.

Chappill and Maslin undertook two different analyses; a generic analysis of the Mimosoideae and an infra-generic analysis of the tribe Acacieae. In the generic analysis of the Mimosoideae each of the five groups of the Acacieae was treated as monophyletic and analysed as such. These groups were *Faidherbia*, *Acacia* subgenus *Acacia*, *Acacia* subgenus *Aculeiferum* section *Spiciflorae*, *Acacia* subgenus *Aculeiferum* section *Filicinae* and *Acacia* subgenus *Phyllodineae*. The classification used was that of Pedley (1978). For the infrageneric analysis each species/taxon was analysed individually. Both analyses were based upon morphological, phytochemical and pollen characters. The majority of the morphological data were gathered from

herbarium specimens and the other data were collected from various papers (see refs. in Chappill and Maslin, 1995).

The cladogram resulting from their cladistic analysis of the Mimosoideae can be seen in figure 2.1. Their results suggested that the tribe Acacieae was polyphyletic. The tribe could be seen to consist of three distinct groups; *Faidherbia*, *Acacia* subgenus *Acacia*, and *Acacia* subgenus *Aculeiferum* with subgenus *Phyllodineae*. Each of these groups appeared distinct. The first two were nested within the Ingeae and the third basal to the Ingeae. The results appear to support Pedley (1986) who suggested that the Acacieae and the Ingeae should be united. The characters used in Chappill and Maslin's study which support the uniting of Acacieae and Ingeae are; numerous stamens; dimorphic pollen grains with proximal and supplementary pores present; only eight polyads per anther and the presence of albizziine in the seeds.

Genera of the Ingeae to which the groups within the Acacieae appear to be related are suggested by the cladogram. *Acacia* subgenus *Acacia* is closely related to the genus *Calliandra sens. str.*, *Faidherbia* is related to the genus *Wallaceodendron*, while subgenera *Aculeiferum* and *Phyllodineae* pair together, basal to the united Acacieae and Ingeae.

The infrageneric analysis is shown in figure 2.2. The results from this analysis are less clear, and often contrary to results of the previous analysis. The results show *Acacia* subgenus *Acacia* to be a monophyletic group, while subgenus *Aculeiferum* is not monophyletic, coming out basal to, and within subgenus *Phyllodineae*. Both *Faidherbia* and genera within the Ingeae appear in positions contrary to those in the Mimosoideae analysis. *Faidherbia* appears between one group of subgenus *Aculeiferum* and subgenus *Phyllodineae*.

The authors discuss the possible nomenclatural implications of their analyses. They assume that names will have to be changed, and they discuss possible alterations. Chappill and Maslin concentrate on the fate of the generic name *Acacia*. It has always been assumed that following

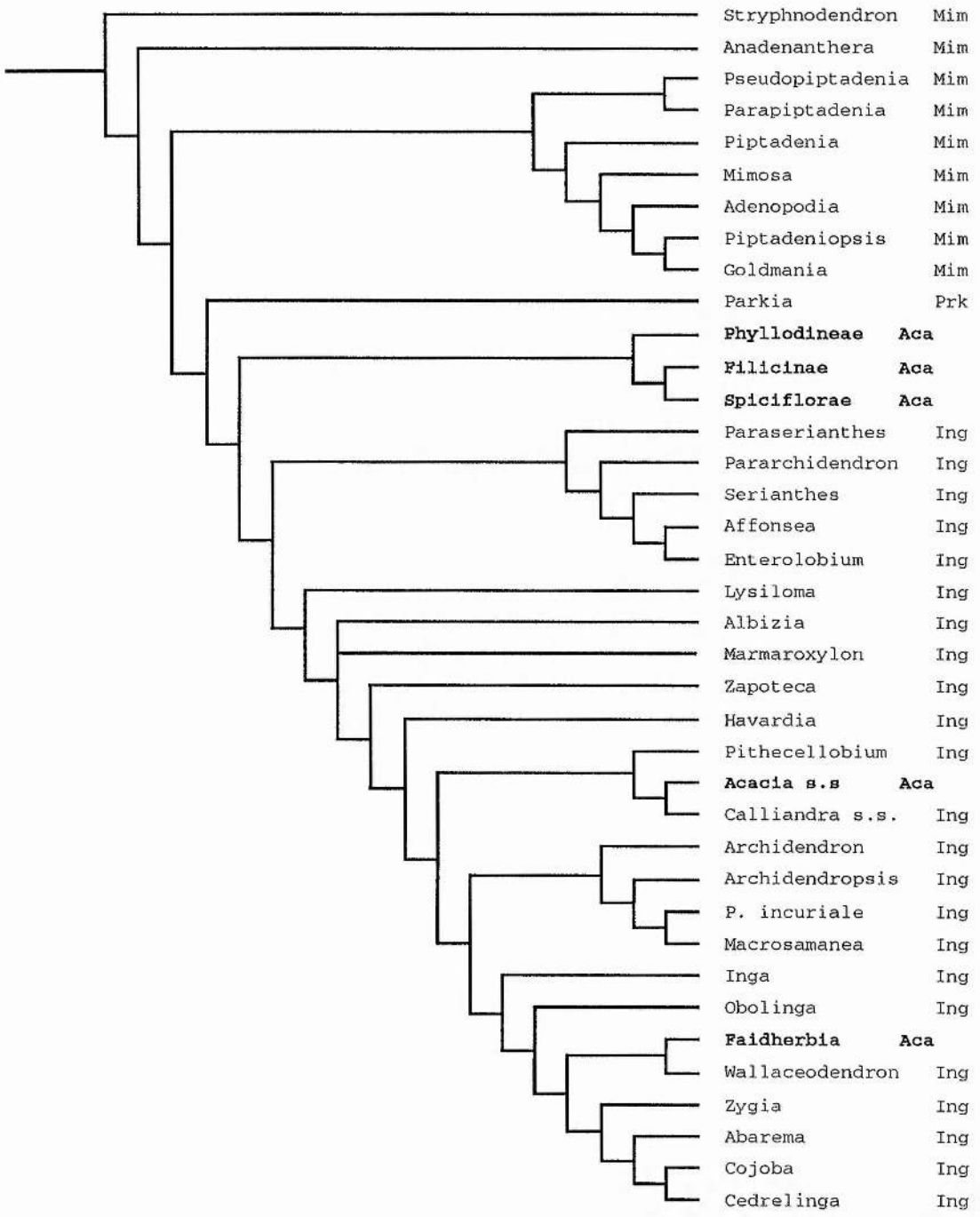


Figure 2.1. From Chappill & Maslin (1995). Part of a 50% majority rule consensus tree for the genera of the Mimosoideae relevant to the Ingeae and Acacieae. The letters after the generic names indicate their tribal placement in the classification of Polhill and Raven (1981): Mim - the tribe Mimoseae; Ing - the tribe Ingeae; Aca - the tribe Acacieae; Prk - the tribe Parkieae. For the genus *Acacia* the names of the major infrageneric groups are shown, following Pedley's (1978) scheme: **Phyllodineae** - *Acacia* subgenus *Phyllodineae*; **Filicinae** - *Acacia* subgenus *Aculeiferum* section Filicinae; **Spiciflorae** - *Acacia* subgenus *Aculeiferum* section Spiciflorae; **Acacia s.s.** - *Acacia* subgenus *Acacia*. *P. incuriale* presumably refers to *Pithecellobium incuriale*.

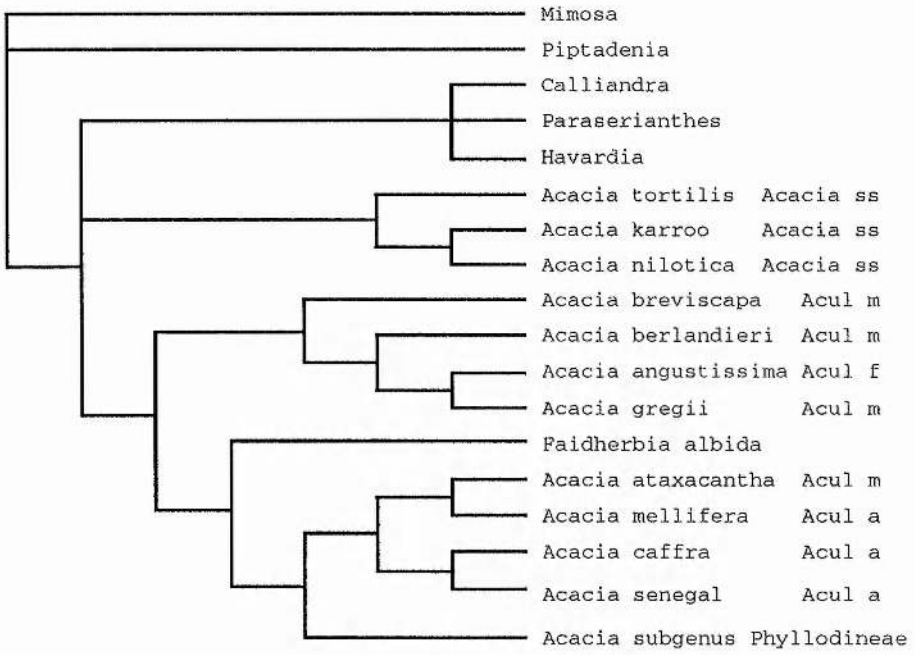


Fig 2.2. From Chappill and Maslin (1995). A strict consensus tree from the analysis of selected *Acacia* species and outgroup genera. For the subgroups of *Acacia* the classification of Pedley (1978) is used. Acul f - subgenus *Aculeiferum* section *Filicinae*, Acul m - subgenus *Aculeiferum* section *Monocanthea* and Acul a - subgenus *Aculeiferum* section *Aculeiferum*. In the original cladogram there were many species from subgenus *Phyllodineae* included, as they formed a monophyletic branch I have reduced them to just one branch.

fragmentation, the name *Acacia* would be applied to the taxon *Acacia* subgenus *Acacia*. After all, this is consistent with the lectotypification of the name by Britton and Rose (1928) on *Acacia nilotica* DEL. On this assumption the name would be applied to approx. 200 spp., the remaining species, approximately 250 in subgenus *Aculeiferum* and approx. 900 in subgenus *Phyllodineae* needing a new generic name or names.

Chappill and Maslin, however, suggest that it would be better, in the interests of stability, to re-typify *Acacia* on a taxon in subgenus *Phyllodineae*. The name *Acacia* would thus be used for the largest group of species meaning fewer nomenclatural changes. This is in keeping with Article 14 of the International Code for Botanical Nomenclature (ICBN).This idea, i.e. to restrict

Acacia to Australian species, was first proposed by Bentham (1840). They also suggest that there are other advantages for non-Australian workers if the name *Acacia* is applied to subgenus *Phyllodineae*. If three genera were eventually recognised, as suggested by Pedley (1986), then the subgenera *Acacia* and *Aculeiferum* would each receive a new generic name. Chappill and Maslin take the view that it would be "easier for a country to accept a change of generic name for all indigenous *Acacia* species than it would be to accept a change for only half of them". "Therefore, to retain the current lectotypification and restrict *Acacia* to subgenus *Acacia* but to provide a new name for subgenus *Aculeiferum* may be more confusing than to provide new generic names for both of these subgenera."

Finally they re-discuss the question of whether *Racosperma* MARTIUS is validly published? No new suggestions are presented.

To summarise, Chappill and Maslin's results and conclusions are

1) The tribes Acacieae and Ingeae should be amalgamated. This is supported by numerous characters.

2) The taxonomic position of *Faidherbia* is equivocal. The generic analysis supports it as a distinct genus within the Ingeae, but the infrageneric analysis places it within a paraphyletic assemblage of subgenus *Aculeiferum* species.

3) Fundamental differences exist between *Acacia* subgenus *Acacia* and *Acacia* subgenera *Aculeiferum* and *Phyllodineae*.

4) *Acacia* subgenera *Aculeiferum* and *Phyllodineae* probably form a monophyletic group. In all likelihood subgenus *Phyllodineae* is monophyletic within this group, subgenus *Aculeiferum* needs further study before a firm conclusion can be reached concerning its monophyly.

5) Within subgenus *Phyllodineae*, existing classifications need to be reassessed, to derive "meaningful infrageneric categories that truly reflect the evolutionary history of this group" (Chappill and Maslin, 1995).

Although believing their results lend support to previous suggestions (Guinet, 1969; Pedley 1986) that *Acacia* comprises more than one genus the authors considered it inadvisable to undertake any formal splitting of the genus *Acacia* at present.

As the paper has only recently been published, their results have not been discussed in a wider arena. However there are three aspects of the paper which appear to be open to criticism. These are, the cladistic analysis, their results and the nomenclatural suggestions.

Starting with the cladistic analysis and the data set, it is apparent that there is a large amount of missing data. The matrix for the generic analysis of the Mimosoideae was not included so I cannot comment upon it. However the matrix for the infra-generic analysis was included. Of the possible 6862 character/taxa combinations, 1945 were scored as unknown or missing. This approximates to 28% of the data. This large amount of missing data is reflected in the problems the authors had in finding/searching for a minimal tree. Due to memory constraints on the computer only 10,500 trees were saved. It is impossible to say how many unsaved trees were discarded. This constraint means that the accuracy of the consensus tree derived from the 10,500 saved trees may be compromised. This is a problem recognised by the authors, a problem that perhaps would have been solved had they not had a fixed presentation date. They promise "a full analysis... at a later date when more of the missing data have been found".

Considering now the methodology of the cladistic analysis, there is an important question not addressed in the manuscript, namely how was character polarity decided? This is one of the most important aspects of a cladistic analysis. Another critical feature of cladistics is that synapomorphic character states must be identified and that they alone provide the basis for clade identification (Avise, 1994), i.e. characters which are false

synapomorphies, due either to convergence or parallelism, should not be included in the analysis.

Despite this there are characters included which fall into this category. Armature has evolved in *Acacia* as a defence mechanism against predation (Brown, 1960). All but a few species of *Acacia* in Africa, where grazing pressure is very high, have well developed armature. The converse is true in Australia, where there has been a "long continued absence or scarcity of effective large browsers", and very few species are armed. Attributes such as this are under great selective pressure and are not of great value in assessing evolutionary trends. Inflorescence shape is another character that is known to be 'convergent'. Chappill and Maslin divide this character into three states; globular; oblongate and spicate. In subgenus *Acacia* all three states are present. In the other subgenera *Aculeiferum* and *Phyllodineae* all three states are also present. This character, inflorescence shape, is thus not a stable character in Acacieae since it moves freely between all character states. Characters such as this should not be included in any cladistic analysis.

Considering now the results presented; the generic analysis of the Mimosoideae appear to be of most importance, since the results of the inter-specific analysis are unreliable because of the missing data and unresolved relationships.

The implications of Chappill and Maslin's cladogram for the classification of the Ingeae differ from a recent cladistic treatment of the tribe, also based on morphological characters (Arce, 1989). Without going into great detail, many species relationships indicated by Chappill and Maslin are contrary to those of Arce. For example Arce (1989) suggests that "together *Affonsea* and *Inga* are ... one genus, with *Inga* being the correct name". Chappill and Maslin's cladogram (see figure 2.1) however places *Affonsea* and *Inga* a long distance from each other.

Also there are similar contradictions between Chappill and Maslin (1995) and traditional classifications of the Ingeae such as that of Neilsen (1979).

Finally, their nomenclatural 'considerations' should be considered. They suggest that *Acacia* be retypified on a taxon in subgenus *Phyllodineae* so that fewer nomenclatural changes occur. This apparently is in keeping with Article 14 of the ICBN. This was additionally justified with the notion that it would be easier for "non-Australian" workers if both subgenus *Acacia* and subgenus *Aculeiferum* had name changes. This appears somewhat parochial. Why "non-Australian" (presumably this means African and Latin American) workers would find it difficult to understand the splitting of the genus, and the resulting nomenclatural change, is never explained. The suggestion that it is easier and less confusing to accept a change in generic name for all species of *Acacia* than to accept a change for only half of them is not readily acceptable.

A consensus view of other workers is that stability can only be reached by rigorous applications of the rules of the ICBN. The principle of priority should take precedence in a case like this i.e. the generic name *Acacia* should stay with the taxon the generic description of *Acacia* was based upon, *Acacia nilotica*, in *Acacia* subgenus *Acacia*. It is not permitted to change the name of a plant for reasons of convenience.

2.7 Concluding remarks on the taxonomy of *Acacia*.

Chappill and Maslin's work is the last work that deals with the whole of the tribe Acacieae. If we recall the history of the genus, it is obvious that throughout large periods of its existence the classification of the genus *Acacia* has been in a state of uncertainty, with very little overall consensus having been reached. The problem with the classifications discussed above is that they are based mainly on unsatisfactory morphological characters. Guinet and Vassal (1978) comment that in *Acacia*

"delimitation of subdivisions and clarification of relationships between groups and between species are often beset with difficulties. Taxonomic limits are frequently obscured by continuous variation of characters. Moreover, the levels of specialisation in different characters may be unequal in any taxon. The relationships then overlap to a greater or lesser degree and limits become difficult to establish."

Although Guinet and Vassal were talking specifically about the genus *Acacia*, the same appears to be true for the tribe Ingeae, where morphological characters are also plastic. This makes comparison between the two beset with difficulty.

An independent study of the tribe Acacieae is called for. If morphological characters cannot be used, or as in this case, cannot be relied upon to give accurate representation then we must turn to other characters. The field of molecular taxonomy is one which perhaps can enable us to investigate the phylogeny of *Acacia*. Molecular characters can in some instances be seen to be of greater use in classifications, they are less influenced by environmental changes, independent of each other and no assumptions need to be made concerning their polarity or homology. These features of molecular methods are important in this study, because of the plasticity and convergence shown by morphological characters.

This is the primary objective of my research, i.e. to assess the relationships of the subgenera of *Acacia* and certain genera of the Ingeae and Mimoseae using molecular characters.

Chapter 3
An introduction to the use of cpDNA in phylogenetic analysis.

The aim of this project was to produce a chloroplast DNA (cpDNA) phylogeny of the genus *Acacia* MILLER, this chapter and the following chapter describe the work associated with this.

3.1 Examples of the uses of cpDNA in systematics.

The taxonomic problem outlined in chapter 2 has been addressed in this study by analysing restriction site data from cpDNA. The choice of this molecule was based on previous taxonomic studies in which it had proved useful.

There have been many such studies. A few are reviewed here because they illustrate features of particular interest to the present investigation. These are; a) phylogenetic analysis at higher taxonomic levels; b) analysis of relationships at the specific level and c) analysis of the parentage of hybrids.

3.1.1 Generic Level and above : cpDNA variation in Woody Papilionoideae.

For reasons that will be mentioned later, the uses of restriction site mapping in phylogenetic analyses are limited at higher taxonomic levels by the amount of homoplasy encountered. However because of the slow rate of evolution amongst the woody Papilionoideae restriction site mapping has been useful in assessing the relationships in this subfamily of the Leguminosae (Doyle *et al.*, 1992). In the tribes Millettieae and Robinieae restriction map characters have been useful for the reconstruction of generic phylogenies within tribes. These studies (Lavin and Doyle, 1991; Doyle *et al.*, 1992) suggested that the Robinieae is a monophyletic tribe, comprising two

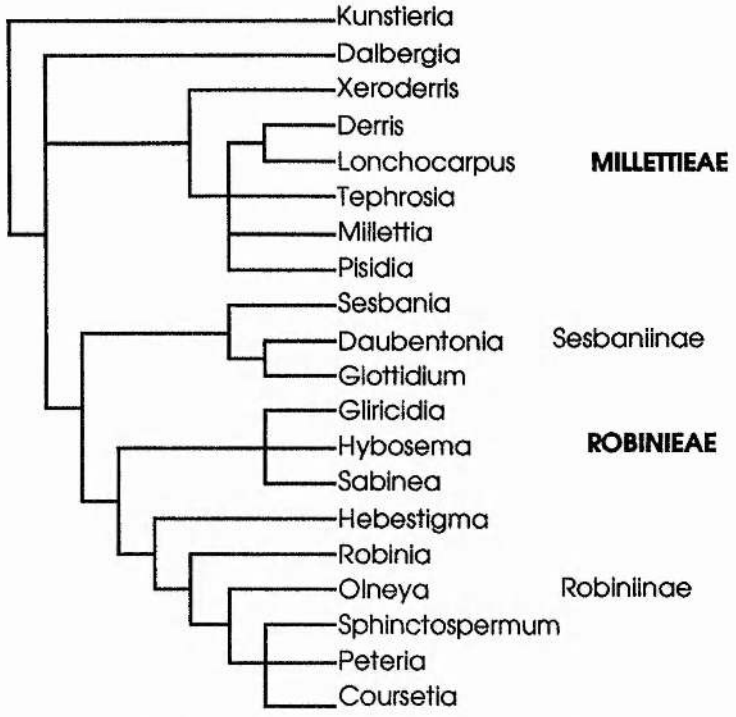


Figure 3.1. Cladistic Relationships of Millettieae and Robinieae. Tribal and subtribal classifications are shown to the right of the genera. (After Doyle *et al., 1992*).

separate lineages that correspond to the traditionally recognised subtribes, Sesbaniinae and Robiniinae. The cpDNA data was consistent with Rydberg's (1924) classification rather than with a recent cladistic treatment (Lavin, 1987), whose study suggested Sesbaniinae was closely related to basal groups in the Robiniinae. This prompted a re-evaluation of the morphological characters of the Sesbaniinae. A different interpretation of two characters, leaf nyctinasty and bracteolate flowers, produced a topology congruent with the chloroplast data. The monophyly of the Millettieae could not be confirmed.

The restriction site characters also allowed for the assessment of certain problematic genera. The genus *Tephrosia* was considered closely related to either the Robinieae (Sousa and de Sousa, 1981) or to genera of the Millettieae (Geesink, 1984; Lavin, 1987). The chloroplast data supported the latter relationship (Lavin and Doyle, 1991), as can be seen from the cladogram (see figure 3.1). The affinities of the genus *Sphinctospermum* had also been

disputed. *Tephrosia* (Lavin, 1987) or *Coursetia* (Polhill and Sousa, 1981; Wood, 1949) had been suggested as possible relatives. As we can see from the cladogram the cpDNA data suggested a close relationship between *Coursetia* and *Sphinctospermum*. In this study, the use of cpDNA had two main consequences. The relationships of the genera as revealed by the cpDNA suggested a re-evaluation of the morphological characters as used by Lavin (1987). When re-evaluated, the new morphological data set was congruent with the cpDNA data. Without the independent viewpoint of cpDNA data it is unlikely that the morphological characters would have been re-examined. The independence of the cpDNA, in the latter part of the study, allowed Doyle *et al.* (1992) to suggest affinities for certain problematical genera.

This example has been chosen for two reasons. Firstly, the taxonomic level of the taxa investigated is similar to that of *Acacia* in the present study. The studies of Lavin and Doyle (1991) and Doyle et al. (1992) suggest that useful taxonomic variation can be found at the generic level using cpDNA restriction site data.

Secondly, the taxa investigated are woody legumes like *Acacia*, albeit in a different subfamily of the Leguminosae. One would therefore expect that *Acacia* too would be amenable to cpDNA restriction site analysis and give useful results.

3.1.2 Specific level : phylogenetics of the genus *Gossypium* (Cotton).

The genus *Gossypium* L. consists of about 50 species of shrubs and small trees found throughout the tropics and subtropics. From these, four species have been extensively cultivated world-wide for either fibre or oilseed. Wendel and Albert (1992) looked at the phylogenetic relationships of 40 species of *Gossypium* using restriction site variation. The two methods they used were a Wagner parsimony analysis and a character state weighting approach. The Wagner parsimony analysis resulted in four equally

parsimonious trees. The character state approach, with gain/loss weight ratios of between 1.001:1 and 2.3:1, identified two equally parsimonious trees, identical to two of the Wagner parsimony trees. These two trees differed in the placement of one taxon, *G. longicalyx*.

The cladistic relationships as illustrated by the cpDNA data were congruent with both cytogenetic groupings and geographical clustering. The three major monophyletic clades revealed corresponded to three continents, Australia (representing the C- and G-genomes), the Americas (D- genome) and Africa (A-, E- and F-genomes). There was also a certain degree of congruence between the cpDNA phylogeny and traditional taxonomic treatments though at lower infrageneric ranks there was disagreement over placement of certain taxa. Wendel and Albert attributed these inconsistencies to reticulate evolution among the diploids.

In addition to the phylogeny of *Gossypium*, cpDNA data allowed Wendel and Albert (1992) to look at the biogeography of the genus. Their data suggested that *Gossypium* originated in either Africa or Australia. The initial splitting of the African-Australian species was estimated by Wendel and Albert (1992) to be during the mid to upper Oligocene. The New World seems to have been colonised twice, an early long-distance dispersal from Africa leading to the evolution of the D-genome diploids, followed by dispersal of the maternal A-genome, ancestor of the allopolyploid species in America.

Wendel and Albert's (1992) study has given us novel insights into the evolution of the genus *Gossypium*. The relationships of the taxa investigated prompted a new look at the morphological characters, suggesting that reticulate evolution has occurred. This viewpoint is supported by other data, i.e. cytogenetic and biogeographic data. These in combination with the cpDNA prompted the re-evaluation. In addition the cpDNA data allowed Wendel and Albert (1992) to date the cladogenesis events, something which could only have been performed with molecular data. The maternal parents

of the allopolyploids in America were also identified, using cpDNA, pointing to a secondary dispersal of *Gossypium* from Africa. Finally, the study confirmed the applicability of the character-state weighting to phylogeny reconstruction.

In this example of the use of cpDNA data in taxonomy there are three main areas which parallel this present study. Firstly the taxonomic level and number of accessions used correlates with those of this study. Hopefully in addition to investigations at the generic level, there will be enough variation to study the species relationships within the genus *Acacia* (an initial aim of this thesis) similar to that of Wendel and Albert's (1992) in *Gossypium*. Wendel and Albert (1992) also illustrate that it is possible to use a large number of accessions and still be able to analyse these accessions meaningfully. This is important in this present study as a large number of accessions will have to be analysed in order to sample the range of genetic variation present in *Acacia* and the Ingeae.

Secondly useful biogeographic information has been obtained using cpDNA variation at the species level. This is an area that will be considered in the present study.

Finally, Wendel and Albert's (1992) study has illustrated that a character-state weighting approach to phylogeny construction is practical. They have been able to compare this approach with that of a 'normal' Wagner parsimony approach. The advantages and disadvantages of both weighted parsimony and Wagner parsimony will be discussed later in this chapter. Due to the perceived advantages of the weighted parsimony approach over Wagner parsimony, in the present study it seems desirable to consider applying a weighted parsimony approach similar to that applied by Wendel and Albert (1992).

3.1.3 Determining the parentage of hybrids and polyploids: speciation in *Tragopogon*.

This example is relevant to Chapter 5, but is included here because cpDNA was the chief tool used in the investigation and the results obtained encouraged a similar investigation in *Acacia laeta*.

Tragopogon (Compositae) provides two "classic examples of recent allopolyploid speciation" (Soltis and Soltis, 1989), *T.mirus* OWNBEY and *T.miscellus* OWNBEY. Ownbey (1950) demonstrated that the parents of *T.mirus* are *T.dubius* and *T.porrifolius* and those of *T.miscellus* are *T.dubius* and *T.pratensis*. *Tragopogon* is an Old World genus, but the progenitors of the hybrids were introduced into North America and are now naturalised. The hybridisation event has probably occurred within the last 50 years (Soltis and Soltis, 1989). Restriction fragment analysis of the cpDNA provided additional evolutionary information concerning the origin of these allotetraploids. Soltis and Soltis (1989) identified six restriction site mutations and three length mutations that unambiguously differentiated the parental diploids. Previous studies (Ownbey and McCollum, 1953, 1954; Brehm and Ownbey, 1965; Roose and Gottlieb, 1976) suggested that *T.mirus* arose independently at least three times. The cpDNA data of Soltis and Soltis (1989) suggest that; i) *T.porrifolius* has always been the maternal parent of *T.mirus*, ii) there has been a minimum of two independent origins of *T.miscellus*, i.e. populations from Pullman WA have *T.dubius* as the maternal parent and all other populations have *T.pratensis* as the maternal parent. In addition to this, that the two rare diploid species *T.porrifolius* and *T.pratensis* are frequently the maternal parents of the hybrids has led Soltis and Soltis (1989) to suggest that "pollen load" is an important factor in determining the maternal and paternal parents of allopolyploids.

As with many cpDNA studies of taxa previously studied, Soltis and Soltis (1989) confirmed previous hypotheses about the evolution of the taxa concerned. In this study it confirmed the parents of the hybrids and the

suggestion of independent origins of several *T.mirus* populations. However, Soltis and Soltis (1989) were able to document events that only molecular data, and in particular cpDNA data, could reveal, such as the direction of the crosses that resulted in the hybrids, i.e. identifying the sexes of the parents. This insight allowed Soltis and Soltis (1989) to put forward another hypothesis; i.e. because the maternal hybrids of both hybrids are relatively scarce compared with the paternal parents, "pollen load" might play an important role in determining male and female parents of hybrids. The molecular data in this case have proved stimulating to research in floral biology.

3.1.4 Introgression : Evidence from a hybrid swarm of Azaleas in Georgia.

Kron *et al.* (1993) investigated introgression in a population of Azaleas in Georgia comprising *Rhododendron flammeum, R.canescens* and their interspecific hybrids. Only one restriction site difference in the cpDNA was found, i.e. *R.flammeum* possessed a *Ban* II restriction enzyme site absent in *R.canescens*. The authors found evidence of extensive localised cytoplasmic introgression. 33 individuals of *R.flammeum* were investigated. Of these, 26 individuals had the *R.canescens* cpDNA phenotype i.e. individuals in this population that were morphologically indistinguishable from *R.flammeum* possessed the chloroplast genome of *R.canescens*. Kron *et al.* suggested that some of the observed variation in some populations of *R.flammeum* may be due to past introgression from *R.canescens*.

This study revealed that individuals identified morphologically as *R.flammeum* had *R.canescens* type cpDNA. This shows the utility of cpDNA in revealing novel insights into the evolution of the *Rhododendron* taxa studied here.

Again this example is not directly applicable to producing a phylogeny of *Acacia*, but is included as an example to illustrate the range of cpDNA in resolving taxonomic problems.

3.1.5 Concluding remarks.

From the few examples above, it can be seen that cpDNA data are applicable over a wide range of taxonomic problems. The first two investigations considered illustrate the utility of cpDNA data in a taxonomic investigation such as the one this thesis is mainly concerned with. This range of applicability in combination with the fairly simple and well understood practical aspects of such an investigation has convinced us to use cpDNA to investigate the phylogeny of *Acacia*. In the following sections the practical details concerning the use of cpDNA as a phylogenetic tool are covered.

3.2 The use of cpDNA in phylogenetic analysis.

The seminal work of Zuckerkandl and Pauling (1965) initiated the current spate of interest in molecular systematics. The authors put forward two principles. These were i) that molecular change was likely to occur at an equal rate in all lineages and ii) that phylogenies could be deduced from the pattern of molecular change. These principles, combined with an increasing understanding of molecular biology, have opened a new field in biology; that of molecular taxonomy.

Many different molecules are being utilised to reconstruct phylogenies, but in this brief introduction I shall only be considering the use of cpDNA. There are many comprehensive reviews on the use of this and other molecular tools for phylogeny reconstruction. A general review is that of Avise (1994).

3.2.1 The chloroplast DNA molecule.

The physical attributes of the chloroplast genome have been well documented (Palmer, 1985; Palmer, 1987; Palmer *et al.*, 1988; Clegg and Zurawski, 1992). In most land plants it is in the region of 120-160 kb , though naturally there are exceptions to this e.g. *Pelargonium hortorum* has a genome size of 217 kb (reported in Palmer, 1985). In addition the complete

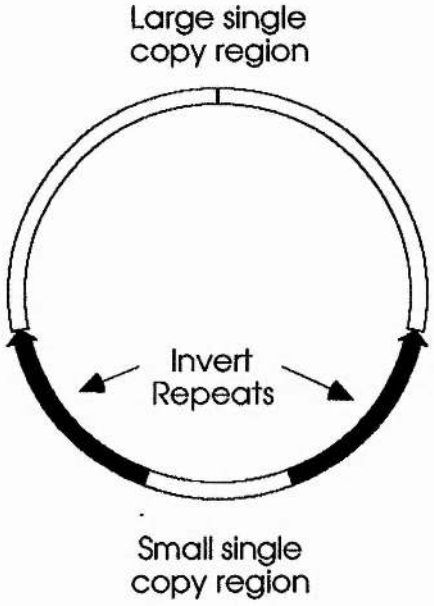


Figure 3.2. Showing the relative positions of the invert repeats and large and small single copy regions of the chloroplast DNA.

cpDNA sequence of three land plants has been obtained; *Nicotiana tabacum* (Shinozaki *et al.*, 1986); *Marchantia polymorpha* (Ohyama *et al.*, 1986); and *Oryza sativa* (Hiratsuka *et al.*, 1989).

These studies and others have enabled us to determine the structure of the chloroplast genome. Generally the chloroplast genome can be divided into 4 parts (see figure 3.2). Two identical, but inverted, regions called the 'invert repeat' of approximately 25 kb (Palmer *et al.*, 1988) separate the chloroplast genome into two regions of approx. 87 kb and 18 kb (Palmer *et al.*, 1988), the large and small single copy regions respectively. The genome is packed with 120-139 genes (Palmer *et al.*, 1988 and Shinozaki *et al.*, 1986 respectively) coding for products whose functions are primarily in photosynthesis and transcription - translation of the chloroplast genome.

3.2.2 Why use cpDNA?

What features of cpDNA have made it a useful tool for studying evolution? Palmer *et al.* (1988) viewed its "conservative mode of evolution" as "extremely valuable", whereas Clegg and Zurawski (1992) consider it as "well suited ... because (it) is a relatively abundant component of total plant DNA, thus facilitating extraction and analysis". However, a general consensus would be that cpDNA is a valuable tool because it has a suite of useful features rather than one in particular. These are (in no order); a) its conservative mode of evolution; b) the extensive background of molecular information on the chloroplast genome; c) the large number of copies in photosynthetic plant cells; d) its (usually) uniparental inheritance and e) its small size and stable structure.

These features will be considered in detail.

a) Its conservative mode of evolution. The analyses of Wolfe *et al.* (1987) concerning evolutionary rates in the different genomes of a plant cell suggested that plant cpDNA evolved at only half the rate of plant nuclear DNA, supporting the view that the chloroplast genome evolves slowly. This conservatism has both practical and fundamental advantages (Clegg and Zurawski, 1992). The fundamental advantage of such conservatism is the ability to probe and resolve phylogenetic relationships at "deep levels of evolution" (Clegg and Zurawski, 1992), typically from the interspecific level to the generic level. The more practical advantage is that chloroplast probes developed for one study can be used for examining a wide range of taxa. This feature means that the expensive and time consuming job of preparing probes does not have to be done for each new taxon investigated. For example, Sytsma and Gottlieb (1986) successfully used *Petunia* (Solanaceae) cpDNA probes to measure chloroplast variation in species of *Clarkia* (Onagraceae). This shows that cpDNA probes can possibly be used across wide taxonomic barriers.

b) The background of knowledge on the chloroplast genome structure. This enables the investigation of changes in gene content and structural organisation, as well as rates of cpDNA evolution, in taxa previously unstudied (Clegg and Zurawski, 1992)

c) The large number of cpDNA molecules in the chloroplast. Palmer (1987) estimates that there are between 20 - 200 copies of the cpDNA genome in each mature chloroplast. This makes practical aspects of the study easier. It is easy to extract, detect and analyse cpDNA.

d) Uniparental inheritance of cpDNA. Although there are exceptions to this rule (see Harris and Ingram (1991) for a review of this important topic) cpDNA is usually maternally inherited in angiosperms. There is an absence of recombination, which means that the cpDNA molecule is inherited clonally except for mutation. In plants where biparental inheritance of plastids occurs, recombination between the chloroplast genomes has never been observed. Rather the plastids simply sort out somatically (Palmer *et al.*, 1988). Thus, because of the "historical information" (Palmer, 1987) contained within the cpDNA it is an excellent marker for evolutionary studies and can provide insights into the origins of hybrid and polyploid complexes (Palmer *et al.*, 1988)

e) Its small size and stable structure. In addition to the large numbers of cpDNAs the size and stable structure also facilitate its use. The small size of the molecules means that they can be compared, in their entirety, on a single gel and can be mapped relatively easily (Palmer, 1987).

3.2.3 Limitations on the use of cpDNA.

We can see from the above information that cpDNA can be a powerful tool for studying evolution. However, as data have gradually accumulated, through studies utilising cpDNA, some of the initial claims for cpDNA have had to be re-evaluated.

Initially, because of the conservative mode of evolution of cpDNA, it was thought that it would be useful only for studies at the species level and above. Palmer *et al.* (1988) viewed this as a "serious drawback" as the amount of useful information at the intraspecific level that could be obtained would be limited. Time and hindsight have shown this to be too broad a statement. Numerous studies have found evidence of intraspecific variation (see Harris and Ingram (1991), table 1 for details). An extreme example is described by Hosaka and Hanneman (1988). Hosaka and Hanneman, as part of their investigation into the origin of cultivated potato, looked at 113 accessions of *Solanum tuberosum* subsp. *andigena*. In these accessions they found five different cpDNA types. Hosaka and Hanneman (1988) attributed this range of polymorphism to the cultivated nature and vegetative propagation of potato. The fact that intraspecific variation exists shows us that it is not satisfactory to pick a single individual to represent a taxon, unless one can be certain that levels of cpDNA variation are low. Likewise when a small sample size is used the probability of not detecting cpDNA variation may be high (Baum and Bailey, 1989). This infraspecific cpDNA variation is likely to cause difficulties in phylogeny reconstruction at lower levels of the taxonomic hierarchy, e.g. the species level (Harris and Ingram, 1991). The problems can be exacerbated if the changes are synapomorphic rather than autapomorphic. For example, the investigation of Doyle *et al.* (1990) of *Glycine* subgenus *Glycine* revealed considerable intraspecific variation. The result of this variation was that the selection of a single accession from each of the species *G.latifolia*, *G.microphylla* and *G. tabacina* could have resulted in any of the 6 possible resolutions for this 3-taxon phylogeny.

However, intraspecific variation may be useful, for it has potential for analysing other evolutionary events e.g. auto-vs-allopolyploid speciation (Soltis *et al.*, 1989a, 1989b, Soltis and Soltis 1989) and introgressive speciation (Lumaret *et al.*, 1989; Rieseberg *et al.* 1990; Whittemore and Schaal, 1991).

Another factor which has been re-evaluated is the uniparental inheritance of cpDNA. While on the whole this is true there are exceptions and it must never be assumed that inheritance is uniparental. Harris and Ingram (1991), in a survey of 398 species from 88 families, found that 23% of these species always exhibited biparental transmission of chloroplasts while a further 4% sometimes demonstrated this mode of transmission. The use of cpDNA as a uniparentally inherited marker involves the assumption that the effects of hybridisation can be ignored, since the chloroplast genome of only one parent is being followed (Harris and Ingram, 1991). This assumption can lead to difficulties when constructing cladistic phylogenies if "there is, or has been, a strong biparental plastid transmission pattern" (Harris and Ingram, 1991). The final *caveat* is that of Birky (1978), who noted the importance of minor events over evolutionary timescales, "even if there are very low levels of parental chloroplast gene transmission and recombination, these must be measured for they become very important over long periods even though they are negligible when we look at the results of a single mating". Thus although it would seem that maternal inheritance appears to be the predominant mode of chloroplast inheritance, it would be rash to assume that maternal inheritance is the case without having first proved it. This is especially true in studies concerning introgressive and polyploid speciation.

3.3 What characters can we use?

How is the cpDNA molecule used to provide information i.e. what characters do we use? Fundamentally we are looking for changes in the sequence of nucleic acid bases in the cpDNA genome between the taxa of interest. These differences are then used to compare taxa. There are two methods of obtaining this information; indirectly via restriction site changes, or directly through sequencing.

Restriction endonucleases cleave double-stranded DNA at specific sequences, usually 4,5 or 6 bp long. For example the enzyme *Eco*R I (isolated from the bacterium *E.coli*), cuts only at the sequence, 5'-GAATTC-3' and at no other sequence. Many different enzymes, from different bacteria, have been found that cut at a variety of different specific sites (it is thought that these enzymes provide a protective role against foreign DNA for the individual bacterium). In cpDNA there are two main sources of DNA sequence variation i.e. point mutations and sequence rearrangements. Restriction enzymes have the ability to detect both these mutations.

Point mutations may either create or destroy a recognition sequence for a particular restriction enzyme. These mutations are the most common form of mutation used in phylogenetic analysis of cpDNA (Soltis *et al.*, 1992). They are easily identified. See figure 3.3.

Sequence rearrangements fall into one of two categories, insertions or deletions, and inversions. Insertions and deletions may be further sub-divided depending upon the size of the DNA inserted or deleted. Palmer *et al.* (1988) state that small deletions or insertions of DNA (1-10 bp) are probably the most common sources of variation in cpDNA but often remain undetected as they cannot be resolved by the tiny differences in migration distance on an autoradiograph. Larger insertions or deletions (50-1,200 bp) can be detected more easily. They can be seen as similar changes in length of a fragment for a number of different restriction enzymes but with no change in fragment number (see figure 3.4). These rearrangements usually occur in spacers between genes. Because they tend to cluster in these "hotspot" regions displaying high levels of variability, they pose problems for phylogenetic analysis (Palmer *et al.*, 1988). The assignment of exact homology is difficult and these two problems can result in a high degree of homoplasy. For this reason small length mutations are not usually included in a phylogenetic analysis (Palmer *et al.*, 1988). The final class of length mutation is those involving large

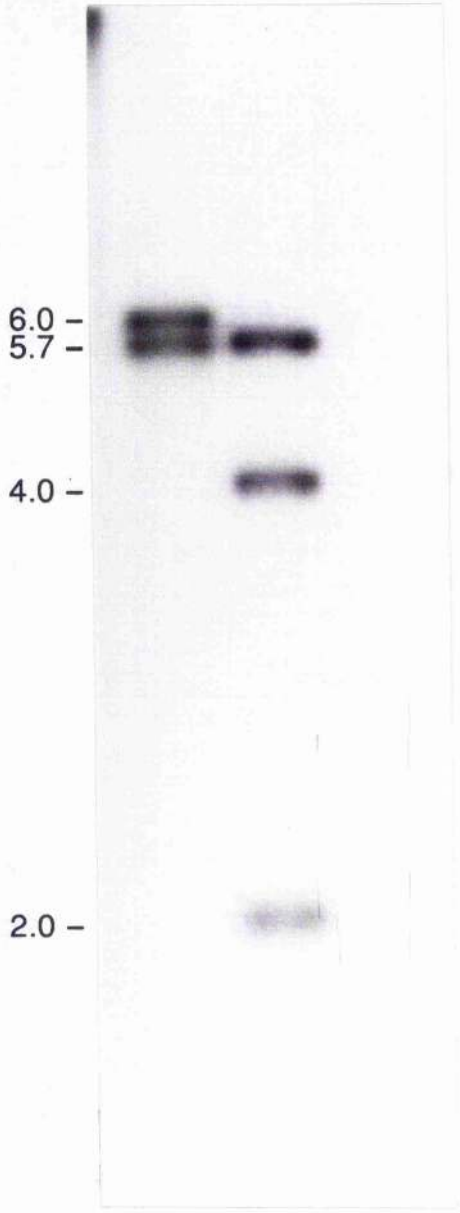


Figure 3.3. This picture of an autoradiograph illustrates one of the two main sources of DNA sequence variation.

a) (opposite) This picture of two samples illustrates a point mutation which has created or destroyed a restriction enzyme site. The autoradiogram was taken as part of the current investigation into the phylogeny of Acacia. The restriction enzyme used was *Nsi* I and the probe is the cpDNA probe MB 3 (details of the probe and restriction enzyme can be found in Appendix B).

In lane 1 is *Acacia nilotica* subsp. *subalata* and in lane 2 is *A. sieberana* var. *woodii*. In lane 1 there are two bands, at 6.0 kb and at 5.7 kb. In lane 2 there are three bands, at 5.7 kb, 4.0 kb and 2.0 kb. Both samples share the 5.7 kb band but differ at the 6.0 kb, 4.0 kb and 2.0 kb bands. The most likely explanation for the difference between *A. nilotica* subsp. *subalata* and *A. sieberana* var. *woodii* is that a restriction enzyme site which creates the 2.0 kb and 4.0 kb bands in *A. sieberana* var. *woodii* has been lost in *A. nilotica* subsp. *subalata*, leaving the 6.0 kb band. The reason for believing that a loss of a restriction site is involved is that the majority of other *Acacia* species surveyed have the 4.0 kb and 2.0 kb bands. We can illustrate this difference diagrammatically in figure 3.2 b) below.

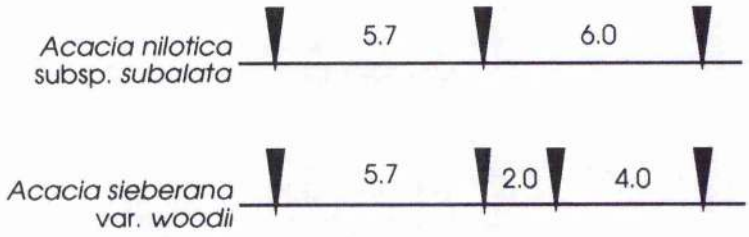


Figure 3.3 b) (above) This diagram shows the putative restriction site changes as detailed above. The restriction enzyme site in *A. sieberana* var. *woodii* which creates the 2.0 kb and 4.0 kb bands has been lost in *A. nilotica* subsp. *subalata* ,resulting in the 6.0 kb band found in this accession.

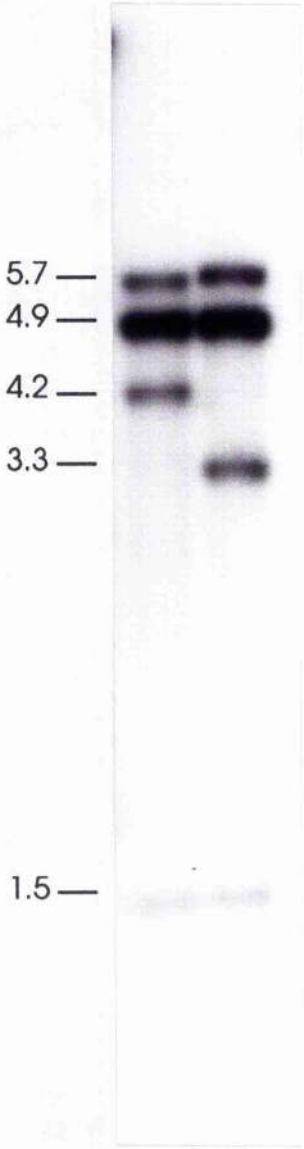


Figure 3.4. This picture of an autoradiogram illustrates the other major source of DNA sequence variation.

a) (opposite) This picture of an autoradiogram illustrates a length mutation. In lane 1 is *Acacia nilotica* subsp. *subalata* and in lane 2 is *A. sieberana* var. *woodii*.. The restriction enzyme used is *Nsi* I. and the cpDNA probe usedwas MB 7.

The only difference between the two accessions are the bands at 4.2 kb and at 3.3 kb. *A. nilotica* subsp. *subalata* has the 4.2 kb band and *A. sieberana* var. *woodii* has the 3.3 kb band. Due to the lack of change in fragment number a possible explanation for this difference is a mutation in the length of either the 4.2 kb or 3.3 kb band. The mutation is possibly an insertion of 0.9 kb in the 3.3 kb band, as the majority of other *Acacia* species surveyed have the 3.3 kb band, not the 4.2 kb band.

The test of this hypothesis is to examine other restriction enzyme profiles using the same probe, MB 7. If it is a length mutation, then a increase of 0.9 kb in the length of a DNA fragment will also be visable using the different restriction enzymes.. We can illustrate the differences diagramatically in figure 3.3 b) below.

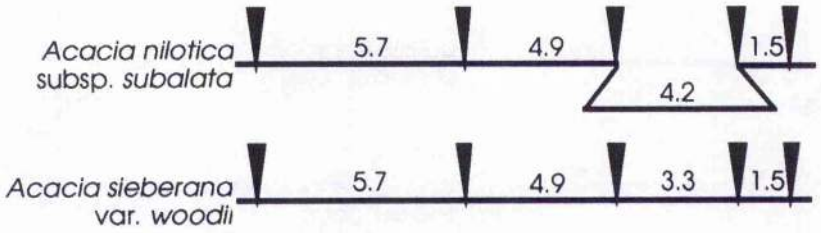


Figure 3.4 b) (above). This diagram shows the putative insertion mutation as detailed above. The 0.9kb insertion has taken place inside the 3.3 kb band found in *A. sieberana* var. *woodii*, resulting in the presence of a 4.2 kb band in *A. nilotica* var. *subalata*.

pieces of DNA, these can be very valuable biosystematic markers, e.g. the deletion of one of the invert repeats in a group of Papilionoid legumes (Doyle *et al.*, 1992)

Inversions are the rotation through 180° of a DNA sequence. The detection of inversions is complicated because they change the relative positions of a large number of restriction sites.

Both inversions and large deletion/insertion events are relatively rare in coding regions of the chloroplast genome (Palmer *et al.*, 1988). Because of this they are important evolutionary events and can act as useful phylogenetic markers.

DNA sequence analysis allows one to compare bases individually. Many of the problems of restriction site analysis are avoided. In a restriction site analysis changes at any one of the bases in the recognition sequence of the enzyme will result in a site loss. This can lead to homoplasy. The problem of assigning homology to insertion/deletion events is also avoided by sequencing. However DNA sequencing is a significantly slower and more expensive way of gathering information (Palmer *et al.*, 1988) compared to restriction site mapping. It is best used for studies at the family level and above where restriction site mapping is beset by excessive homoplasy and confounded by length mutations (Palmer *et al.*, 1988). The methodology of DNA sequencing is reviewed by Hillis *et al.* (1990). It has not been used in this present study.

3.4 How are restriction site characters used?

The raw results of a restriction site analysis are fragments on an autoradiograph. These autoradiographs can be analysed in several ways. Bremer (1991) defined two main types of analysis, each of which can be subdivided. The two main groups of methods are; RFLP or fragment methods and mapping methods.

3.4.1 RFLP methods.

The two subdivisions of the RFLP method are:

i) Fragment Direct Analysis (FDA). In this method the fragments on individual radiographs are visualised using a total cpDNA probe. Presence or absence of these bands is then scored for each taxon, the bands being the characters used. Very rarely, when enough pure cpDNA is available the fragments can be visualised directly on the agarose gel with ethidium bromide.

ii) Fragment Occurrence Analysis (FOA). In an attempt to reduce the amount of homoplasy caused by similar sized fragments being scored as the same character, FOA uses smaller separate cpDNA probes from known areas of the chloroplast genome. Restriction fragments of the same size and position on the cpDNA molecule can then be used as characters.

3.4.2 Mapping Methods.

With these methods the interpretation is taken one step further. The fragments are reconstructed to give "more or less detailed restriction maps", the restriction sites are then aligned and used as characters (Bremer, 1991).

i) Site Occurrence Analysis (SOA). In SOA the pattern of restriction sites is used. Whether a restriction site occurs or is absent from a defined position is used as a character in the analysis. Presence or absence is then tabulated for each taxon. The underlying process behind the occurrence or absence of a restriction enzyme site may be different, e.g. substitutions in the recognition site of the restriction enzyme, or a deletion or insertion in the recognition site may cause the change in site occurrence (Bremer, 1991).

ii) Site Mutation Analysis (SMA). In an attempt to make up for the shortcomings in SOA, this method of analysis makes some assumptions about the processes involved. The gain or loss of a site is only scored as a character when it is thought to be caused by the same mutation i.e. the loss of a site

through a point mutation would not be considered similar to the loss of the same site by a deletion event. Short of actually sequencing the whole of the cpDNA genome it is often difficult to determine the underlying mutation which has caused the loss or gain of a restriction site.

3.4.3 Which method is best?

Palmer (1987) reviewed the advantages of site mapping methods over fragment methods of analysis. He saw the map methods as having several advantages over fragment methods. Firstly, they allowed for the comparison of more-divergent cpDNAs than can possibly be critically analysed by the fragment (RFLP) methods (Palmer, 1987). A second advantage is that mapping allows discrimination between site mutations and length mutations as the possible causes of fragment differences (Palmer, 1987). The third and final advantage of using mapping methods that Palmer (1987) considered was again linked to their greater resolution. Often restriction enzymes that produce too many fragments to be analysed using fragment methods can be analysed using mapping methods. This can mean that fewer restriction enzymes need to be used in mapping studies to cover the same amount of DNA sequence (Palmer, 1987). Such advantages have led most workers to use mapping approaches, in particular the SOA method. The major disadvantage of the mapping approach is that it is more labour intensive and time consuming than fragment methods (Palmer, 1987).

Bremer (1991) took the process of comparison further and explicitly compared the four methods for scoring restriction site data. Was there only one accurate way of scoring the data? Bremer (1991) took data from a previous study (Bremer and Jansen, 1991) and constructed four data matrices, corresponding to the four methods of analysis proposed i.e. FOA, FDA. SOA and SMA. Bremer then subjected each of the data sets to a phylogenetic analysis. The result of her study was that the four different scoring

methods yielded solutions with different topologies, resolutions, number of shortest trees, number of steps, consistency and retention index values. The major discrepancies between the trees concerned the basal branches of the trees.

In general the preference of most workers for SOA over FOA results from the fear that without map construction, RFLP analyses will include non-homologous fragments i.e. fragments of similar length from different parts of the genome. However using small probes (1-10 kb) there is a low probability that two non-homologous fragments would have the same size (Bremer, 1991). Also if time is taken by workers to identify non-homologous fragments (e.g. Moretti *et al.* 1993) then this fault in FOA can be overcome. Of the 944 fragments identified in Bremer's study (1991), none of the fragments was shown to be homologous after mapping.

Another objection to using fragments as characters is that they are not evolutionarily independent of each other. Non-independent characters ought not to be used in cladistic analysis as they may bias the results (Swofford and Olsen, 1990). A hypothetical case is discussed by Bremer (1991). An outgroup displays a 10 kb fragment and two ingroup taxa share a restriction site, which has cut the 10 kb fragment into two smaller fragments, say 4 kb and 6 kb. With mapping methods the restriction site would be scored as a character, in this example the two ingroup taxa would share the gain of one restriction site, this would correspond to one step on the resulting phylogenetic tree (Bremer, 1991). In the RFLP methods where fragments are used as characters, the restriction site mutation would result in three characters: the presence or absence of the 10 kb, 6 kb and the 4 kb bands.

However, restriction sites can also be non-independent (Bremer, 1991), e.g. consider a 200 bp deletion that encompasses 6 restriction sites. All the taxa in which the deletion event has occurred will share the loss of 6 restriction sites, although these losses are dependent on one event.

A consequence of using non-independent characters is that there will be an artificial increase in homoplasy. This may often obscure the phylogenetic relationships of the taxa being investigated. However, the rise in homoplasy did not obscure the phylogeny of *Dioon* (Zamiaceae, Cycadales). Moretti *et al.* (1993) demonstrated that 3 times as much data was generated than was needed to support their conclusions on the phylogeny of *Dioon*

Finally returning to Bremer's (1991) study, her conclusions were; i) although the methods differ in the degree of difficulty in determining homology, the number of dependent characters and the time consumption of the analysis, none of the methods will systematically bias the resulting cladograms; and ii) the RFLP analyses are much less time consuming but probably less accurate than the site analyses (mapping methods).

The choice of method is thus dependent on a trade off between accuracy and resources (time) (Bremer, 1991). Some authors have chosen FOA methods for the sake of experimental parsimony (Moretti *et al.* 1993; Caputo *et al.*, 1991) or through pragmatism (Harris *et al.*, 1993). However, the majority of authors still use site methods, in particular SOA, because of their accuracy and because of the lower levels of homoplasy encountered.

3.5 Reconstructing phylogenies.

After scoring the restriction sites or fragments a data matrix is constructed with the characters consisting of either restriction site or fragment occurrence, these being coded as either presence or absence in each taxon studied[1]. From this it is possible to analyse the data and obtain a phylogenetic tree using one of three groups of methods; i) distance methods; ii) parsimony methods; iii) maximum likelihood methods.

1See Appendix C for an example of a restriction site data matrix.

3.5.1 Distance methods.

For these we need to convert our data matrix into one that consists of estimated pairwise genetic distances between taxa. For example, let us consider two hypothetical taxa A and B. If out of the *n* sites/fragments surveyed they are similar at *m* sites/fragments then the simplest estimate of their genetic distance (d) would be $d = 1 - \left(\frac{m}{n}\right)$. From these distance matrices there are 4 main algorithms that can be used to produce phylogenetic trees; i) UPGMA cluster analysis (Sneath and Sokal, 1973); ii) the Fitch-Margoliash Method (1967); iii) Neighbour-Joining Method (Saitou and Nei, 1987) and iv) the Distance Wagner Method (Farris, 1972). All four methods are reviewed in Avise (1994). All of these methods cluster together taxa according to overall similarity or distance. These methods are relatively simple to understand and are good examples of the phenetic approach to phylogeny reconstruction (Avise, 1994)

In the conversion of the restriction site/fragment data matrix to a distance matrix there is a concomitant loss of information concerning the pattern of gains and losses of sites/fragments. This pattern of gains and losses can provide useful information on the phylogenetic tree. For instance it is possible to see how many restriction site characters support a clade.

3.5.2 Parsimony methods.

The first question that can be addressed is the nature of parsimony. Felsenstein (1983) defined it as "a method of inferring phylogenies (evolutionary trees) by finding that phylogenetic tree on which the observed characters could have evolved with the least number of evolutionary changes". Evolutionary change is the gain or loss of a fragment or restriction site. The most parsimonious tree or trees are those in which the number of gains and losses of restriction sites or fragments is minimised The main assumption is that simple hypotheses of character change are preferred to

more complicated ones (Swofford and Begle, 1993). Like distance methods, parsimony methods encompass several related methods that differ in their assumptions about how character state transformations occur. The most commonly employed are:

a) Wagner Parsimony (Farris, 1970; Kluge and Farris, 1969). This method of parsimony is the simplest. There are no constraints on the characters, free reversibility of the characters is allowed. That is, the character may change state in any direction, with no cost involved in that change. The characters may be binary or multistate, if they are multistate then the states must be ordered, e.g. for a taxon to move from state A to D it must move through states B and C. A modification of this method by Fitch (1971) allowed any state to change to any other state without cost i.e. the change from state A to state D could be done in one step. No character polarity is assumed with this method so any of the states may be the ancestral one.

b) Dollo Parsimony (Farris, 1977). Dollo parsimony places constraints on character state change. Initially the ancestral condition for each binary or multistate character is specified. Every character state derived from that ancestral condition must be uniquely derived. That is each character state is only allowed to have originated once on the phylogenetic tree, any required homology must be reversals to a more ancestral condition (Swofford and Olsen, 1990). In effect this means that parallel or convergent gains of apomorphies are not allowed, but multiple losses of states are.

c) Camin-Sokal Parsimony (Camin and Sokal, 1965). This is the most stringent form of parsimony. It assumes that all evolutionary change is irreversible. This approach takes Dollo parsimony further by not allowing reversions to the ancestral condition. It is not, however, widely employed for molecular character sets, as these characters violate Camin-Sokal parsimony assumptions.

d) General Parsimony (Swofford and Olsen, 1990). All the methods discussed above have implicit assumptions about the costs of character transformation, for example Camin-Sokal parsimony sets an infinite weight to the loss of a character. General parsimony is an approach which offers flexibility in assigning costs to gains or losses of characters (Avise, 1994). The weights or costs assigned would in theory be gathered from independent data.

The parsimony approach to phylogenetic reconstruction is in harmony with the cladistic approach to phylogeny reconstruction. The principle of cladistics is to cluster terminal taxa by discovering for each taxon a sister group of taxa with which it is united by one or more synapomorphies.

All the above parsimony methods, with the exception of Camin-Sokal parsimony, have been applied to restriction enzyme data sets. There is, however, no consensus as to which is the best parsimony method,. All have their weaknesses. Dollo parsimony with its stipulation that derived character states may evolve only once presents a problem for restriction site data . The restriction sites are treated as individual characters, yet they consist of a string of nucleotides that themselves undergo evolution (Wendel and Albert, 1992). The drawback is, with restriction sites there is a possibility that a site can independently evolve, a possibility that Dollo parsimony does not recognise. This is "biologically unrealistic" (Wendel and Albert, 1992). Wagner parsimony, as detailed above, allows free reversibility of characters. This applied to restriction sites means that the loss of a site is equally as likely as a gain of a site. Again this is "biologically unrealistic". Thus we can see that neither Dollo or Wagner parsimony methods are 100% suitable for restriction site (or fragment) data. However the Wagner method produces more accurate topologies (Albert *et al.*, 1992).

A more realistic parsimony method is that of general parsimony. With evidence from independent data it is possible to assign costs to the gains and losses of sites. This means that parallel gains of restriction sites are allowed, but the lower probability of this happening, relative to restriction site losses, is reflected in the higher 'cost' of this event. The topic of general parsimony is discussed later under the heading of character state weighting.

3.5.3 Maximum Likelihood.

The final method of phylogeny reconstruction to be considered is maximum likelihood. This method can be thought of as a logical extension of the general parsimony method. Instead of dealing with the costs of character state change, maximum likelihood methods try to estimate the probability of change. By their very nature they are statistically complex. A maximum likelihood approach to phylogenetic analysis assesses the likelihood (probability) that a "defined evolutionary model" (Swofford and Olsen, 1990) will give us the known pattern of character states. The method selects those phylogenies with the highest likelihood. The evolutionary model used in the analysis depends on the sort of data being used. Felsenstein (1992) describes a model for restriction site characters and Swofford and Olsen (1990) describe how to apply this method to nucleotide sequence data.

Before the existence of computers, likelihoods were hard to compute. Only recently with the availability of fast cheap computers have maximum likelihood methods begun to be used with actual data. However, because the methods are numerically complex, a problem still exists with the excessive computing time.

3.6 Analysing the data.

Many of the algorithms used, to analyse the data and produce a phylogenetic tree, are numerically complex and/or iterative. This means that

doing the computations 'by hand' is very time-consuming and prone to error. To cope with this problem a number of computer programmes have been developed. At present there are two 'packages' for phylogenetic analysis, and several other programmes designed to deal with individual problems. The two packages are PHYLIP (PHYLogenetic Inference Package) (Felsenstein, 1993) and PAUP (Phylogenetic Analysis Using Parsimony) (Swofford, 1993); PHYLIP is a suite of small programmes devoted to solving individual problems, while PAUP is a parsimony programme with several interactive features. PAUP was the programme favoured in this study for several reasons; it was faster and more efficient at finding phylogenetic trees (Fink, 1986), its simplicity of operation and the additional features it offered such as allowance for missing data, the calculation of consensus trees, a range of tree statistics and bootstrap analyses.

It is important that the phylogenetic analysis programme eventually chosen is not treated as a black box, i.e. one does not just enter the data, press a few buttons and end up with a phylogenetic tree. It is important that the user of the programme knows how the programme works so that one can be aware of the limitations of any solutions. For this one must really read the manual of the programme concerned, but a few fundamental features of PAUP, a parsimony programme, will be discussed below.

3.6.1 Searching for the shortest tree

Ordinarily the goal of the search is to find all of the equally parsimonious trees that exist for a particular data set. In PAUP there are three groups of methods for doing this. Two guarantee an exact solution, i.e. exhaustive searches and branch-and-bound searches and the third approximates to an exact solution, i.e. heuristic searching methods.

T	B(T)
3	1
4	3
5	15
6	105
7	945
8	10,395
9	135,135
10	2×10^6
15	8×10^{12}
20	2×10^{20}
50	3×10^{74}

Table 3.5. This table shows the number of trees that exist for several values of *T*. The equation used to calculate B(T) is equation 3.1. (From Swofford and Begle, 1993).

---

i) Exact methods.

Exhaustive searches. This is the simplest method to understand. All the possible phylogenetic trees are evaluated and the shortest tree(s) is selected. The total number of distinct, unrooted, terminally labelled, strictly bifurcating, trees for *T* terminal taxa is given by the formula:

3.1 $$B(T) = \prod_{i=3}^{T} (2i - 5)$$ (Swofford and Olsen, 1990)

Table 3.5 shows the value of B(T) for several values of T. We can see that as the number of taxa increases the number of possible trees that exist increases dramatically. This means that trying to analyse more than 11 taxa is not really feasible, as the time required to search through all of the possible trees would be prohibitive.

Branch-and-Bound searches. Fortunately an exact algorithm that does not require exhaustive searches is available (Swofford and Begle, 1993). It differs from the exhaustive search methods in that the length of each tree is not calculated at the time of its construction, thus considerably reducing computer time.

Working through a search for five sample taxa will help illustrate how this search works (example from Swofford and Begle, 1993). We begin with the

only possible unrooted tree for three taxa (A) (see figure 3.6). Next an additional taxon is connected to tree A, giving tree B1, to this tree is added the fifth taxon giving tree C1.1. If more than five taxa were being analysed then we would continue to add taxa in this manner until we had added all the taxa. We next backtrack one node on the search tree (B1) and generate a second tree by the addition of the fifth taxon to tree B1, i.e. C1.2. When the fifth taxon has been added in all possible positions (C1.1 - C1.5) then we backtrack all the way to tree A and begin on tree B2. Again all of the five possible trees are constructed (C2.1 - C2.5). Finally we backtrack to tree A again and move forward and construct all possible trees for tree B3 (C3.1 - C3.5). If we had calculated the length of each tree when it had been constructed then we would have performed an exhaustive search. However in branch-and-bound methods all trees lengths are not calculated, or all trees even constructed.

Suppose that *L* represents an upper limit for the length of the shortest tree(s). We can obtain an initial estimate of *L*, by evaluating a random tree; if we know that a tree of length *L* exists then the length of the optimal tree(s) cannot exceed this value (Swofford and Begle, 1993). If, as we are moving through the search tree towards its tips, we encounter a tree whose length exceeds L, then we need proceed no further along this path as connecting additional taxa cannot possibly decrease the length of this tree. If this happens then the evaluation of trees down that search path can be stopped, and searching down another path can begin. By decreasing the number of trees that are constructed and evaluated the searching time is greatly reduced (Swofford and Olsen, 1990).

If at the end of the search path the length of the tree obtained is equal to the upper limit *L* then this tree is also retained. If however, the length of the tree is less than *L* this tree is the shortest tree found so far, and the upper bound on the length of the trees has been improved. This bound improvement

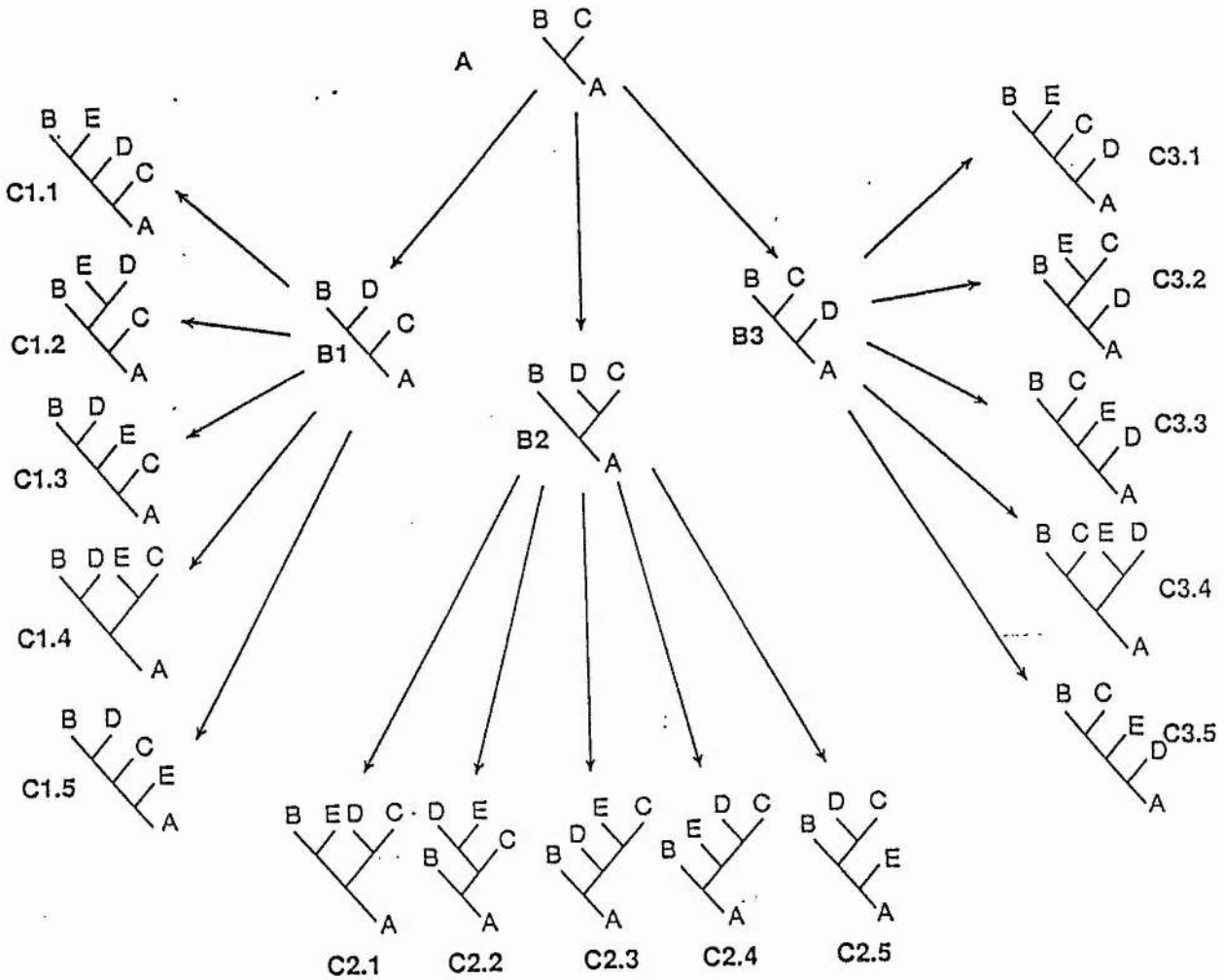


Figure 3.6 (from Swofford and Olsen 1990). The search procedure for a branch-and -bound search. See text for explanation of figure.

is important because it allows other search paths to be terminated earlier. When the branch-and-bound search has been completed, all the optimal trees will have been identified.

Although the above description of a branch-and-bound search is accurate, it is a simplification of the procedures actually implemented. PAUP has several refinements or "cheats" (Swofford and Begle, 1993) to the branch-and-bound algorithm. These include: using heuristic methods (see below) to obtain an initial upper limit to the search tree and adding divergent taxa early on, which increase the length of the initial trees in the search path.

Several factors influence the running time of the branch-and-bound method. The 'quality' of the data is the most important factor. Data sets with large amounts of homoplasy or large amounts of missing data will run considerably slower than 'clean' data sets (Swofford and Olsen, 1990). Also the type of character is important, e.g. undirected characters run faster than directed characters (Swofford and Begle, 1993). Weighted characters run considerably slower, due to the computational time required to calculate branch lengths (Swofford and Begle, 1993). Finally the type of computer used is critical to the length of the search. Supercomputers naturally operate much faster than the average desktop computer. In practice, data sets with more than 30 taxa and a large number of characters (e.g. > 200) are prohibitively large.

ii) Heuristic methods.

When the data set is too large to permit the use of a exact searching method, optimal trees can be found using heuristic methods. The guarantee of finding the optimal tree is exchanged for reduced computing time. One must not think that heuristic methods cannot find optimal trees. They often do. It is just the guarantee of optimality that is lost.

A dictionary definition of a heuristic method is "(a) system of education under which the pupil is trained to find out things for themselves" (Fowler and Fowler, 1964). This essentially means that the pupil discovers things by trial and error, and this is how heuristic searching methods operate. Initially a tree, comprising all of the taxa, is constructed and a search for the minimal length of this tree is carried out by rearranging the terminal taxa in predetermined ways. When no further way to improve the tree is found then the search is stopped. There is no way of knowing whether this is the shortest tree overall, i.e. we do not know whether we have found a local optimum or a global optimum.

The heuristic method can be broken down into two distinct steps , as outlined above. Firstly an initial tree or set of trees is obtained by a process known as stepwise addition. Secondly these trees are subjected to rearrangements in an attempt to find a shortest tree. This process is called branch swapping.

Stepwise addition.

Three taxa are chosen for an initial tree as detailed below. Next, one of the unplaced taxa is selected for addition. This taxon can be added in any of three positions, each of which is evaluated and the tree whose length is the shortest is retained for the next round of addition. In this round there are five possible positions for the taxa to be placed. Again each one is evaluated and again the tree whose length is minimal is retained. This continues until all the taxa have been added. This of course is an oversimplification of the process. How do we decide on the three initial taxa and then on the order of addition for the remaining taxa? PAUP provides four options for specifying the addition sequence.

a) As is (Swofford and Begle, 1993). The order of addition of the taxa is according to their position in the data matrix. The initial three taxon tree

comprises the first three taxa in the matrix, then the fourth is added, then the fifth, etc. This method is not very effective.

b) Closest (Swofford and Begle, 1993). All the possible three taxon trees are evaluated, and the shortest are selected for the next round. In this all of the remaining taxa are considered for addition to the tree. They are tried in all three possible addition points. The taxon and position giving the shortest tree is retained for the next addition sequence. This sequence is similar to the previous sequence in that all the remaining taxa are considered for addition, except that now there are five possible positions for the taxon. Again the best taxon-branch combination is selected for the next round. This process continues until all the taxa have been added. This addition sequence can require a great deal of computer time due to the number of possible trees that must be evaluated (Swofford and Begle, 1993).

c) Simple (Swofford and Begle, 1993). This is an implementation of the "simple algorithm" of Farris (1970). As with the 'As is' method the sequence of addition of taxa is determined prior to the beginning of the process, but it differs in the way taxa are ordered. Initially the distances between each taxon and a selected reference taxon are calculated. The initial three-taxon tree comprises the reference taxon and the two taxa closest to it. The remaining taxa are then added according to their distance to the reference taxon. The closest are added at the beginning and the furthest at the end.

d) Random (Swofford and Begle, 1993). As its title suggests the selection of the initial three taxa and the order of addition for the remaining taxa is random.

None of the above methods seems to work best for all data sets. Swofford and Begle (1993) recommend trying all of the above methods, in an attempt to produce different starting points for the subsequent branch swapping algorithms. The random addition option is not the most effective way of finding a minimal tree compared with the other stepwise addition

options, but when iterated it is useful in obtaining different branching points for branch swapping (Swofford and Begle, 1993).

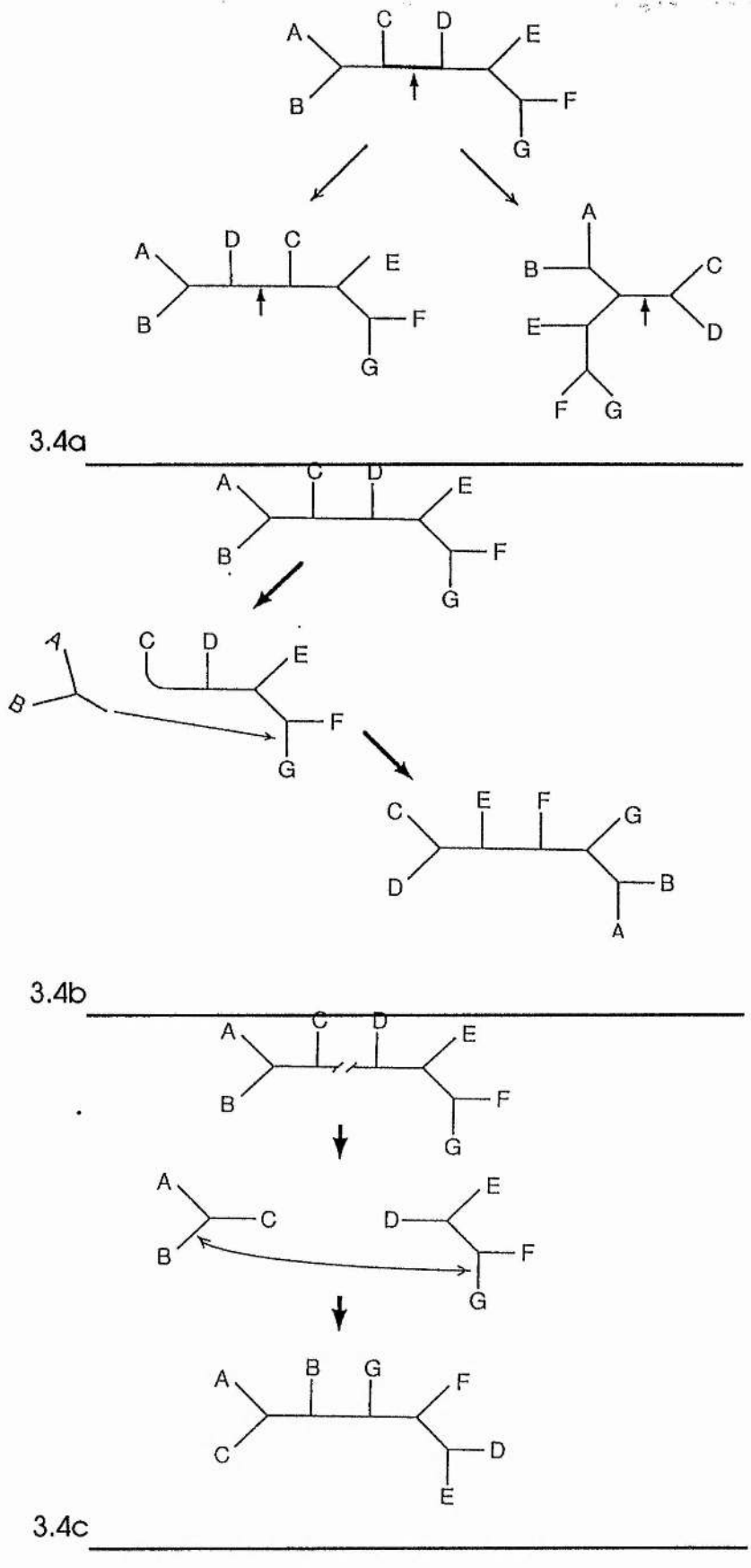
The biggest drawback of stepwise addition algorithms is that they are short-sighted, i.e. they cannot predict the effects of the addition sequence. It is not implausible to suggest that the placement of one taxon may be optimal given the taxa that have already been added. However, it may not be the best possible position for that taxon when all the other taxa have been added. In all the above algorithms for stepwise addition, once a taxon has been added it cannot be removed and repositioned until the stepwise addition process has finished. It then becomes branch swapping, which is considered below. A consequence of this is that a tree can be caught in a 'local optimum', it is often difficult to escape this and get to the 'global optimum', as offered by exhaustive and branch-and-bound searches.

Branch Swapping.

The short-sightedness and tendency to fall into local optima limits the use of stepwise addition sequences. They generally do not find optimal trees unless the data are very clean(Swofford and Begle, 1993). Branch swapping is a method of trying to improve the tree or trees selected by the stepwise addition process. In general, any one of these rearrangements amounts to a "stab in the dark" (Swofford and Begle, 1993).

PAUP offers a choice of three different branch-swapping methods. These in order of increasing effectiveness are: a) nearest neighbour interchanges; b) subtree pruning and regrafting and c) tree bisection-reconnection. See figure 3.7 for illustration of how each of these procedures work.

Each internal branch in the tree has four subtrees attached to it. In nearest neighbour interchanges (NNI) these subtrees are exchanged between each other. If we consider the example in figure 3.7a, the illustration



Figures 3.7 a) – c), from Swofford and Olsen (1990). Examples of the branch swapping algorithms, see text for details of the rearrangments.

on the left has had subtrees C + D exchanged, and on the right (this requires a little more visualisation) subtrees C + E(F,G) have been swapped.

In subtree pruning and regrafting (SPR) a subtree is cut from the main tree, and then added to a new location elsewhere on that tree. All possible subtree removals and attachment points are evaluated. (See figure 3.7b)

In tree bisection and reconnection (TBR) the tree is again cut, this time however the two subtrees are distinct from each other, the two subtrees are then reconnected to each other by joining a branch from each tree. All possible bisections and reconnections are evaluated. (See figure 3.7c).

3.6.2 Consensus Trees.

Often more than one optimal tree is found. A way of presenting all the data found in all of the trees is to produce a consensus tree. These are "hierarchical summaries" (Swofford and Begle, 1993) of the information common to all of the optimal trees, and in general, will be longer than any of the minimal trees they represent. Of the several different methods of calculating consensus trees, only two will be considered here. The strict consensus tree is the most conservative method and the easiest to interpret. In a strict consensus tree only those groupings or clades that appear in all the minimal trees being considered are represented. This method may however be too strict and give an unresolved consensus tree. We may want to see how often groupings or clades appear in the rival minimal trees. In a majority rule consensus tree groupings that appear above a pre-specified percentage of trees (50%) are represented. This means that a clade may be represented on the consensus tree even if this conflicts with some of the minimal length trees. Majority rule consensus trees are often easier to interpret, and the percentage of times a specific clade occurs within the minimal trees can be an indication of the robustness of the clade.

3.6.3 Measuring the reliability of the tree.

Several methods exist for testing the reliability of the tree or trees that have been found. These include: i) the consistency index (Kluge and Farris, 1969) which gives an indication of how well a particular tree explains the data. If a tree explains the data as well as any tree could the consistency index would equal 1; ii) the retention index (Farris, 1989) which indicates how well characters fit the tree that describes them. When a character fits the tree as poorly as possible its retention index will be 0; and iii) the homoplasy index, which provides an indication of the amount of homoplasy present in the tree (Swofford and Olsen, 1990).

How robust are the clades?

In addition to calculating a majority rule consensus tree, there are a number of procedures available for testing the robustness of individual clades. Two of these are the bootstrap and decay indices. A full explanation of the complex bootstrap procedure will not be attempted here. For a complete description see Felsenstein (1985). In the bootstrap procedure a new data matrix is constructed. It is randomly constructed using characters and character states from the original data matrix. Thus, some characters will not be present and some will be present more than once. Once the new data matrix is complete, a minimal tree for that data matrix is searched for and saved. This procedure is then repeated a predetermined number of times, usually 100, each time with a new random data set. When the iterations have been completed the topology of all the saved minimal length trees are compared. A majority-rule consensus tree is then calculated from these trees. If a particular clade occurs in 95% or more of the trees resulting from the replicate, one can conclude that the group is significantly supported (Swofford and Olsen, 1990).

The production of decay indices involves saving trees that are longer than the shortest tree. This is initially done with trees one step longer, then two

steps, etc. until the shortest tree becomes an unresolved bush. In this way it is possible to visualise the robustness of each clade. Robust clades are those that stay resolved while the tree length increases.

3.6.4 Character State Weighting.

So far most phylogenetic analyses of restriction site characters have used either Wagner or Dollo parsimony . To recap, Wagner parsimony permits a restriction site to be gained or lost with equal weight, but with Dollo parsimony a restriction site can evolve only once on the phylogenetic tree, but can be lost many times. Both these parsimony methods are illogical when applied to restriction site data, as it not impossible to have a parallel gain of a restriction site. It is just less likely than a loss of a restriction site. This can be illustrated when we consider the 'mechanics' of a restriction site. In, for example, a 6 bp restriction enzyme recognition site, any one of 18 substitutions (Albert *et al.*, 1992) can led to the loss of this restriction enzyme site. However, a gain of a restriction site is much less probable. It requires 1 specific substitution event in a sequence of six base pairs, five of which are already in their 'proper' order. This is why a parsimony model that takes into account the relative probabilities of site gains -vs- site losses is preferable. This can be done by weighting individual character state transformations i.e. the loss of a restriction site will 'cost' less than a gain of a restriction site. Albert *et al.* (1992) investigated the problem of what weights to attach to each event. The specifics of their calculations will not be detailed here. However their practical recommendations for data produced using 6 bp restriction endonucleases were that; there should be a "weighting of gains over losses by a factor of approximately 1.3 for low-level analyses" (Albert *et al.*, 1992) and by a "factor of approximately 2 for high-level analyses". Low level analyses were those examining species within a genus or even genera within a family. High level analyses were those investigating taxa at the Division level.

Having considered the use of cpDNA in phylogenetic analysis, the next chapter details its use in the construction of a phylogeny for the genus *Acacia* and related genera.

Chapter 4
A cpDNA phylogeny of the genus *Acacia* and related genera in the Mimosoideae.

4.1 Aims.

In Chapter 2, where we discussed the taxonomy of the genus *Acacia*, we concluded that there was no consensus about the classification of the genus. This is due in part to the use of morphological characters in determining the relationships of the taxa investigated. Chloroplast DNA restriction site mapping, as discussed in Chapter 3, is a method of investigating the phylogeny of *Acacia* and related genera that can overcome many of the difficulties of previous works. The aim of this chapter is to produce a phylogeny of *Acacia* and related genera using cpDNA restriction site characters.

4.2 Material used in this investigation.

The choice of material for this study was partly dictated by the material available, so pragmatism played a large part in the material used. However, having said that, there was an overall scheme of collection. In a genus of over 1200 species, it is difficult to analyse all the species, so some criteria for the selection of taxa are needed. As mentioned in Chapter 2, Maslin and Stirton (in press) have produced a list of 'critical species'. These critical species, based on an evaluation by "various experts" of *Acacia* (see Acknowledgements in Maslin and Stirton (in press)), are species which they believe are representative of natural groupings within the genus. Their criteria for selection were as follows. "1. The species in total should represent, as far as possible, the taxonomic

variation within the genus. 2. It should be possible to acquire experimental material of the nominated species with minimal cost and effort. 3. Other things being equal, preference is given to taxa that have already been studied," (Maslin and Stirton, in press).

Initially I analysed accessions as available (from collections made by C.Fagg, OFI) to optimise the procedures. These were mostly species from subgenus *Acacia* in Africa. However, as the project expanded more species were sought, and obtained, from the list compiled by Maslin and Stirton, thus relying on their more direct experience of the genus. From subgenus *Acacia* where Maslin and Stirton (in press) proposed 14 critical species or groups of species, I analysed species from 8 (57%). From subgenus *Aculeiferum* where Maslin and Stirton (in press) proposed 25 critical species or groups of species, I analysed species from 11 (44%). Given the limited availability of *Acacia* material for DNA analysis (apart from economically important species), I believe this is a respectable sample of *Acacia*. In addition Maslin and Stirton (in press) do not list and 'group' all *Acacia* species. Undoubtedly there are taxa I have analysed that correspond to 'critical groups' of species, but are not catalogued by Maslin and Stirton. The reason many taxa from the critical list were not analysed, was due to a lack of material for DNA analysis. This is a problem which will need to be addressed, should further work to be planned. The Australian species of the genus are another large grouping of taxa. There has been no argument concerning the monophyly of subgenus *Phyllodineae*, so species from this subgenus were selected according to a personal assessment made by Bruce Maslin (Western Australian Herbarium) in response to a request for his experienced judgement.

The choice of outgroups for the study was again dictated by the material available. The taxa selected from the Ingeae were designed to represent the range of taxa present and available. The outgroups from

the Mimoseae were those thought to be basal to the Acacieae and Ingeae (Guinet, 1990). A full listing of the taxa studied can be found in Appendix A, along with their authorities and the origin of the material. A list of the taxa finally used and their taxonomic position can be found in table 4.1.

Finally a note on the use of only one accession per taxon studied. I am aware of the problems in using one accession per taxon, namely that intraspecific variation is ignored. However in this study I did not feel it necessary to include more than one accession per taxon for three reasons. i) Preliminary investigations of taxa from subgenus *Acacia*, where more than one accession per taxon was investigated revealed very little intraspecific variation. ii) The level of analysis being performed concerned the relationships of subgenera and genera, not species relationships. At this level intraspecific variation should not adversely affect phylogeny reconstruction. iii) It seemed more important and time efficient to survey a wide range of species, rather than concentrate on a large sample of a small number of species, in order to provide a comprehensive overview.

4.3 Methods used in this investigation.

4.3.1Experimental Methods.

The full details of the methods used can be found in Appendix B. However, I will summarise them here. DNA was extracted from either fresh material grown from seedlings or dried material (see Harris and Robinson, 1994 for specific details of the use of dried material). This DNA was then cut using a variety of enzymes. The enzymes were selected on the folllowing criteria; a) previous investigations had demonstrated the applicability of these enzymes (Gillies, 1994; Harris, 1990; Harris et al. 1993): b) the recognition sites of the enzymes were 6 bp long, so hopefully a 'manageable' number of restriction fragments would be produced: c) the

Table 4.1. Taxa used for the chloroplast DNA analysis. The taxonomic groupings are based on Maslin and Stirton (in press). They based the groupings on a combination of the schemes proposed by Vassal (1972) and Pedley (1978). In addition the geographic origin of the taxa is indicated, New World includes subtropical and tropical America and Africa includes Arabia.

subtribe Acacieae	
Acacia subgenus *Acacia*	
Acacia abyssinica	Africa
Acacia arenaria	Africa
Acacia dolichocephala	Africa
Acacia drepanolobium	Africa
Acacia erioloba	Africa
Acacia exuvialis	Africa
Acacia hebeclada subsp.*chobiensis*	Africa
Acacia karroo	Africa
Acacia leuderitzii var. *leuderitzii*	Africa
Acacia nilotica subsp. *kraussiana*	Africa
Acacia nilotica subsp. *subalata*	Africa
Acacia nilotica subsp. *tomentosa*	Africa
Acacia rhemmiana	Africa
Acacia seiberana var. *seiberana*	Africa
Acacia seiberana var. *woodii*	Africa
Acacia seyal var. *fistula*	Africa
Acacia seyal var. *seyal*	Africa
Acacia tortilis subsp. *heteracantha*	Africa
Acacia tortilis subsp. *raddiana*	Africa
Acacia tortilis subsp. *spirocarpa*	Africa
Acacia xanthophloea	Africa
Acacia amentacea	New World
Acacia caven	New World
Acacia choriophylla	New World
Acacia cucuyo	New World
Acacia daemon	New World
Acacia farnesiana	New World
Acacia macracantha	New World
Acacia pennatula	New World
Acacia pringlei	New World
Acacia roigii	New World
Acacia schaffneri	New World
Acacia subgenus *Aculeiferum*	
section Aculeiferum	
Acacia burkei	Africa
Acacia caffra	Africa
Acacia galpinii	Africa
Acacia mellifera subsp. *detinens*	Africa
Acacia nigrescens	Africa
Acacia persiciflora	Africa
Acacia polycantha subsp.*campylacantha*	Africa
Acacia senegal var. *leiorhachis*	Africa

Acacia senegal var. *senegal*	Africa
section Monacanthea	
Acacia ataxacantha	Africa
Acacia breviscapa	Africa
Acacia chariessa	Africa
Acacia montigena	Africa
Acacia gaumeri	New World
Acacia glomerosa	New World
Acacia greggii	New World
Acacia mammifera	New World
Acacia riparia	New World
Acacia sericea	New World
section Filicinae	
Acacia angustissima	New World
Acacia chamelensis	New World
Acacia rosei	New World
Acacia tequilana	New World
Acacia subgenus *Phyllodineae*	
section Botrycephalae	
Acacia mearnsii	Australia
section Alatae	
Acacia alata	Australia
section Phyllodineae	
Acacia paradoxa	Australia
Acacia pycnantha	Australia
section Plurinerves	
Acacia melanoxylon	Australia
Acacia koa	Hawaii
Faidherbia	
Faidherbia albida (syn. *Acacia albida*)	Africa
subtribe Ingeae	
Albizia	
Albizia harveyi	Africa
Albizia schimperana	Africa
Albizia versicolor	Africa
Albizia saman	New World
Albizia tomentosa	New World
Calliandra	
Calliandra calothyrsus	New World
Enterolobium	
Enterolobium cyclocarpum	New World
Pithecellobium	
Pithecellobium dulce	New World
subtribe Mimoseae	
Piptadenia	
Piptadenia viridiflora	New World
Prosopis	
Prosopis juliflora	New World

recognition sites were all different: d) the recognition site needed six specfic base pairs, e.g. the restriction enzyme *Nsp* I was not considered as its recognition sequence is 5'-PuGATG↓Py-3', where Pu is any purine and Py is any pyrimidine: and finally, e) the cost of the restriction enzyme was taken into account i.e. expensive enzymes were not considered. The fifteen enzymes finally used were: *Apa* I, *Bam*H I, *Bcl* I, *Bgl* II, *Cla* I, *Eco*R I, *Eco*R V, *Hin*D III, *Nsi* I, *Pst* I, *Pvu* II, *Stu* I, *Sma* I, *Sst* I and *Xho* I. Initially the enzymes *Sno* I, *Dra* I and *Mlu* I were used, however they were found not to be suitable as they did not cut the cpDNA often enough to be of use. The specific cutting sites of the fifteen enzymes finally used can be found in Appendix B. The cut DNA fragments were then separated on an agarose gel and transferred to a nylon membrane by the method of Southern (1975). This nylon membrane was sequentially probed with non-overlapping probes from *Vigna radiata* (mung bean). The mung bean set of probes was selected from the cpDNA probe sets available as it is taxonomically the closest to *Acacia*. The nearer taxonomically the sample DNA and the probe DNA, the easier the detection of restriction fragments is. Other sets that were available were a *Petunia* (Solanaceae) set, a *Lactuca* (Compositae) set and a *Nicotiana* (Solanaceae) set. The mung bean probes were a gift from Jeff Doyle (Cornell University, USA). Probes MB8 and MB 9 were used together, all the other MB probes, MB1, MB2, MB3, MB 5, MB7, MB11, MB 12 were used singly. See Appendix B for details of these probes. Together these probes covered over 89% of the chloroplast genome. The resulting autoradiographs were then used to construct restriction maps of the taxa. The SOA method of Bremer (1991) was used, mainly because of the perceived greater accuracy of this method of data analysis.

4.3.2 Data analysis.

The restriction enzyme sites were scored as either present (1) or absent (0) or missing/undetermined (?) for each of the taxa studied. The resulting data matrix was analysed by PAUP 3.1.1 (Swofford, 1993) on an Apple Macintosh Powerbook 165c (33 MHz, with maths co-processor). The amount of data to be analysed meant that heuristic searching methods had to be adopted. Preliminary trials with the branch-and-bound algorithm suggested that this method would be prohibitively time consuming. Initially the characters were analysed under Wagner parsimony, which assigns no costs to character change. The addition sequence that was chosen was the random addition option, which was iterated twenty times. Although not the most effective addition sequence, when iterated it is very effective at obtaining different starting trees for the branch swapping algorithms. The branch swapping algorithm used was tree-bisection and reconnection (TBR).

In addition to the Wagner parsimony analysis, a character state weighting approach was attempted. The methodology of Albert *et al.* (1992) was used; an ANCSTATES statement specifying the ancestral condition for all the characters as unknown was entered; only trees which were compatible with the constraint tree were analysed. This constraint tree forced outgroup monophyly, where *Piptadenia viridiflora* was the outgroup and a step matrix implemented the desired weighting of character state change. The weighting was 2:1, i.e. the cost of gaining a restriction site was twice the cost of losing a restriction site. This follows the recommendations of Albert *et al.* (1992) and Wendel and Albert (1992). The minimal tree was searched for in a similar fashion to the Wagner analysis, i.e. a heuristic search with 20 random addition sequences and TBR branch swapping.

PAUP was also used to; calculate the consensus trees (strict and majority rule); perform the bootstrap analysis; construct decay analyses and also calculate tree statistics such as the consensus index.

4.4 Results.

Restriction maps were constructed for all the taxa apart from *Calliandra calothyrus*. This species appeared to have a chloroplast genome highly divergent from the rest of the taxa studied. The reasons for this were not investigated. It was decided not to construct a restriction map for this species because of the speculative nature of any such map, and the excessive time that would be needed to construct it. Similarly restriction maps could not be constructed from the restriction fragments for five probe/restriction enzyme combinations (PEC). These were, *Bcl* I/MB 11, *Bcl* I/MB 12, *Cla* I/MB12, *Eco*R I/ MB 11 and *Eco*R I/MB 12. Again the reasoning for not constructing maps in these areas was the excessive amount of time it would take and the level of supposition needed. Maps for the other 115 PEC (eight probes x fifteen restriction enzymes) were constructed.

At least 559 restriction sites were identified, of these 391 (70%) were found to differ between the taxa. Of the 391 restriction sites that differed 137 (35%) are autapomorphic, i.e. they are only present in one taxon. Assuming a genome size of 150 kb (Palmer, 1985), then by surveying 559 restriction sites of 6bp in length, 3354bp were surveyed, i.e. 2.2% of the genome. Hillis *et al.* (1994), modelling the number of nucleotides needed to find a correct solution (at a probability of greater than 99%), found that uniformly weighted parsimony methods required at least 2,000 nucleotides. The restriction sites that differed between the taxa can be found in Appendix C, along with the PECs that identified them.

A binary data matrix was constructed from the restriction site characters. This can also be found in Appendix C. Initially *Prosopis juliflora* was included with the taxa studied, as an outgroup, and a restriction map was constructed for this taxon. However, in preliminary 'runs' of PAUP it was found that this taxon was slowing down the analysis considerably. It was therefore decided to remove it from the analysis. This was a pragmatic necessity, as the search for the minimal length tree was expected to take a long time and any unnecessary time increase could not be justified. *Prosopis juliflora was* included as an additional outgroup taxon in the Mimoseae, *Piptadenia viridiflora* was also present as an outgroup, so the loss of *P.juliflora* did not compromise the investigation.

A parsimony analysis of the data matrix resulted in 720 minimal length trees of length 714 steps. It took about six weeks to find the minimal length trees. No additional minimal length trees were found after the thirteenth addition sequence, Swofford and Begle (1993) suggest that if no new trees are found for several addition sequences in a row then the chances are that all the minimal length trees have been found. The strict consensus tree and the majority rule tree can be seen in figures 4.2 and 4.3 respectively. A consistency index of 0.536, a retention index of 0.831 and a homoplasy index of 0.464 were returned. The significance of these values will be discussed later.

The character state weighting approach was attempted, but it had to be terminated. After searching for over 1000 hours (~ 2 months) it had failed to progress past the first addition sequence, and did not appear to be near completing this sequence. So due to time constraints the programme run was stopped.

Likewise, the bootstrap analysis could not be completed due to length of computational time it required. This does not affect the result of the Wagner parsimony search in any way.

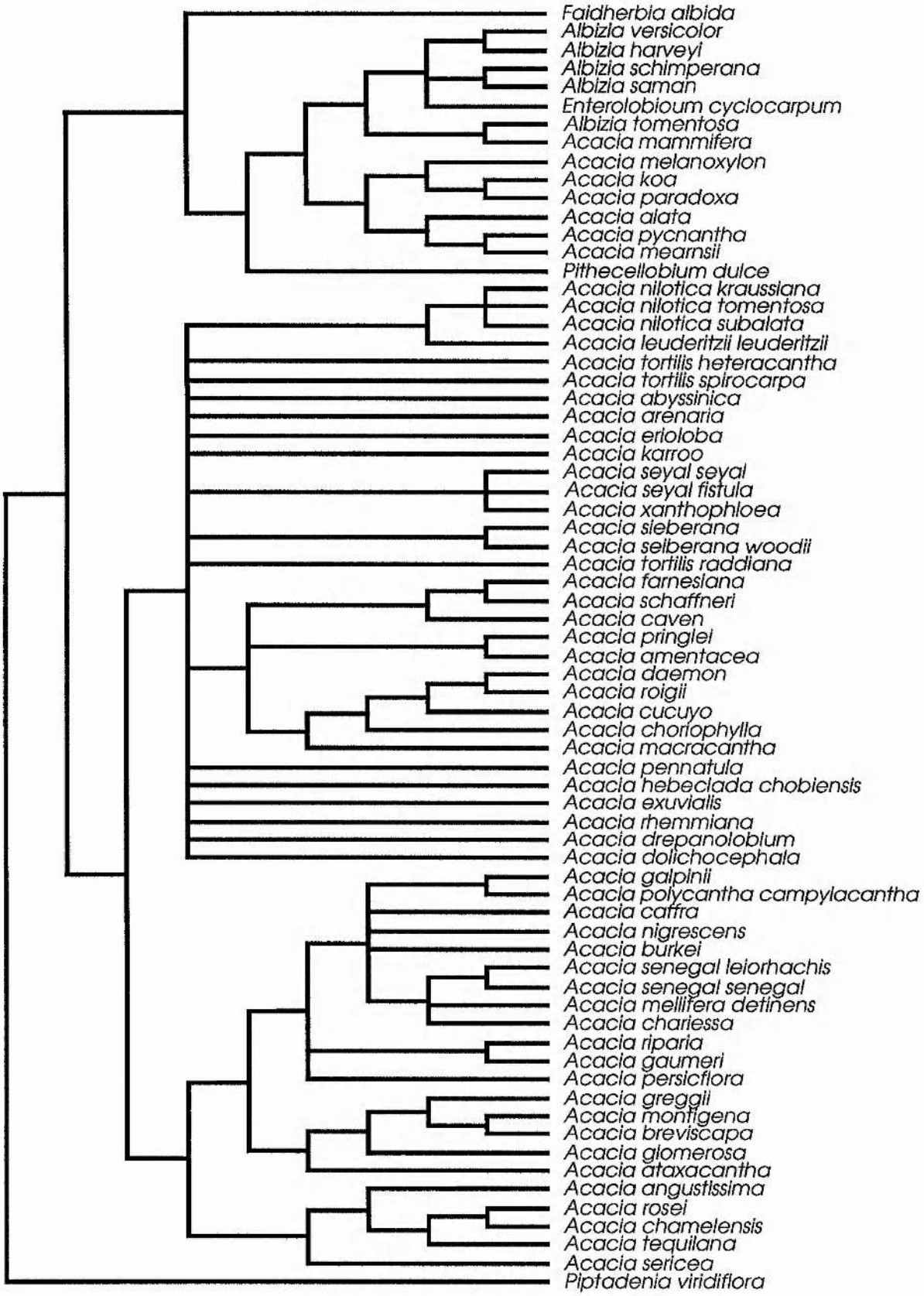


Figure 4.2. Strict consensus tree of 720 trees, length= 714 generated by the heuristic search of PAUP of the cpDNA restriction site data.

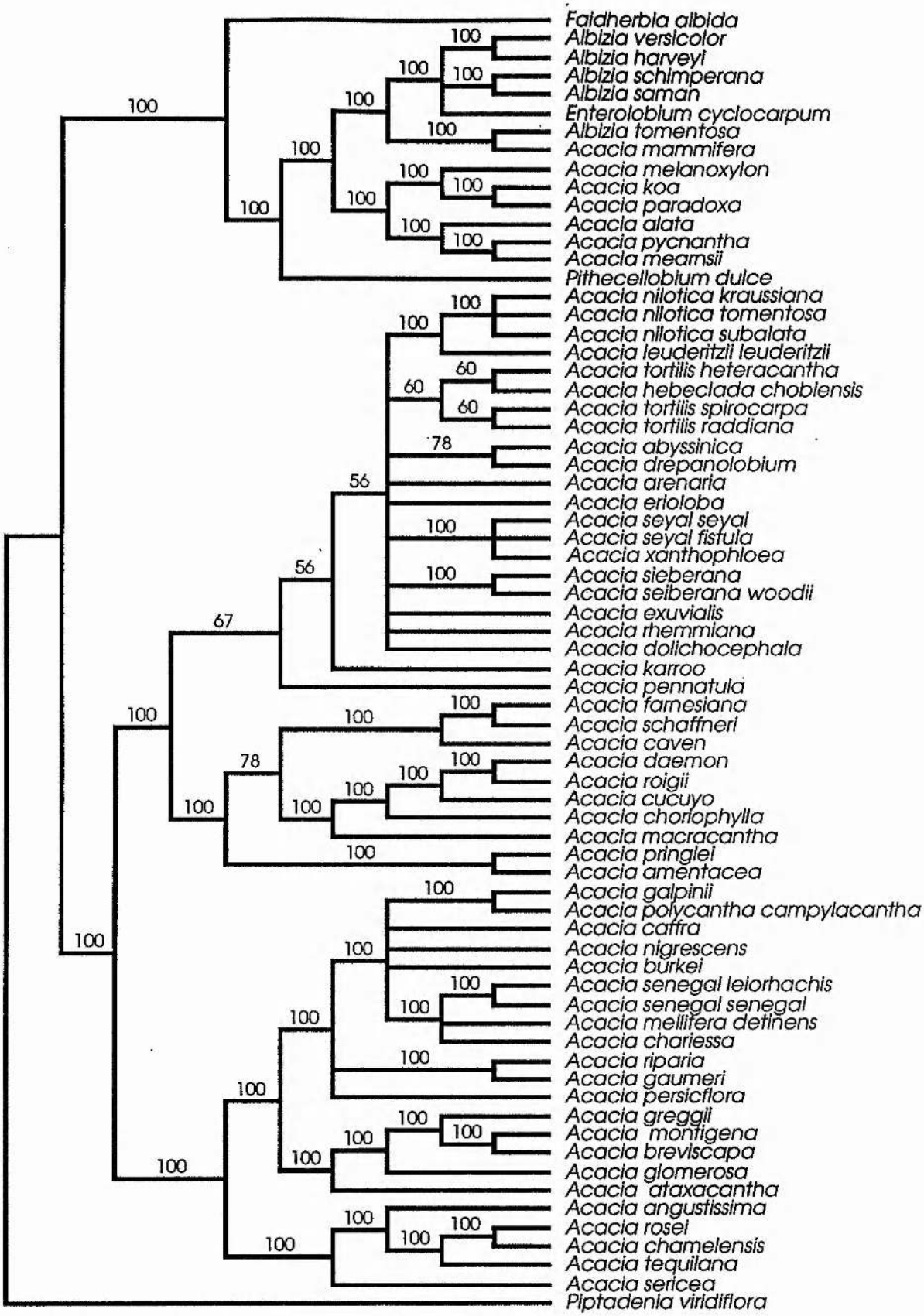


Figure 4.3. 50% majority rule consensus tree of 720 trees, length= 714 generated by the heuristic search of PAUP of the cpDNA restriction site data. The figures above the branches indicate the percentage of times that the branch was found in the 720 trees.

The decay index of the minimal length tree can be found in figures 4.4 to 4.8. Figure 4.4 shows the strict consensus tree of those trees that are less than or equal to 716 steps, 2 steps away from the minimal length tree. Figure 4.5 shows the strict consensus tree of those trees found that are less than or equal to 719 steps, 5 steps longer than the minimal length trees. Figure 4.6 shows the strict consensus tree of those trees found that are less than or equal to 724 steps, 10 steps longer than the minimal length trees. Figures 4.7 and 4.8 show the strict consensus and the majority rule consensus trees respectively of those trees found that are less than or equal to 734 steps, 20 steps longer than the minimal length trees. Due to computer memory limitations all of the trees that were found that were longer than the minimal length trees are a subset of all the possible trees of that length. The significance of the decay analysis will be discussed later.

From the strict and majority rule consensus trees of the 720 minimal length trees found we can identify three clades. All of these are supported by 100 % of the minimal length trees.

Clade one consisted of all the accessions of subtribe Ingeae investigated, all the accessions of subgenus *Phyllodineae* investigated plus *Faidherbia albida* and *Acacia mammifera*. This clade was not lost on any of the trees longer than the minimal length trees investigated.

Clade two contained all of the accessions of subgenus *Acacia* that were investigated. This clade was supported on all trees found up to 5 steps-or-less longer than the minimal length trees. It is also found in 99% of the trees that were less than or equal to 734 steps, 20 steps longer than the minimal length trees.

Clade three consists of all the accessions of subgenus *Aculeiferum* investigated, with the exception of *Acacia mammifera*. This clade is not well supported on any of the strict consensus trees up to 20 steps-or-less

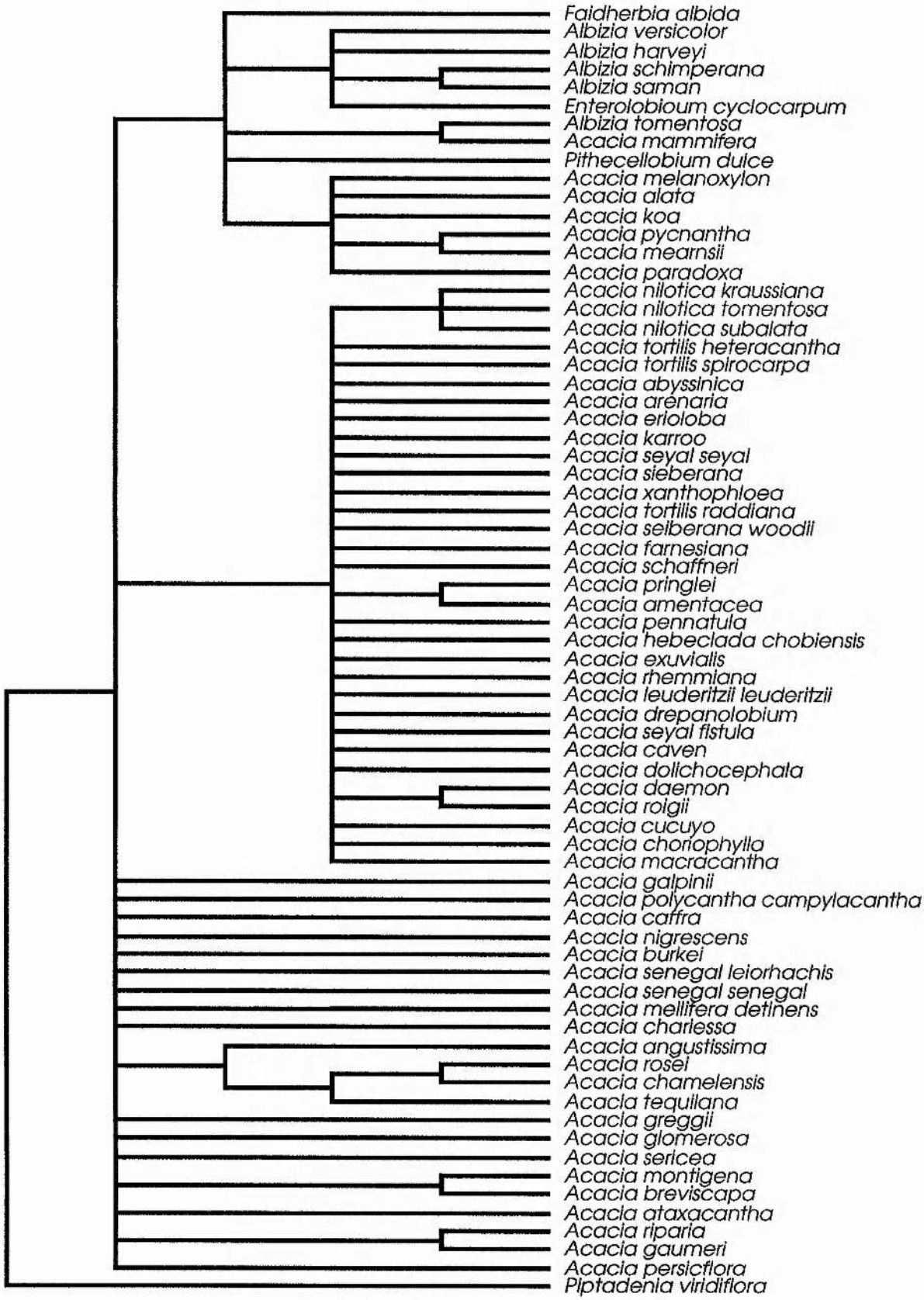


Figure 4.4. The strict consensus tree of the 5304 trees equal to or less than length 716, 2 steps longer than the minimal length trees.

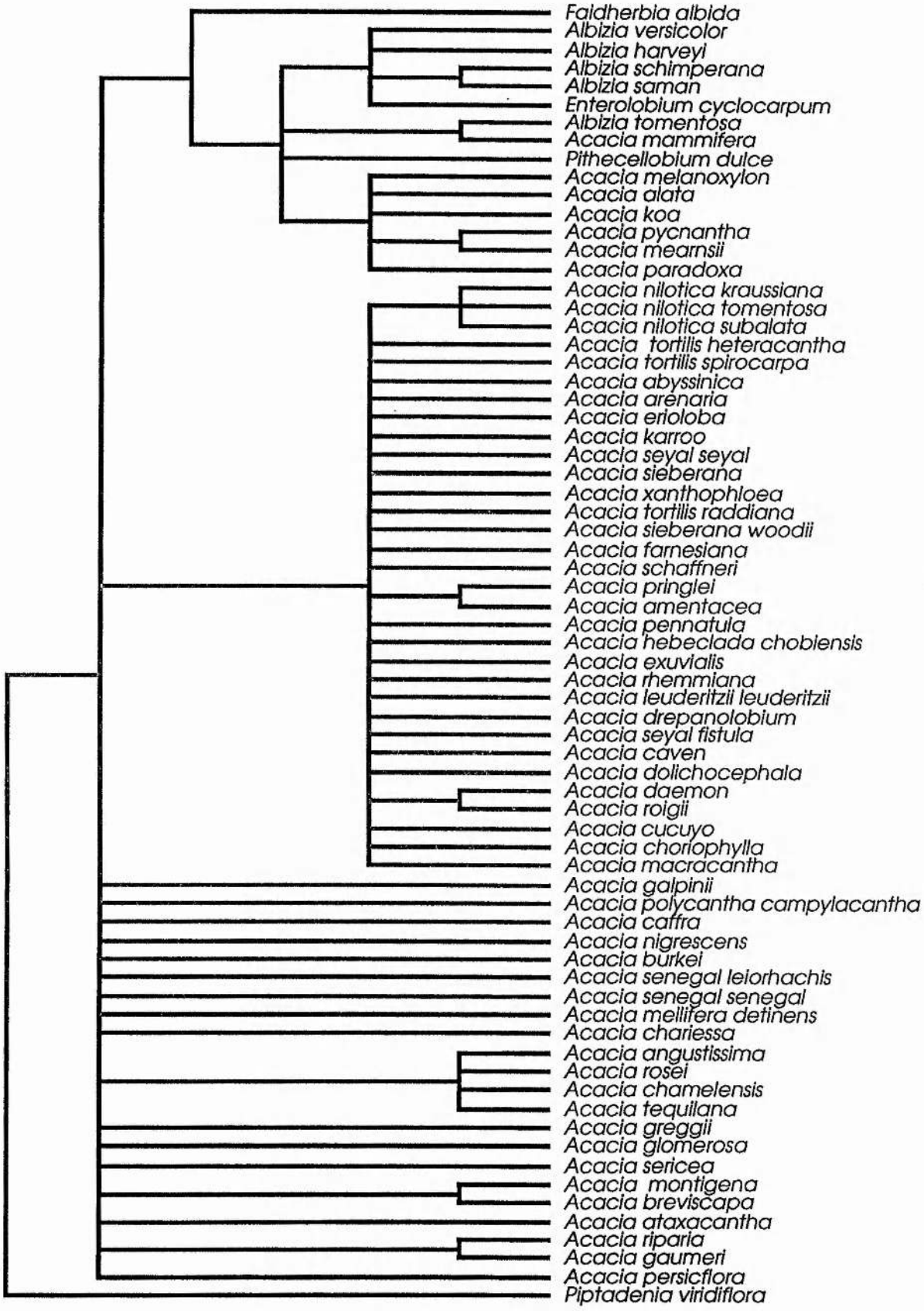


Figure 4.5. The strict consensus tree of the 7272 trees equal to or less than length 719, 5 steps longer than the minimal length trees.

Figure 4.6. The strict consensus tree of the 7199 trees equal to or less than length 724, 10 steps longer than the minimal length trees.

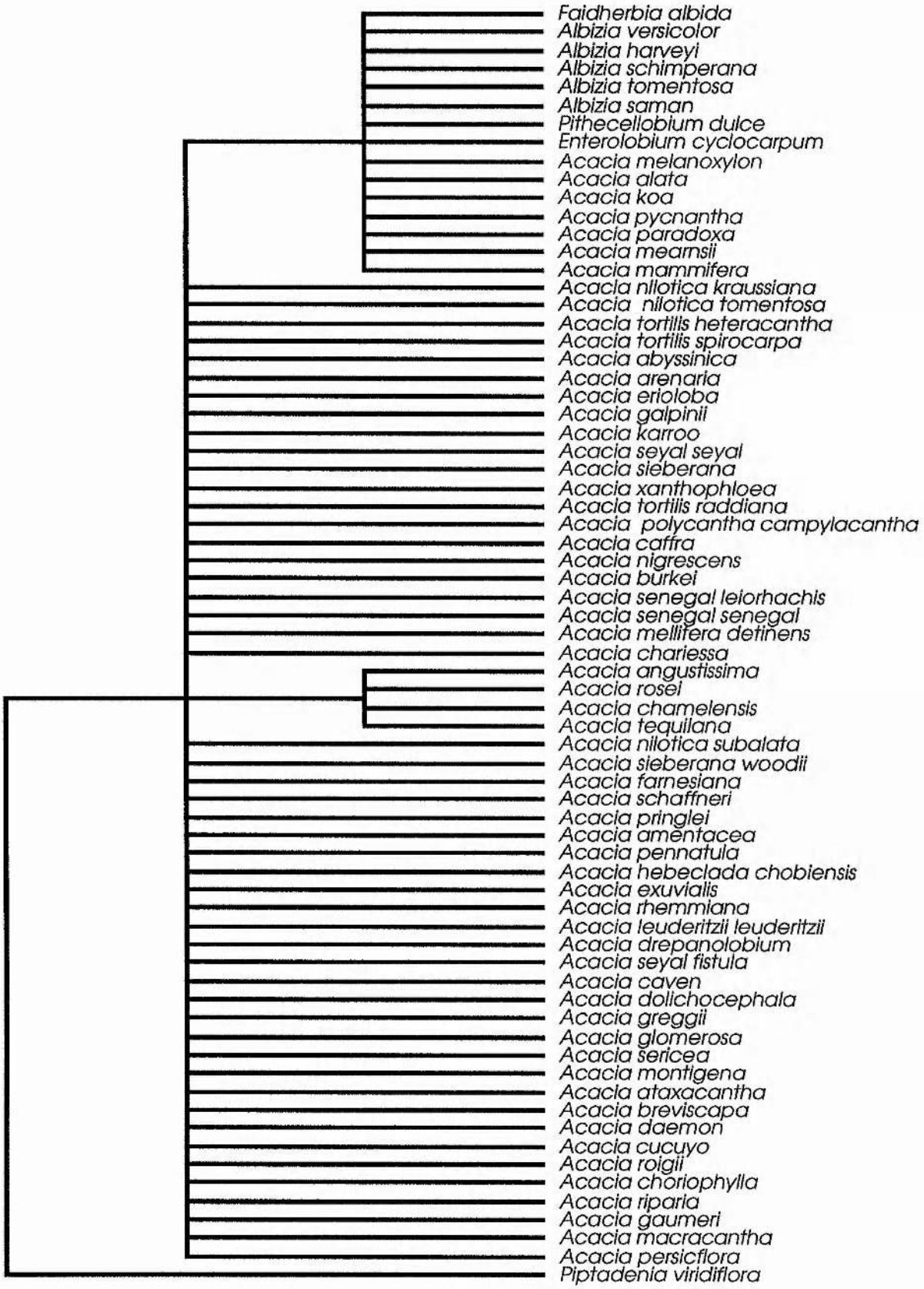


Figure 4.7. The strict consensus tree of the 7963 trees equal to or less than length 734, 20 steps longer than the minimal length trees.

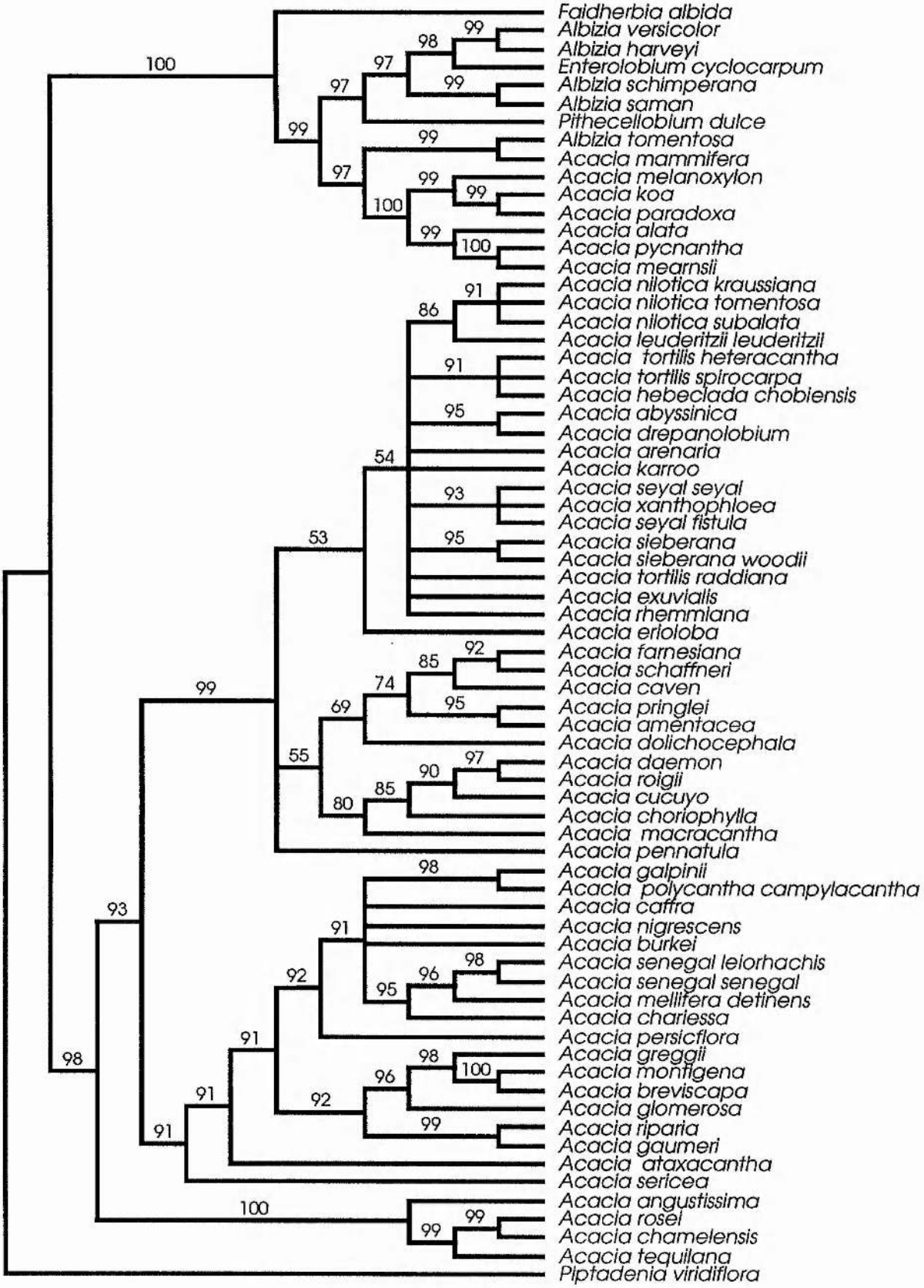


Figure 4.8. The majority rule consensus tree of the 7963 trees equal to or less than length 734, 20 steps longer than the minimal length trees.

longer than the minimal length trees. However, in the majority rule tree, derived from the trees found that were 20 steps-or-less longer, we can see species from section *Filicinae* separating out from this clade, basal to clades two and three.

4.5 Discussion.

The first topic for discussion must be robustness of the trees, i.e. is any tree an accurate representation of the data? The consistency index of 0.536 indicates that there are about twice as many character state changes as there would be if homoplasy was absent from the data set. How does this figure compare with other data sets? Sanderson and Donoghue (1989) researched 60 published data sets, investigating the relationship between consistency index, number of taxa, number of characters and the taxonomic rank of the investigation. They found that the consistency index was highly negatively correlated with the number of taxa studied. As the number of taxa studied increased the resulting C.I. of the phylogenetic tree returned decreased. Neither the number of characters nor the taxonomic level of the study was correlated with the C.I., in the studies Sanderson and Donoghue (1989) investigated. They suggested this was due partly to the increased probability of character state change as the total number of branches of the tree increased and also partly to the fact that the number of possible states for the character was limited. In restriction site occurrence analyses the number of character states is limited to two, the site either being present or absent. Sanderson and Donoghue (1989) calculated from the investigations they looked at a formula to investigate C.I. from the number of taxa investigated. Unfortunately, this returns a negative value for 71 taxa, due to inaccuracy in extending the formula beyond the range of data it was

based on. However, roughly calculating a C.I. directly from the graph (figure 1B in Sanderson and Donoghue (1989)) gives a value of approximately 0.3 for 71 taxa. Two studies with a similar, though lower, number of taxa to those investigated by Sanderson and Donoghue (1989) were those by Kellogg and Campbell (1987) who had a C.I. of 0.37 for 65 taxa and Stevens (pers. com. to Sanderson and Donoghue (1989)) who had a C.I. of 0.32 for 68 taxa. Very few other studies have been published in which a large number of taxa have been investigated. Comparison of my C.I. is therefore difficult. On the limited data available it does however appear that the C.I. obtained in this study is well above the value that would be expected as average for a study with the same number of taxa.

In addition to the C.I., the value returned for the retention index (R.I.), 0.831, is high. The retention index indicates how well the characters fit the tree that describes them. A value of 0.831 indicates that the characters fit the minimal phylogenetic trees found very well.

The decay indices also indicate that the tree found is robust. The three clades identified in the minimal length trees exist in over 90% of the trees that are 20 steps-or-less than the minimal length trees. The most robust clade is that of clade one. This consists of the subtribe Ingeae plus subgenus *Phyllodineae*, *A.mammifera* and *F.albida*. This clade occurs in all of the decay index trees studied. Next in terms of robustness is the clade containing the accessions of subgenus *Acacia* studied. This clade becomes unresolved only when trees 10 steps-or-less longer than the minimal length trees are investigated. Even when trees 20 steps-or-less longer are considered this clade is returned in 99% of them. Finally the third clade, that of subgenus *Aculeiferum* appears to be the least robust, becoming unresolved when we look at those trees 2 steps-or-less longer than the minimal length trees. If we look at trees that are 20 steps-or-less longer than the minimal length tree, we can see from the majority rule

consensus tree that section *Filicinae* is basal to the combined clades of subgenus *Acacia* and subgenus *Aculeiferum*. It is the instability of the section *Filicinae* clade which is probably causing the subgenus *Aculeiferum* clade to become unresolved at only 2 steps longer than the minimal length tree.

To summarise, therefore, it would appear that the trees are indeed reliable. This can be seen from the higher than average C.I. value and high R.I. value for 71 taxa. The clades identified on the consensus trees derived from the minimal length trees also appear robust. This is shown by the high percentage of trees that are '20 step-or-less longer than the minimal length trees' which these clades occur in.

The next question to address is the lack of computing power which unfortunately meant that the weighted analysis could not be completed. This is not a shortcoming of the computer used for the analysis, but rather a reflection of the complexity and amount of data that had to be analysed. It appears that D.Swofford is extending PAUP so it will able to be run on many of the mainframe computers that Universities possess. This release will be eagerly awaited as many data sets similar to the one in this study will become analysable.

Having established the legitimacy of the data, it is now possible to examine the implications of the trees found for i) the relationships of the subgenera of *Acacia* to each other and the other taxa studied and ii) the relationships of taxa within each of the subgenera of *Acacia*.

As mentioned earlier we can identify from the strict consensus tree three clades. Two of these clades relate to subgenera *Acacia* and *Aculeiferum*. If for the moment one disregards *A.mammifera* which will be discussed later, then we can see that both subgenera are monophyletic.

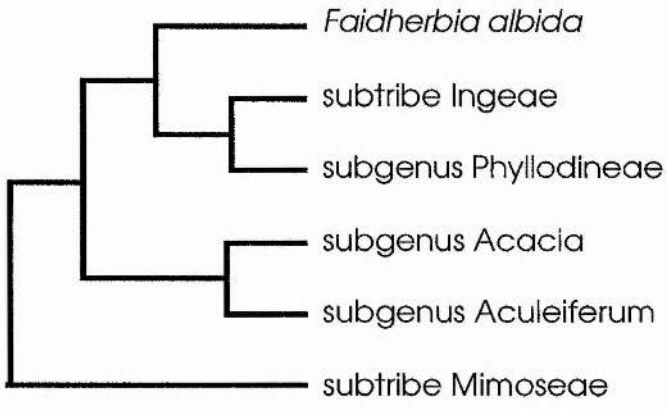


Figure 4.9. This figure shows a simplified version of the strict consensus tree of figure 4.2.

Within the other clade subgenus *Phyllodineae* is also monophyletic. The monophyly of all three subgenera of *Acacia* allows us to represent the strict consensus tree of the 720 minimal length trees more simply. This can be seen in figure 4.9.

What are the implications of the results? The monophyly of each of the subgenera has never been disputed. Each subgenus is clearly defined by a wide range of morphological and chemical characters, so it is not unexpected that each subgenus has been found to be monophyletic. The position of section *Filicinae* from subgenus *Aculeiferum*, has been a matter of discussion. This study suggests that this section could be basal within subgenus *Aculeiferum* and so will be discussed later in the context of the infra-subgeneric relationships of subgenus *Aculeiferum*.

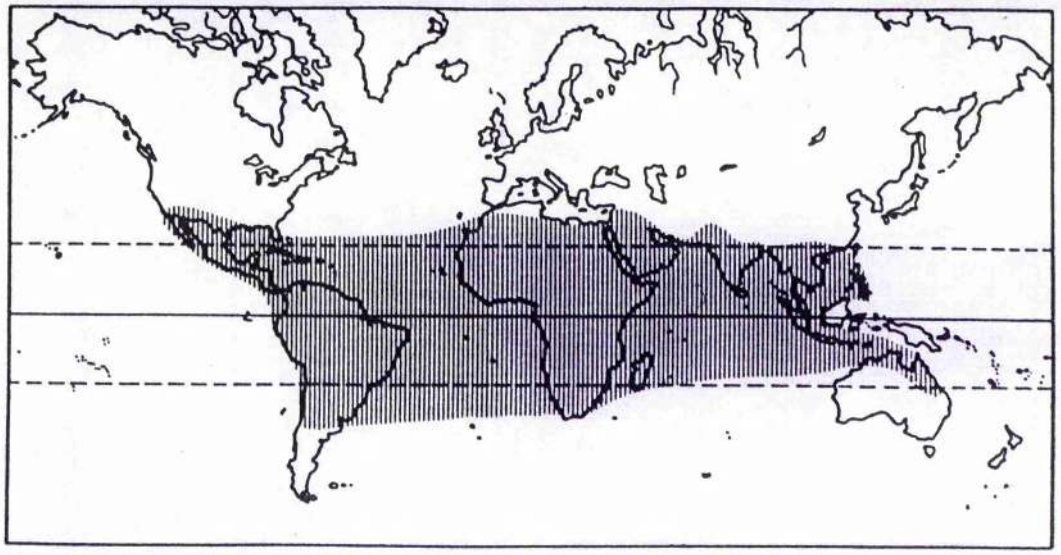
If we now look at the relationships of the higher level taxa to each other, we can observe several implications for the phylogeny of *Acacia sens. lat.* Firstly it appears that *Acacia sens. lat.* may not be monophyletic, i.e. the three subgenera do not appear to share a common ancestry. This can be seen in figure 4.9. Here we can see that subgenus *Acacia* and subgenus *Aculeiferum* appear closely related, being grouped together in the cladogram. No other taxa investigated seem to be closely related to these two subgenera. Subgenus *Phyllodineae* is not close to these two

subgenera, but rather it appears to be related to taxa within the Ingeae, coming out within the tribe Ingeae. That subgenus *Phyllodineae* is within the Ingeae is not obvious from figure 4.9, but if we look at figure 4.2 we can see that species from subgenus *Phyllodineae* come out between *Pithecellobium dulce* and *Albizia* spp. Due to the lack of reference taxa within the Ingeae (this analysis has looked at only 3 genera from a possible 17 genera (Neilsen, 1981)) the precise relationships of this subgenus are hard to define. The relationships of the taxa within the Ingeae will be discussed shortly. Finally the position of *Faidherbia albida* can be seen from figures 4.9 and 4.2, it appears to be basal to the tribe Ingeae and subgenus *Phyllodineae*. Both subgenus *Acacia* and subgenus *Aculeiferum* are distributed from the New World, through Africa to S.E. Asia. Subgenus *Phyllodineae* on the other hand is mainly restricted to Australia (see Map 4.1 overleaf). The results of this investigation indicate that the two subgenera of *Acacia* which share the same distribution are closely related. The distribution of taxa within the subtribe Ingeae is pantropical, with a few genera being represented in all continents.

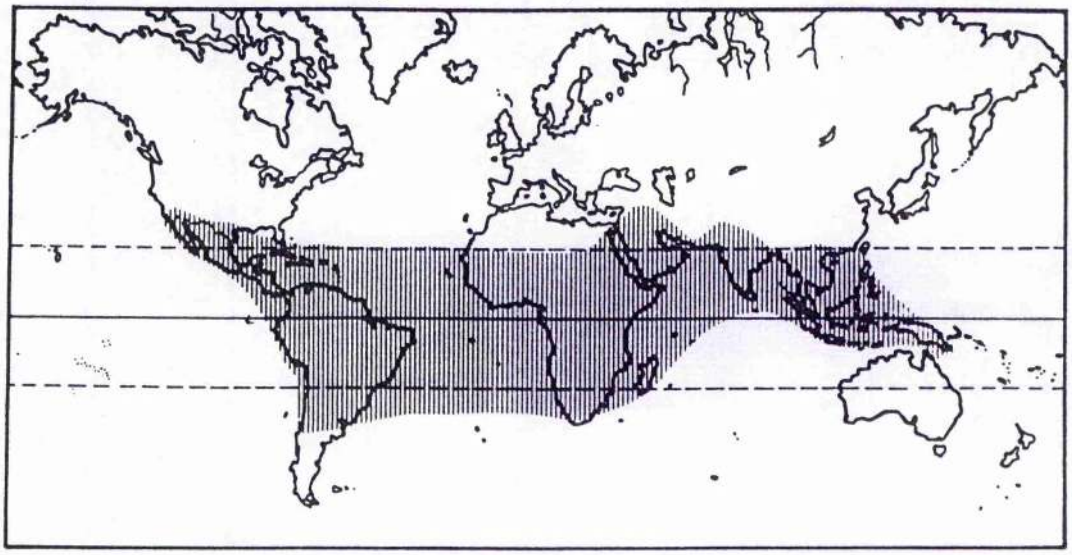
The indication that *Acacia sens. lat.* is not monophyletic, is not unexpected. As detailed in Chapter 2, two recent attempts at a classification of *Acacia sens. lat.* have suggested that it is not monophyletic. However, the conclusions of these previous investigations differ from the results obtained in this one.

Pedley[1] (1986) suggested that subgenus *Acacia* and subgenera *Phyllodineae* and *Aculeiferum* were derived from different lineages within the Ingeae. In addition Pedley (1986) suggested that subgenus *Acacia* had some affinity with the genus *Pithecellobium*, and subgenera

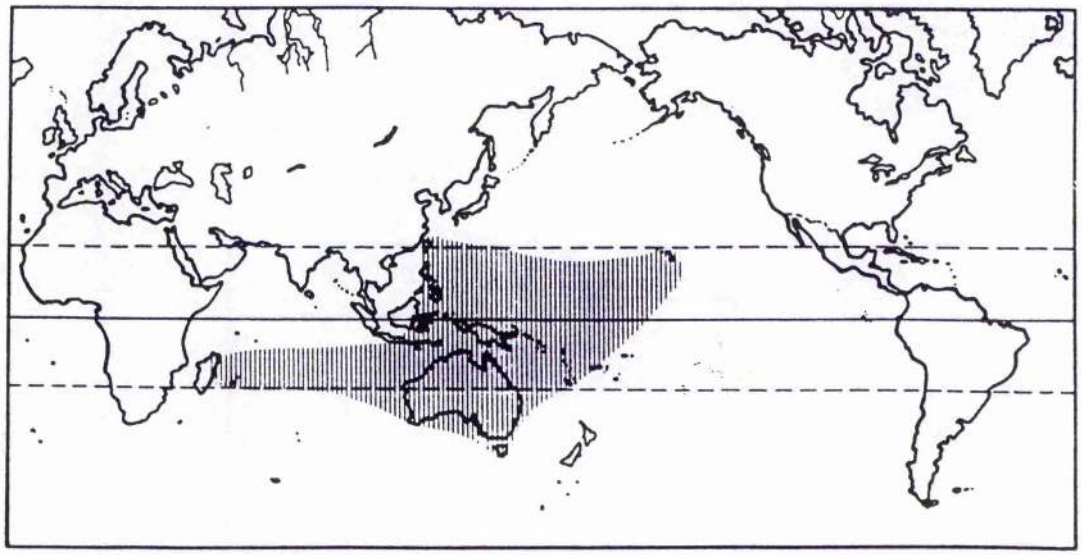
[1] Pedley (1986) promoted the three subgenera of *Acacia* to genera. To clarify matters, in this discussion I will use the subgeneric names in place of Pedley's generic names. These are; subgenus *Acacia* for *Acacia*; subgenus *Aculeiferum* for *Senegalia* and subgenus *Phyllodineae* for *Racosperma*. See Chapter 2 for further details.



a) The distribution of subgenus *Acacia* (excluding *A.farnesiana* in Australia).



b) The distribution of subgenus *Aculeiferum*.



c) The distribution of subgenus *Phyllodineae*.

Map 4.1 a), b), and c). This series of maps show the global distribution (hatched area) of the subgenera of *Acacia*. From Ross (1981).

Aculeiferum and *Phyllodineae* had affinities with *Calliandra*. Pedley (1986) concluded that subgenus *Aculeiferum* and subgenus *Phyllodineae* were closely related, both being clearly distinguished from subgenus *Acacia*.

The cladistic study of Chappill and Maslin (1995) had similar conclusions concerning the monophyly of *Acacia sens. lat.* and the relationships of the subgenera to each other. From the cladogram they produced (figure 2.1, and also in condensed form in figure 4.10) it can be seen that *Acacia sens. lat.* is not monophyletic. Subgenus *Acacia* is closely related to *Calliandra* and also *Pithecellobium*. Subgenera *Aculeiferum* and *Phyllodineae* are closely related to each other basal to the tribe Ingeae and subgenus *Acacia*.

Apart from the polyphyly of *Acacia sens. lat.*, the results obtained from the present chloroplast DNA restriction site investigation bear little relationship to the results of either Chappill and Maslin (1995) or Pedley (1986). Comparison with Pedley's results cannot be attempted, as his method of generating the phylogeny and classification is not directly comparable to the methods I have used, i.e. Pedley (1986) analysed his data subjectively without any clearly defined criteria for assessing relationships.

Chappill and Maslin (1995) view their investigation as a "preliminary analysis". There is a conflict between their results of the analysis of the subfamily Mimosoideae and their results from an infra-generic analysis of the tribe Acacieae. This, in addition to the unusual relationships of taxa within the Ingeae they postulated, means that few firm conclusions can be drawn from their analysis at present. The addition of much of the missing data from both analyses would help resolve relationships.

It is, however, possible to compare the preliminary results of Chappill and Maslin (1995) with the results of the cpDNA analysis presented here. First I will compare results for the tribe Ingeae.

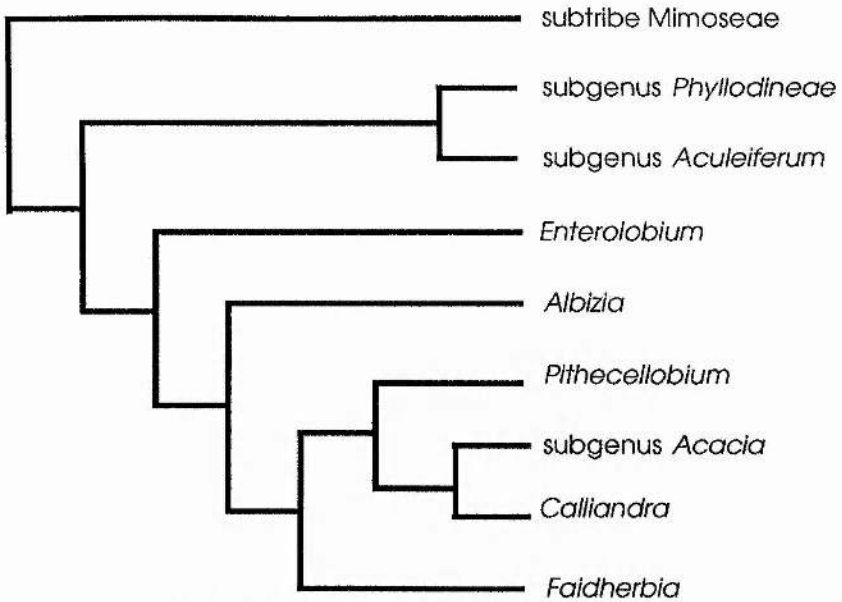


Figure 4.10. This figure shows a simplified version of cladogram in figure 2.1. The cladogram in figure 2.1 is taken from Chappill and Maslin (1995). In this figure only taxa relevant to this study have been kept, the other taxa have been 'pruned' off.

From the strict consensus tree (figure 4.2) it can be seen that in the present investigation *Faidherbia albida* is basal to the tribe Ingeae and could possibly be united with it. Chappill and Maslin (1995) suggested that the genus *Faidherbia* was either related to the genus *Wallaceodendron* within the Ingeae (figure 2.1) or that it was basal to subgenus *Phyllodineae* and some *Aculeiferum*, yet within the genus *Acacia sens. lat.* (see figure 2.2). Neither of these positions for *F.albida* had been suggested before. The genus *Faidherbia* is usually seen as either basal to the genus *Acacia sens. lat.* yet still within the tribe Acacieae (e.g. Guinet and Vassal 1978), or as belonging within the tribe Ingeae (Guinet 1990). The results of the present analysis give weight to Guinet's (1990) suggestion, although whether to include *Faidherbia* within the Ingeae remains questionable as only a few taxa from the Ingeae have been studied.

Within the Ingeae the relationships of the few genera analysed in the present study again present a contrast to the relationships of these taxa in the cladogram of Chappill and Maslin (1995, see figure 2.2 and

4.10). The importance of the genus *Pithecellobium* between the two phylogenies is difficult to assess, due to the lack of reference taxa in the Ingeae in the present analysis. However, the relationships of the two other taxa I investigated contrast with those of Chappill and Maslin (1995). The cladogram of Chappill and Maslin (1995) suggests that *Albizia* and *Enterolobium* are not closely related (see figure 4.10). The results of my cpDNA analysis suggest they are very closely related. In fact my analysis suggests *Enterolobium* is within the genus *Albizia* (see figure 4.11). This agrees with Neilsen (1981) who in his description of taxa within the Ingeae wrote that *Enterolobium* is "hardly distinguishable from some of the indehiscent *Albizia* groups". His reason for not advocating transfer of *Enterolobium* to *Albizia* was probably due to widespread cultivation of *Enterolobium* spp. for timber production.

The results obtained in the present investigation concerning the relationships of the subgenera of *Acacia*, are again in contradiction to the results of Chappill and Maslin (1995), as well as the conclusions of other previous investigations (e.g. Pedley, 1996; Guinet and Vassal, 1972; Guinet, 1990). My results suggest that subgenus *Acacia* and subgenus *Aculeiferum* are closely related to each other. Neither subgenus is closely related to any genera in the tribe Ingeae. The other subgenus, subgenus *Phyllodineae*, appears not to be closely related to the other subgenera of

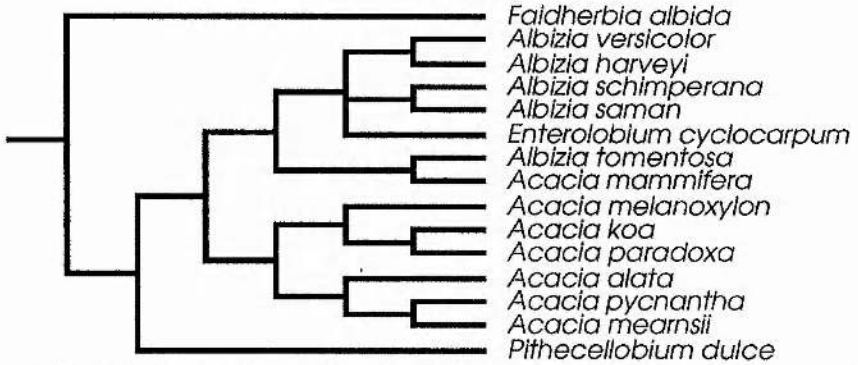


Figure 4.11. This figure shows the clade from the strict consensus tree (figure 4.2) pertaining to subgenus *Phyllodineae*, *Faidherbia albida* and species from the Ingeae.

Acacia, rather it appears to be related to genera within the Ingeae. The results of Chappill and Maslin (1995) suggest a close relationship between subgenus *Aculeiferum* and subgenus *Phyllodineae*. Their results also suggest that subgenus *Acacia* is not closely related to either subgenus, but rather belongs in the Ingeae (see figure 4.10). The close relationship between subgenus *Aculeiferum* and subgenus *Phyllodineae* is one which has been suggested in the majority of previous attempts at classifications, as well as in numerous observations on the genus (e.g. Pedley, 1986 and refs. therein; Guinet, 1990; Brain, 1987 and 1990; Conn *et al.*, 1989). This hypothetical relationship between the two is based on a series of shared morphological and chemical characteristics. Similar pollen morphology and the types of free amino acids found in the seed are the main characters which distinguish subgenera *Aculeiferum* and *Phyllodineae* from subgenus *Acacia* (Pedley, 1986; Guinet, 1990; Chappill and Maslin, 1995). Table 4.12 lists these differences and similarities.

It is the pollen characters which provide the strong link between subgenus *Aculeiferum* and subgenus *Phyllodineae*. Both subgenera share the absence of columellae with either a porate or extraporate aperture. Taxa from subgenus *Acacia* have columellae with a colporate aperture. If we look at these characters in the Ingeae we can see that all the possible pollen characters of the Acacieae are also present. This suggests that some of these characters have arisen at least twice during the evolution of the Mimosoideae.

The free amino acids of the seeds offer less satisfactory evidence of a close relationship between subgenus *Aculeiferum* and subgenus *Phyllodineae*. They do illustrate however that subgenus *Acacia* is different from subgenus *Aculeiferum* and subgenus *Phyllodineae*. The amino acid characters which link subgenus *Aculeiferum* and subgenus *Phyllodineae*

Character	subgenus *Acacia*	subgenus *Aculeiferum*	subgenus *Phyllodineae*	tribe Ingeae
Pollen				
-aperture type	Colporate	Porate (infrequently extraporate)	Extraporate (infrequently porate)	Colporate in *Calliandra sens. str.* Porate in the remainder of the Ingeae except extraporate in some species of *Cojoba*, *Marmaroxylon*, *Obolinga* and *Zygia*
-columellae	Present	Absent	Absent	Present in *Calliandra sens.* str. Absent in the remainder of the Ingeae.
-exine ornamentation	Smooth	Smooth	Reticulate (rarely areolate)	Areolate or sometimes smooth
Free amino acids in seeds				
- s-carboxethyl cysteine	Absent	Polymorphic	Polymorphic	Polymorphic
- s-carboxiso propylcysteine	Absent	Polymorphic	Polymorphic	Polymorphic
-albizzine	Absent	Polymorphic	Polymorphic	Polymorphic
- willardine	Absent	Absent	Polymorphic	Absent
- alphabeta-diaminopropionic acid	Absent	Polymorphic	Polymorphic	Polymorphic
- djenkolic acid	Absent	Polymorphic	Polymorphic	Absent
- N-acetyl djenkolic acid:	Absent	Present	Present	Present
- pipecolic acid	Absent	Polymorphic	Polymorphic	Absent
- 4-OH-pipecolic acid	Present	Polymorphic	Present	Polymorphic
- 5-OH-pipecolic acid	Absent	Absent	Polymorphic	Polymorphic
- 2,4-cis-4,5-trans-diOH pipecolic acid	Absent	Polymorphic	Polymorphic	Absent

Table 4.12. This table shows the distribution of pollen characters and seed free-amino acid characters which are thought to closely unite subgenus *Aculeiferum* and subgenus *Phyllodineae*. The data for the pollen characters have been taken from Table 2 in Maslin (1988), the free amino acids of the seed data were taken from the data matrix for the infrageneric analysis of Chappill and Maslin (1995). The character attributes for the pollen data are self explanatory. For the amino acid data polymorphic refers to the condition where the amino acid is present in some of the species of the taxa but not in others of the same taxa e.g. Willardine is found in section *Monacanthea* and *Filicinae* of subgenus Aculeiferum but not in section *Aculeiferum*. In addition to this the Ingeae descriptions for the amino acids are only based on the genera *Calliandra*, *Havardia* and *Paraserianthes* as Chappill and Maslin only included data from these genera in their infrageneric analysis.

would appear to have evolved more than once during the evolution of the tribes Acacieae and Ingeae.

During the evolution of both the pollen characters and amino acid characters several character states have arisen more than once, for example, the presence of albizzine in the seeds. It is absent from subgenus *Acacia*, but is present in some, but not all taxa from subgenus *Aculeiferum*, subgenus *Phyllodineae* and subtribe Ingeae. The wide

spectrum of otherwise unrelated taxa that have albizzine present in their seeds effectively rules out a single origin of albizzine. This is true whichever phylogeny you believe is correct. The possibility therefore exists that the similarities between subgenus *Aculeiferum* and subgenus *Phyllodineae* as suggested by these characters are false synapomorphies due to parallel evolution or convergence, rather than a common shared ancestry. If this hypothesis is correct then there can be no objection to the unrelated positions of subgenus *Aculeiferum* and subgenus *Phyllodineae* suggested by this cpDNA analysis. When the analysis of Chappill and Maslin (1995) is completed, the phylogeny of the Mimosoideae as suggested by morphological characters will be more clear. It may be that the relationships of the taxa as suggested by Chappill and Maslin (1995) will be reinforced, however the opposite is also possible- i.e. a new set of relationships will be suggested. Until that time further comparisons between the results of the two data sets are impossible , and we must proceed on the premise that the cpDNA results are a good 'best approximation' to the phylogeny of Acacieae, and may be more reliable than the morphological data.

Before we proceed to the taxonomic and biogeographic implications of the cpDNA results, an alternative to deciding which phylogeny is 'correct' will be discussed. This is to suggest that both phylogenies are 'correct'. The competing phylogenies (the one presented here and the analysis of Chappill and Maslin (1995)) have been investigated using different characters. The morphological characters represent the evolution of the morphological features of the Mimosoideae, and the cpDNA characters represent the evolution of the cpDNA molecule. Neither phylogeny shows the true phylogeny, each being an approximation to it.

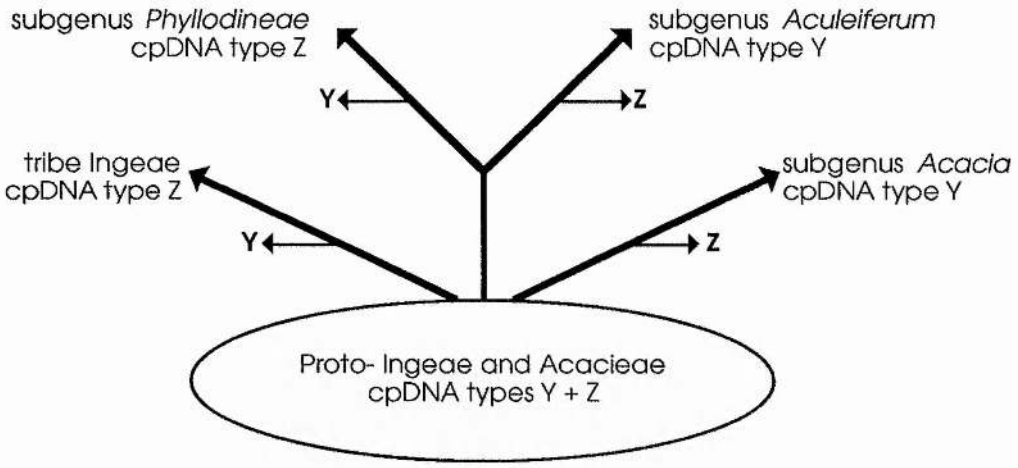


Figure 4.13. This figure represents the possible evolutionary paths of the cpDNA molecules of the proto Ingeae and Acacieae. See text for details.

It should be recognised that a scenario can be envisaged which would explain the apparent dichotomy of the morphological and molecular results. In the proto- Ingeae and Acacieae there possibly existed a range of polymorphic cpDNA types. As taxa began to evolve from this group they perhaps still had the original cpDNA polymorphisms. As time progressed through random losses and lineage sorting in some of the taxa, the cpDNA polymorphisms were lost and one type of cpDNA became fixed.

Consider the example in figure 4.13. It could be postulated that within the proto Ingeae and Acacieae there existed only two cpDNA types, Y or Z. Subgenus *Acacia* evolved from taxa derived from this proto tribe. This lineage either lost cpDNA type Z during its evolution or it only ever had cpDNA type Y. This is referred to as 'lineage sorting'. Similarly the Ingeae could be derived from taxa which lost the Y type of cpDNA or never had it. Subgenus *Phyllodineae* and subgenus *Aculeiferum* could be derived from taxa in which the chloroplast DNA was polymorphic, the ancestral states Y + Z being present. It was only after subgenus *Phyllodineae* and subgenus *Aculeiferum* had become differentiated that lineage sorting occurred. Subgenus *Phyllodineae* lost the Y type of

cpDNA, and subgenus *Aculeiferum* lost the Z type of cpDNA. The result of this sorting of cpDNA types would be that although subgenera *Aculeiferum* and *Phyllodineae* share a common ancestry their cpDNA types are now dissimilar. In fact, in this example, the cpDNA of these two subgenera are similar to other taxa with which they have no close common ancestry.

With this hypothesis there need be no conflict between the morphological data and the cpDNA data. Each set of data reveals a different aspect of the evolution of the Mimoseae, the morphological data reveal the relationships of the taxa studied, and the cpDNA reveals how the chloroplast DNA in the Mimoseae has evolved. A possible way to investigate this would be to construct a phylogeny based on nuclear DNA characters. At present, a popular nuclear DNA character is the sequence of nuclear ribosomal DNA. Nuclear DNA is not subject to the same evolutionary conditions as cpDNA and so could perhaps offer an independent viewpoint on the problems of the cpDNA-vs-morphological derived phylogenies.

4.6 Infrasubgeneric variation.

4.6.1 Subgenus *Acacia*.

Figure 4.14 shows the clade relating to subgenus *Acacia*, taken from the strict consensus tree (figure 4.2). As mentioned before, these results show subgenus *Acacia* to be monophyletic, all the accessions of subgenus *Acacia* studied being present in this clade. The clade of subgenus *Acacia* is poorly resolved. Within this clade accessions from the New World form a coherent group. The accessions from Africa, apart from several species groups, are relatively undifferentiated. One of the reasons

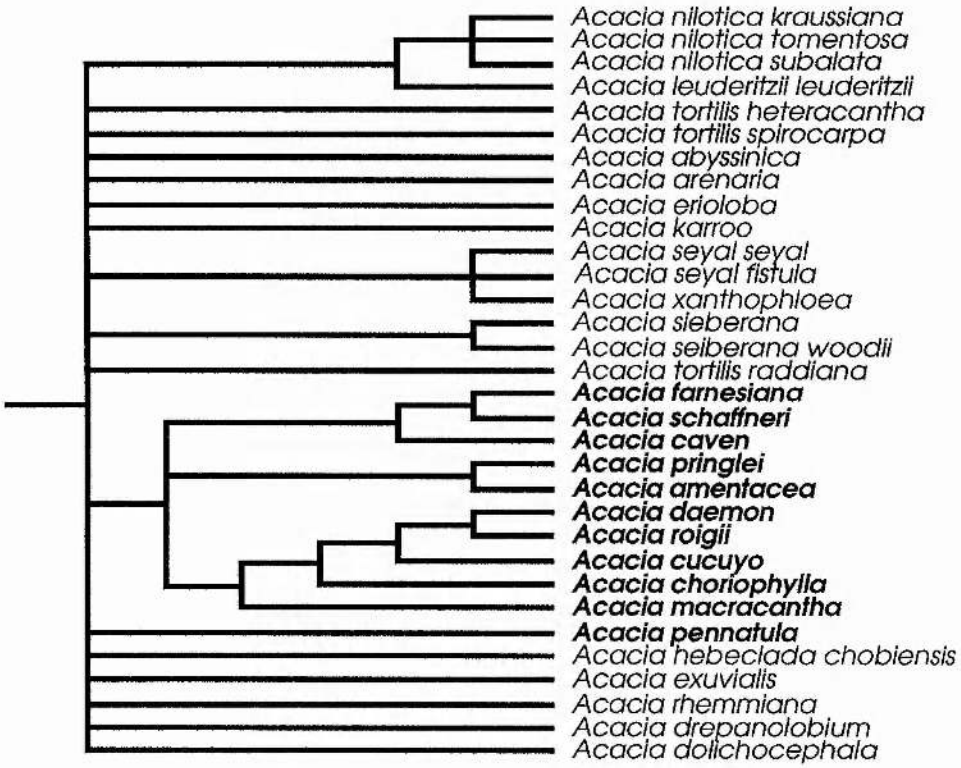


Figure 4.14. This figure shows the clade containing members of subgenus *Acacia* from the strict consensus tree figure 4.2. Accessions from the New World appear in bold typeface.

for this is the lack of variation encountered in African species of this subgenus, e.g. *A.exuvialis* and *A.rhemmiana* have the same cpDNA restriction site patterns, and cannot be distinguished with the characters I have used. In addition many of the differences between species that I observed were autapomorphic and these differences do not contribute to defining species relationships.

The species within the New World grouping of taxa from subgenus *Acacia* are fully differentiated. The only New World accession that was studied but which does not form part of this group is *A.pennatula.* This species is in the undifferentiated part of the subgenus *Acacia* clade. It is difficult to suggest reasons for this.

Previous attempts at subdividing subgenus *Acacia*.

In his 1875 review of the suborder Mimoseae Bentham divided subgenus *Acacia* (he had called it series Gummiferae) into three subseries; Summibracteatae, Medibracteatae and Basibracteatae. The primary discriminating character was the position of the involucel on the peduncle. Ross (1979) reported that this character was variable and was not "suitable for delimiting major groups" within subgenus *Acacia*. Britton and Rose (1928) divided the American species of subgenus *Acacia* (they did not consider the African species in this subgenus) into twelve genera. Although many of the genera of Britton and Rose comprised distinctive groups of species, this 'splitting' of the genus was generally considered excessive. Vassal (1972) divided subgenus *Acacia* into two smaller groupings of species; subsection Pluriseriae and subsection Uniseriae. The Pluriseriae were characterised by having seeds in two or three series within the pod, and the Uniseriae were characterised by having the seeds in one series (Ross, 1979). Most of the species in Africa belong to the Uniseriae, with the exception of *A.erioloba* and *A.farnesiana* (a doubtful native species in Africa). Further ways of dividing the genus are discussed by Ross (1979), though at present there are no satisfactory ways of dividing subgenus *Acacia* in Africa (Maslin and Stirton, in press).

The results of the cpDNA analysis do not provide any evidence for the division of subgenus *Acacia* into sections or other supraspecific taxa at present. They do, however, suggest that the American species of subgenus *Acacia* are distinct from the African species. This can be clearly seen in figure 4.3, the majority rule tree where all but one, *A. pennatula*, of the American species investigated form a distinctive clade.

New World Taxa.

The species of subgenus *Acacia* in the New World can be grouped together into five "reasonably distinct species groups" (Maslin and Stirton, in press). The first group Maslin and Stirton (in press) called the '*A.rigidula*' group. From this group I analysed the species *A.pringlei* and *A.amentacea*. These species group together in the cladogram produced (see figure 4.15). Morphologically they are species which have leaves with one pair of pinnae; whitish, four-parted, reflexed to erect corolla lobes and elongated spicate inflorescences (Lee *et al.*, 1989).

The second group suggested by Maslin and Stirton (in press) was the '*A.constricta*' group. No species from this group were analysed. For more information concerning this group see Clarke *et al.* (1990)

The third group of species was the '*A.farnesiana*' group. Species from this group which were analysed were *A.farnesiana*, *A.schaffneri* and *A.caven*. On the cladogram produced these species are closely related, grouping together on the upper clade (see figure 4.15). Morphologically these species have 2-10 pairs of pinnae; a small petiolar gland and one or more rachis glands. The inflorescences are globose subtended closely by an involucre.

The fourth group of species of subgenus *Acacia* in the New World referred to by Maslin and Stirton (in press) was a group they called the '*Acacia daemon* group'. These species are restricted to the Caribbean

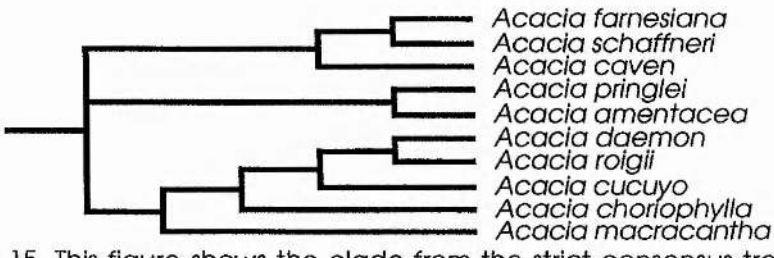


Figure 4.15. This figure shows the clade from the strict consensus tree (figure 4.2) which relates to the species of subgenus *Acacia* investigated from the New World, with the exception of *A. pennatula*.

and are characterised by low stamen number, characteristic fruit structure and a ramified spine system (Maslin and Stirton, in press). From this group four species were analysed, *A.daemon*, *A.choriophylla*, *A.cucuyo* and *A.roigii*. (Maslin and Stirton (in press) did not mention *A.cucuyo* and *A.roigii*, but they both belong to the *A.daemon* group (L.Rico RBG Kew, pers. com.) All four species that were analysed from this group are grouped together on the cladogram (figure 4.15).

The fifth and final group of species, is the Ant-Acacia group (Maslin and Stirton, in press). These are species of *Acacia* which form symbiotic relationships with ants from the genus *Pseudomyrmex*. No Ant-Acacias were investigated in this study. Further information can be found in Ebinger and Siegler (1992). Maslin and Stirton (in press) also included three non-Ant-Acacias in this group, *A.macracantha*, *A.pennatula* and *A.cochliacantha*. Two of these, *A.macracantha* and *A.pennatula* were included in this study. As mentioned earlier *A.pennatula* does not group with the New World species that have been analysed. The reasons for this are unexplained at present, though this accession does appear to lack several of the cpDNA restriction sites which characterise the New World species of subgenus *Acacia*. Whether this accession reflects the species as a whole is a subject which needs further investigation. Since *A.pennatula* does not group with New World species of subgenus *Acacia*, at present there does not appear to be a close relationship between *A.pennatula* and *A.macracantha* as suggested by Maslin and Stirton.

African Taxa

Maslin and Stirton concluded that there were no satisfactory ways of subdividing subgenus *Acacia* in Africa. The cpDNA data was likewise not able to differentiate the accessions analysed into groups of species. This was probably due to the lack of cpDNA restriction site variation in the

taxa from Africa. Many species had similar, or identical, restriction fragment profiles. The only variation that exists has differentiated taxa from each other. On the strict consensus tree there are three groups of taxa that are resolved. The first group consists of *A.nilotica* subsp. *kraussiana*, *A.nilotica* subsp. *tomentosa*, *A.nilotica* subsp. *subalata* and *A.leuderitzii* var. *leuderitzii*. These are all the subspecies of *A.nilotica* studied, and it is not unexpected that they should group together. The presence of *A.leuderitzii* at the base of this clade cannot be satisfactorily explained at present. A close relationship to *A.nilotica* has not been suggested before, and further investigation is required to clarify the relationships within this clade. The second group consists of *A.seyal* var. *seyal*, *A.seyal* var. *fistula* and *A.xanthophloea*. The presence of *A.xanthophloea* in this group is unexpected, as a close relationship between *A.seyal* and *A.xanthophloea* has never been suggested. These two species must, however, be fairly closely related as a hybrid, *A.seyal* var. *fistula* x *A.xanthophloea* has been reported (see Ross, 1979). The third group consists of *A.sieberana* var. *seiberana* and *A.sieberana* var. *woodii*.

Maslin and Stirton (in press) also list two species groups of subgenus *Acacia* in Asia. No accessions of *Acacia* from Asia were analysed in this study.

Summary.

The cpDNA data and the resulting cladogram have been only partly effective in resolving species relationships in subgenus *Acacia*. Within the New World species the cpDNA data were very effective, completely resolving relationships of the species studied. Many of the relationships reinforced previous groupings of species based on morphological comparisons. In species from Africa the cpDNA data did

not resolve any species relationships. At best it differentiated species from each other. The reason for the lack of resolution appears to be the lack of variation in the chloroplast DNA of the African species of subgenus *Acacia*. This lack of variation may also be confounding relationships of the taxa in Africa. There are many methodologies that are available that could possibly overcome this problem. Probably the most cost effective methods would be to use Isozymes to analyse species relationships (Hillis and Moritz, 1990).

4.6.2 Subgenus Aculeiferum.

Figure 4.16 shows the clade relating to subgenus *Aculeiferum* , taken from the strict consensus tree (figure 4.2). All but one of the accessions of subgenus *Aculeiferum* analysed are present in this clade. The accession which is not present on this clade is *Acacia mammifera,* which was linked within the Ingeae (see figure 4.2) next to *Albizia tomentosa*. The reasons for the unusual position of *Acacia mammifera* are at present unknown. *A.mammifera* is in a group of taxa which in general morphology could be in subgenus *Acacia*. However, their pollen morphology indicates that they are in subgenus *Aculeiferum* (L.Rico, pers. com.). In addition the ontogeny of their seedlings fits into neither subgenus (L.Rico, pers. com.). Further investigations will be needed to clarify the precise relationships of this taxon.

The accessions studied within this clade are fairly well differentiated, with relationships between taxa apparent. Unlike subgenus *Acacia*, the groups of species do not appear to belong to geographical groupings.

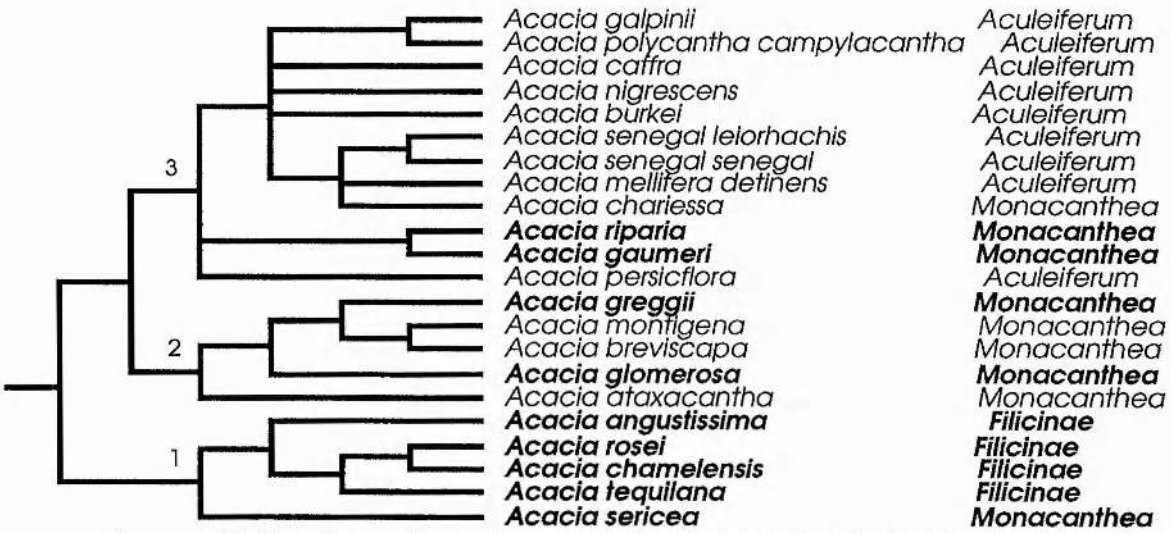


Figure 4.16. This figure shows the clade containing members of subgenus Aculeiferum from the strict consensus tree, figure 4.2. The labels *Aculeiferum*, *Monacanthea* or *Filicinae* refer to the sections of Guinet and Vassal (1972) which that taxa belongs to. The species highlighted in bold print are those which are present in the New World. The numbers refers to the delimitation of groups used in the discussion. Clade 1 is considered to be the section *Filicinae* clade, clade 2 the section *Monacanthea* clade and clade 3 the section *Aculeiferum* clade, see text for further details.

Previous attempts at subdividing subgenus *Aculeiferum*.

Subgenus *Aculeiferum* was established by Vassal (1972). Vassal based his subgenus *Aculeiferum* on the series *Vulgares* and series *Filicinae* of Bentham (1875) who subdivided series *Vulgares* into four subseries. These subseries were based on the geographical position of the taxa and their inflorescence shape, i.e. Old World capitate and spicate flowered species and New World capitate and spicate flowered species. The convenience of these subdivisions of Bentham (1875) does not represent a natural separation of the Old World and New World species (Ross, 1973).

Guinet and Vassal (1972) suggested new subdivisions for subgenus *Aculeiferum*. They subdivided it into three sections, section *Monacanthea*, section *Aculeiferum* and section *Filicinum*. Sect. *Filicinum* is directly referable to Bentham's series *Filicinae*. Bentham's name for this subdivision has been conserved. Section *Filicinae* is distinct from the other two sections by its pollen characters (Guinet, 1990) and by a lack of

"specialised" characters (Guinet and Vassal, 1972). The other two sections were separated from each other on prickle characters. Species in section *Aculeiferum* are armed with prickles in pairs, threes, or solitary prickles near the nodes. Species from this section are confined to Africa and Asia, and also have spicate inflorescences. Ross (1979) viewed it as desirable to further subdivide section *Aculeiferum* into two groups, those with prickles in pairs and those with prickles in threes or solitary.

In section *Monacanthea* the prickles are scattered along the stem. Rarely the specimen is unarmed. Both types of inflorescence are present in taxa from this section, which is present in the New World, Africa and Asia. However, no species is found in more than one area. Ross (1979) suggests that the African species of section *Monacanthea* can further be subdivided into taxa with spicate inflorescences and those with globose inflorescences. In the New World Maslin and Stirton (in press) recognise four major species groups as well as a small number of taxa with uncertain affinities. These groups appear to be founded on the work of D. Siegler (quoted in Maslin and Stirton, in press) and are based on morphological characteristics.

Pedley (1986) amalgamated sections *Aculeiferum* and *Monacanthea* into his section *Senegalia* of the genus *Senegalia*. Section *Filicinae* was retained.

The cladogram shown in figure 4.16 does not appear to differentiate fully the sections of subgenus *Aculeiferum*. There are three clades and these correspond approximately to the sections of Vassal (1972). However, the cpDNA does not support the monophyly of each section.

If we first consider section *Filicinae* we can see that all the taxa in this section are grouping together. Basal to these taxa is *A.sericea*, a

species from the New World in section *Monacanthea*. Section *Filicinae* has been suggested to lie in many positions on the phylogeny of *Acacia sens. lat.* This section was viewed as basal within the genus *Acacia sens. lat.* by Guinet and Vassal (1978). Pedley (1987) suggested that it could be treated as a distinct genus due to its chemical and morphological characteristics. Guinet (1990), analysing the pollen characters of the *Filicinae* suggested that it was closely related to the *Piptadenia* group in the Mimoseae. The cpDNA data presented in this study suggest that section *Filicinae* is within subgenus *Aculeiferum* of *Acacia sens. lat.*, and not outwith *Acacia sens. lat.* The precise relationships of section *Filicinae* cannot at present be fully ascertained. The position of *A.sericea* confounds this issue, since if this taxon were not present then it would appear that section *Filicinae* is basal within subgenus *Aculeiferum*. This observation would correspond well with morphological and chemical data. However, *A.sericea* is present in a basal position to section *Filicinae*, suggesting that species described as belonging to section *Filicinae* have arisen from within subgenus *Aculeiferum*. This needs further investigation.

The next clade to be considered is that which contains the majority of accessions studied from section *Monacanthea*. Not all of the accessions of section *Monacanthea* studied are within this clade. As mentioned earlier *A.sericea* is basal to the Section *Filicinae* grouping, and *A.riparia*, *A.gaumeri* and *A.chariessa* are in the third clade, that relating to section *Aculeiferum*. The reason for the presence of these three taxa in this clade is unknown. It is possible that *A.chariessa* may be grouping with *A.senegal* and *A.mellifera* due to hybridisation, cpDNA capture or simply a misidentification of this accession. This explanation, however, cannot explain the presence of *A.riparia* and *A.gaumeri* in the clade relating to section *Aculeiferum*. Both these species, which Maslin and Stirton (in press) consider closely related, are from the New World and no species from

section *Aculeiferum* is present in the New World. On the majority rule tree of trees that are 20-steps-or-less longer than the minimal length tree, these two taxa are basal to the clade relating to section *Aculeiferum* (see figure 4.8). Again, the relationships of these three taxa require further investigation.

Within the section *Monacanthea* clade the relationships of the taxa do not appear to be geographically based. This can be seen in figure 4.8. *A.montigena* and *A.breviscapa* are both from Africa and have capitate inflorescences. This gives some support to Ross's (1979) suggestion that the African species of section *Monacanthea* could be divided according to the inflorescence shape. However, the two other African species of section *Monacanthea* analysed, *A.ataxacantha* and *A.chariessa*, both with spicate inflorescences, are apparently not closely related to each other or to *A.montigena* and *A.breviscapa*. *A.chariessa* as mentioned earlier is within the clade which appears to relate to section *Aculeiferum*. *A.ataxacantha* is basal to the other accessions in the *Monacanthea* clade. Maslin and Stirton (in press) mention that *A.ataxacantha* is anomalous within section *Monacanthea*. This study suggests that it is correctly placed within section *Monacanthea*, though any further conclusions regarding its affinities are not possible.

The other taxa in this clade are *A.greggii* and *A.glomerosa*. Both are New World accessions from section *Monacanthea*. Maslin and Stirton divided the New World taxa of section *Monacanthea* into four groups. Although *A.greggii* and *A.glomerosa* are in different groups, Maslin and Stirton (in press) view the two groups they belong to as closely related. The cpDNA data do not support this viewpoint. Although *A.greggii* and *A.glomerosa* are close on the cladogram (see figure 4.16), *A.greggii* appears to be closer to the African taxa *A.montigena* and *A.breviscapa* than it is to *A.glomerosa*.

The final clade to be considered relates to section *Aculeiferum*. This section appears to be monophyletic. All the accessions of section *Aculeiferum* studied here appear in this clade. This clade is not as well differentiated as the other clades in subgenus *Aculeiferum*. This is probably due in part to the lack of variation encountered within this section, similar to the poor resolution caused by lack of variation in subgenus *Acacia*. The presence of *A.chariessa*, *A.riparia* and *A.gaumeri* in this clade have been discussed above. *A.persiciflora* appears basal to this clade, though whether this reflects a true division is uncertain at present. Ross (1979) suggested that section *Aculeiferum* could be divided according to the number of prickles present at the nodes. This study provides no evidence to suggest that this divide in armature is reflected in the cpDNA of accessions of section *Aculeiferum*, though not enough accessions with three prickles or a solitary prickle have been analysed to totally refute Ross (1979). *A.senegal* var. *leiorachis* and *A.senegal* var. *senegal* were the only species with three prickles studied. These taxa come out on a clade with *A.mellifera* and *A.chariessa* also present, within section *Aculeiferum* as a whole.

A.galpinii and *A.polycantha* appear to be closely related. Although this relationship has not been suggested before, there is no evidence which suggests that this finding is incorrect. The species are similar morphologically. Ross (1979) suggested that *A.galpinii* and *A.persiciflora* were closely related, as were *A polycantha* and *A.caffra* . The cpDNA data do not support these pairings.

Summary.

The relationships of taxa within subgenus *Aculeiferum* as revealed by this study of cpDNA restriction site variation, are equivocal. Section *Filicinae* appears to be basal within subgenus *Aculeiferum*, though this

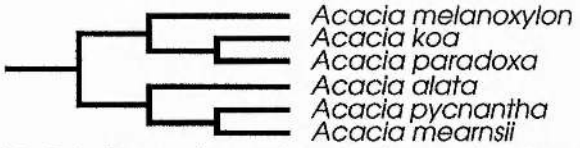


Figure 4.17. This figure shows the clade relating to subgenus Phyllodineae. It has been taken from the strict consensus tree, figure 4.2.

requires further investigation. The two other sections of subgenus *Aculeiferum*, section *Aculeiferum* and section *Monacanthea*, appear to only be partly resolved. Taxa from section *Monacanthea* appear in all parts of the cladogram relating to subgenus *Aculeiferum* (figure 4.16).

No firm taxonomic conclusions can be drawn from the cladogram. It is apparent from the relationships suggested that further investigation of subgenus *Aculeiferum* is required to ascertain the validity and relationships of the sections suggested by Guinet and Vassal (1972).

4.6.3 Subgenus *Phyllodineae.*

Figure 4.17 shows the clade relating to subgenus *Phyllodineae*, taken from the strict consensus tree. All the accessions of subgenus *Phyllodineae* analysed are present, and it appears that this subgenus is monophyletic. The accessions are well resolved, relationships between the taxa being clear.

Previous attempts at subdividing subgenus *Phyllodineae.*

It is not unexpected that in a subgenus with over 900 described species that there should be problems in devising a system of meaningful categories (Maslin and Stirton, in press). The subgenus *Phyllodineae* of Vassal (1972) was an amalgam of Bentham's (1875) series, *Botrycephalae*, *Phyllodineae* and *Pulchellae*. Bentham (1875) further subdivided series

Phyllodineae into eight subseries. All his divisions were based on floral, vegetative and fruit characteristics. Vassal (1972) assigned the name subgenus *Phyllodineae* to these series, and divided his subgenus into four sections with several subsections. The divisions of subgenus *Phyllodineae* by Vassal were confirmed by Pettigrew and Watson (1975) who analysed the morphological characteristics of 171 species from subgenus *Phyllodineae*.

Currently the most generally used classification of subgenus *Phyllodineae* is that of Pedley (1978) who recognised seven sections, which correspond well with Bentham's series. These sections are discussed in detail in Maslin and Stirton (in press). However, a serological study by Brain and Maslin suggests that much of the existing classification of this group is in doubt (Bruce Maslin, pers. com.).

The six species analysed in this present study of cpDNA restriction site variation cannot be considered representative of the other 900 species in subgenus *Phyllodineae*. Therefore, any relationships indicated by these accessions are tentative, and are included here to complete the discussion. From the cladogram (figure 4.17) we can split the six taxa into two groups. The first group is that of *A.paradoxa*, *A.melanoxylon* and *A.koa*. The second group consists of *A.mearnsii*, *A.alata* and *A.pycnantha*.

In the first group *A.melanoxylon* and *A.koa* are both from section *Plurinerves*, but these two are not grouped together as one might expect, the closest taxa in this clade appear to be *A.koa* and *A.paradoxa*, with *A.melanoxylon* basal to them. *A.paradoxa* belongs to section *Phyllodineae*, a very close relationship between section *Phyllodineae* and section *Plurinerves* is not suggested in any previous classifications. These results may be a reflection of the inadequacy of the sample size.

In the second grouping of species, *A.pycnantha* from section *Phyllodineae* would be expected to group with *A.paradoxa*. However *A.paradoxa* is in the first group of species, and *A.pycnantha* appears to be closely related to *A.mearnsii* from section *Botrycephalae*. Recent studies have suggested that section *Botrycephalae* and section *Phyllodineae* are closely related (Maslin and Stirton, in press). Basal to these two species is *A.alata* from section *Alatae*. This section is an un-natural assemblage of species (Maslin and Stirton, in press) and therefore no conclusions can be drawn without much more work.

Summary

No firm conclusions on the relationships of the taxa of subgenus *Phyllodineae* could be reached, due to small number of taxa analysed. The relationships suggested by this study are in most cases contrary to established classifications. Many more taxa drawn from all the sections and subsections of subgenus *Phyllodineae* need to be analysed to clarify relationships in this undoubtedly complex subgenus.

4.7 Taxonomic implications of the cpDNA data.

The results of this present study of cpDNA restriction site variation suggest that certain nomenclatural changes in the genus *Acacia sens. lat.* should be considered, though this thesis is not the correct medium for making such changes.

The apparent polyphyly of *Acacia sens. lat.* revealed by this present investigation is the basis for discussing possible nomenclatural implications. If *Acacia sens. lat.* had been found to be monophyletic then any nomenclatural changes would be harder to justify. However, the cpDNA data suggests that a reassessment of the classification might be

appropriate, though the dangers of using only one type of character as the basis for nomenclatural change must be recognised.

At present the tribe Acacieae consists of the genus *Acacia* and the monotypic genus *Faidherbia*. The genus *Acacia* has three subgenera, subgenus *Acacia*, subgenus *Aculeiferum* and subgenus *Phyllodineae*. The cpDNA data do not support this classification. The cpDNA data suggest that subgenus *Acacia* and subgenus *Aculeiferum* are closely related, but subgenus *Phyllodineae* does not appear to be closely related to them. Instead it appears to be related to taxa within the tribe Ingeae. *Faidherbia* appears to be basal to the Ingeae (See figure 4.9).

Changes to the classification which would be in harmony with the cpDNA data are i) the transfer of subgenus *Phyllodineae* to the Ingeae as a genus, along with the genus *Faidherbia* and ii) the consequent reduction of the tribe *Acacieae* to consist of the genus *Acacia* with two subgenera, subgenus *Acacia* and subgenus *Aculeiferum*.

Subgenus *Phyllodineae*, if transferred to the Ingeae as suggested, would need a new generic name. The generic name *Racosperma* (D.C.) MARTIUS was suggested by Pedley (1986), but Maslin (1988) doubted the validity of this name for subgenus *Phyllodineae* since he thought that the name *Racosperma* might not be linked to a validly published description. Further research and discussion would be needed to produce an acceptable new generic name for subgenus *Phyllodineae*, if transference to the Ingeae was considered appropriate.

Faidherbia has previously been thought to be one of the genera of the tribe Acacieae. The cpDNA data suggest that this is not the case. It appears that *Faidherbia* is closely related to the Ingeae. Further evidence will be needed to ascertain its position relative to other taxa within the Ingeae. There does not appear to be any evidence to suggest that the transference of *Faidherbia* to the Ingeae would be inappropriate. This

transferral has been suggested before, e.g. Guinet (1990). The generic name *Faidherbia* would be retained.

The tribe Acacieae would be reduced to contain only the genus *Acacia* consisting of two subgenera, subgenus *Acacia* and subgenus *Aculeiferum*. The cpDNA data do appear to confirm a real divide between subgenus *Acacia* and subgenus *Aculeiferum*, though whether this divide is strong enough to warrant recognition of two genera is equivocal. Until there is a compelling need to recognise two genera within *Acacia*, it is better to keep the two subgenera *Acacia* and *Aculeiferum*.

Maslin and Chappill (1995) suggested that the generic name *Acacia* should be restricted and be applied to what is at present subgenus *Phyllodineae*. This would mean re-typifying the name *Acacia* on a taxon in subgenus *Phyllodineae*. Their reasons for doing this are discussed at the end of Chapter 2. The cpDNA data suggest that this would be imprudent. Rather, the opposite appears to be the natural course of action, the generic name *Acacia* being kept for subgenus *Acacia* and subgenus *Aculeiferum*, and a new generic name being found for subgenus *Phyllodineae*. The reasons for this are twofold. Britton and Rose (1928) lectotypified the name *Acacia* to *A.nilotica*, a member of subgenus *Acacia*. If the ICBN rules are to be consistently implemented then the name *Acacia* must stay with *A.nilotica* and subgenus *Acacia* . In addition, the present work suggests that subgenus *Acacia* and subgenus *Aculeiferum* appear to be a distinct pair of taxa between the Mimoseae and Ingeae, corresponding to the tribe Acacieae. Subgenus *Phyllodineae* on the other hand appears to be within the Ingeae, rather than the Acacieae and so cannot be considered for the name *Acacia*.

4.8 Biogeographic implications of the cpDNA data.

According to Ross (1981) the identity of the ancestral form of *Acacia* has "been the subject of much speculation and disagreement". Andrews (1914) and Atchison (1948) considered subgenus *Acacia* to be the ancestral form of the genus, because it contained what they considered to be the primitive morphological characters of the genus. This viewpoint was supported by Tindale and Roux (1975) whose work on the chemical content of South African species led them to suggest that subgenus *Aculeiferum* is "generally more advanced" than subgenus *Acacia*.

The first major work to consider the origin and distribution of the genus *Acacia* was that of Guinet and Vassal (1978). This paper has been discussed in Chapter 2. To summarise their conclusions; they believed that the genus *Acacia* originated in western Gondwana, and then spread pantropically. Guinet and Vassal (1978) also noted that subgenus *Aculeiferum* had a preponderance of characters which they considered primitive. Their conclusions were based on the study of morphological characters in the genus. These conclusions supported the earlier work of Robbertse (1974) who investigated the inflorescence and flowering system of South African Acacias. Robbertse (1974) considered that subgenus *Acacia* was more advanced than subgenus *Aculeiferum*. He based his conclusions on his belief that subgenus *Aculeiferum* displayed many primitive floral characteristics.

Ross (1981) developed these ideas and postulated his own. The fact that the rainforested areas of the world were much more extensive during the Palaeocene, led Ross (1981) to suggest that *Acacia* originated in lowland forests. At present *Acacia* is not represented in tropical forests to any degree. It seems that the genus is intolerant of low light levels. Ross (1981) therefore suggested that the ancestral members were climbers or

lianes, as this would enable them to reach and maintain an emergent position in the canopy. Many species in subgenus *Aculeiferum* exhibit this characteristic, e.g. *A.lujae* and *A.ataxacantha* appear to be obligate climbers.

Pedley (1986), in his paper promoting the three subgenera of *Acacia* to genera, considered also the biogeographic aspects of *Acacia*. Pedley first considered the distribution of the Acacieae[2], he outlined two options for the origin of the genus *Acacia sensu* Vassal. Either the genus had originated after the fragmentation of Gondawanaland or the genus had existed at about the time of fragmentation. Pedley (1986) believed that the genus was not suited for long-range dispersal, and so the genus must have existed when Gondwanaland broke up, with both subgenus *Acacia* and subgenus *Aculeiferum* already present in tropical parts of Africa. This is in contrast to the idea put forward by Ross (1981) who believed it unlikely that *Acacia* or its immediate prototype had differentiated before or during the separation of Africa and South America. Pedley (1986) considered "unlikely" the suggestion of Guinet and Vassal (1978) that *Acacia* originated in west Gondwanaland because taxa in tropical America, especially sect. *Filicinae*, had a "preponderance" of primitive characters. Pedley regarded these characters as derived. He also considered that subgenus *Phyllodineae* developed from subgenus *Aculeiferum* in east Gondwanaland (Pedley, 1986) at about the same time as India and Australia-Antarctica separated, and subgenus *Phyllodineae* then diversified extensively in Australia.

The general consensus of opinion at present favours the hypothesis of Guinet and Vassal (1978), i.e. an origin of *Acacia* in west

---

2 Pedley (1986) promoted all three subgenera of *Acacia* to the generic level. Pedley retained the tribe Acacieae for these three genera. In this discussion I will continue to use the subgenera of Vassal (1972) in place of Pedley's genera.

Gondwanaland. There does not appear to be agreement as to when the genus originated .

Do the results of this present analysis shed any light on the biogeography of the genus? The answer to this question is 'possibly', though we must bear in mind the limitations of the data, as discussed above. These limitations mean that many of the conclusions are speculative. For firm conclusions concerning the biogeography of the genus more conclusive data concerning the relationships of the subgenera will be needed.

In all the previous discussions on the biogeography of the genus, it has been assumed that subgenus *Phyllodineae* is very closely related to subgenus *Aculeiferum*, and has probably evolved from this subgenus. This present investigation indicates that this may not be the case. It seems probable that subgenus *Phyllodineae* is not within the Acacieae, but within the Ingeae. The question now becomes, to which genera in the Ingeae is subgenus *Phyllodineae* closely related? This question cannot be answered by this investigation. Too few taxa from the Ingeae have been analysed to enable suggestions for the affinities of subgenus *Phyllodineae*, but this study indicates that any further investigation into the affinities of subgenus *Phyllodineae* should take into account its possible position within the Ingeae.

Subgenus *Acacia* and subgenus *Aculeiferum* are both distributed across the New World, Africa and south-east Asia. The infraspecific relationships of these taxa have already been discussed. It is possible from the cladogram (figure 4.2) to suggest how differentiated each subgenus was when Gondwanaland fragmented. The cladogram does not, however, suggest where the genus *Acacia* originated, or which subgenus evolved first.

The first step is to compare the relationships of taxa within each subgenus in respect of their geographical distribution. In subgenus *Acacia*, the New World taxa and the African taxa are well differentiated. The New World taxa form a distinctive clade on the cladogram relating to subgenus *Acacia*, separate from the African accessions (see figure 4.14). In subgenus *Aculeiferum* the relationships of taxa do not appear to follow geographic boundaries. This can be seen in figure 4.16 where taxa from each continent are mixed together, although very closely related species come from similar areas, e.g. *A.montigena* and *A.breviscapa*. The only clade that comes from a single geographical area is that relating to section *Filicinae*, species of which are restricted to the New World.

The observation that some taxa in subgenus *Aculeiferum* which appear to be closely related occur in disjunct geographical areas, suggests that this subgenus may have been relatively well differentiated at the specific level when Gondwanaland fragmented. Before Gondwanaland fragmented these sibling taxa, or their progenitors, would perhaps have had adjacent or similar distributions. When Gondwanaland fragmented, these taxa were separated, the sibling taxa being separated as the fragments of Gondwanaland parted. This would account for the observation that taxa which appear to be closely related are separated by the Atlantic ocean. This situation is most apparent in section *Monacanthea*, the only section recognised by Vassal (1972) which has a pantropical distribution. Taxa in this section do not appear in any way to be geographically arranged. This situation corresponds, though not for the same reasons, with the hypothesis forwarded by Pedley (1986) who believed that the genus *Acacia* existed when Gondwanaland fragmented, due to the inability of long range dispersal of seed in the genus.

The extent of differentiation of taxa at the specific level appears, at the time of fragmentation of Gondwanaland, to be less in subgenus *Acacia*. This subgenus probably had the majority of its specific differentiation after Gondwanaland fragmented. This would account for the divide between New World and Old World taxa that appears to exist. If this subgenus had been as fully differentiated as subgenus *Aculeiferum*, at the time of fragmentation, then we might expect the interspecific relationships to mirror those in subgenus *Aculeiferum*. It is possible that the fragmentation and climatic changes associated with the continental movement, such as the increasing aridity of Africa, provided the stimulus for the differentiation of subgenus *Acacia*.

Even though subgenus *Aculeiferum* appears to be more differentiated than subgenus *Acacia* at the time of fragmentation, it is impossible to suggest, based on the cpDNA data, which of these subgenera is ancestral. With further investigations into the interspecific relationships of each subgenus, combined with a greater understanding of the morphological data the picture of evolution in the genus *Acacia* will hopefully become clearer.

4.9 Summary of Chapter 4.

The aim of this chapter was to produce a phylogeny of the genus *Acacia* based on cpDNA restriction site characters. This was successfully accomplished with a robust phylogenetic tree being produced (see figure 4.2). From this cpDNA phylogeny relationships between the subgenera of *Acacia* as well as other related taxa were suggested. The cpDNA data suggested that the two subgenera of *Acacia* with similar distributions, subgenus *Acacia* and subgenus *Aculeiferum*, were closely related. The third subgenus of *Acacia*, subgenus *Phyllodineae*, which is mostly

restricted to Australia, was found not to be closely related to either of the two other subgenera of *Acacia*. Subgenus *Phyllodineae* appeared to be closely related to taxa within the Ingeae, the sister tribe of Acacieae. The monotypic genus *Faidherbia*, thought to be within the tribe Acacieae appeared to be basal to the Ingeae.

The classification of *Acacia* as suggested by the cpDNA data challenges previous classifications of the genus. The major difference is in the degree of relatedness between subgenus *Aculeiferum* and subgenus *Phyllodineae*. Previous classifications, based mainly on morphological characters, have always suggested a close relationship between these two taxa. This relationship is supported by a suite of morphological characters. The cpDNA analysis presented here suggests that subgenus *Aculeiferum* and subgenus *Phyllodineae* are not closely related. The dichotomy between the morphology based and cpDNA based phylogenies could not be resolved and must await further investigation. It is possible that factors such as lineage sorting or hybridisation at an early stage between the diverging lines of *Acacia* are confounding the phylogeny.

The phylogenetic tree was resolved enough to enable the interspecific relationships of taxa in subgenus *Acacia* and subgenus *Aculeiferum* to be clarified. Within subgenus *Phyllodineae* only six species from around 900 species were analysed. This lack of taxa meant that interspecific relationships in this subgenus could not be commented on. Within subgenus *Acacia* the taxa analysed from the New World formed a distinctive clade (see figure 4.14). Taxa from Africa could not be resolved to the same level, this was probably due to the lack of cpDNA variation between African taxa. The accessions from the New World were fully resolved and species groupings as suggested by morphological analyses (see Maslin and Stirton, in press) correlated well with relationships as

suggested by the cpDNA data. Within subgenus *Aculeiferum* the infraspecific relationships suggested by the cpDNA data were less clear. Section *Filicinae* of Vassal (1972) was well supported as a monophyletic group, though it appears to be derived from within another of Vassal's (1972) sections of subgenus *Aculeiferum*. This needs further investigation. The other two sections of Vassal were only minimally supported by the cpDNA data. There was a degree of interrelationship between the two sections, though the majority of taxa analysed from these two sections formed distinctive clades (see figure 4.16). Unlike subgenus *Acacia* where the accessions appeared to be divided according to their geographical origin, in subgenus *Aculeiferum* the accessions appeared not to be grouped together according to their origin.

The lack of geographic partitioning within subgenus *Aculeiferum* suggests that this subgenus was well differentiated at the specific level when Gondwanaland fragmented during the mid-Cretaceous period, about 10^8 years b.p. (Raven and Axelrod, 1974). In contrast subgenus *Acacia* appeared to be less well differentiated at the specific level at this time. This is illustrated by the divide between African and New World accessions studied.

The implications of the cpDNA data for a classification of the genus *Acacia* are discussed, but until there is either a congruence between the cpDNA data and morphological data, or an explanation as to their dichotomy, no nomenclatural changes are advocated.

To conclude, the aims of this chapter were accomplished, i.e. to produce a cpDNA phylogeny of *Acacia* and compare it to other classifications.

Chapter 5
Acacia laeta, an interspecific hybrid ?

5.1 Introduction to *A. laeta*.

Although some natural interspecific hybrids are known to occur, and hybridisation is suspected in several instances, the full extent of hybridisation in the African Acacias is unknown. Where hybridisation has been reported it is only between closely related species. No 'cross-subgeneric' hybrids have been found.

Acacia laeta R.BR. *ex* BENTH. has long been thought to be of hybrid origin. *A. laeta* is a shrub or tree up to 6 metres high, found in tropical west Africa and north-east Africa from Egypt southwards to Tanzania. Aubreville (1950) observed close affinities between *A. laeta* and *A. senegal* (L.) WILLD. in habitat, habit, prickles, inflorescence and pod characters and suggested that *A. laeta* is a hybrid between *A. senegal* and another species. Jackson & Peake (1955) suggested that the other species was *A. mellifera* (VAHL.) BENTH., i.e. *A. laeta* was a hybrid between *A. senegal* and *A. mellifera*. However, Brenan (1959) suggested that *A. laeta* was a hybrid of *A. mellifera* and another *Acacia* sp., most probably *A. goetzei* HARMS. All three putative parents, *A. mellifera*, *A. senegal* and *A. goetzei* are sympatric. Plate 5.1 (overleaf) shows *A. laeta* and two of its putative parents, *A. senegal* and *A. mellifera*.

The first investigation of *A. laeta* was by El Amin (1976). He only considered *A. senegal* and *A. mellifera* as parents. He gathered 33 morphological characters from specimens throughout Sudan. After comparing these characters in the three taxa he noted that:

10 characters are similar in all three species;

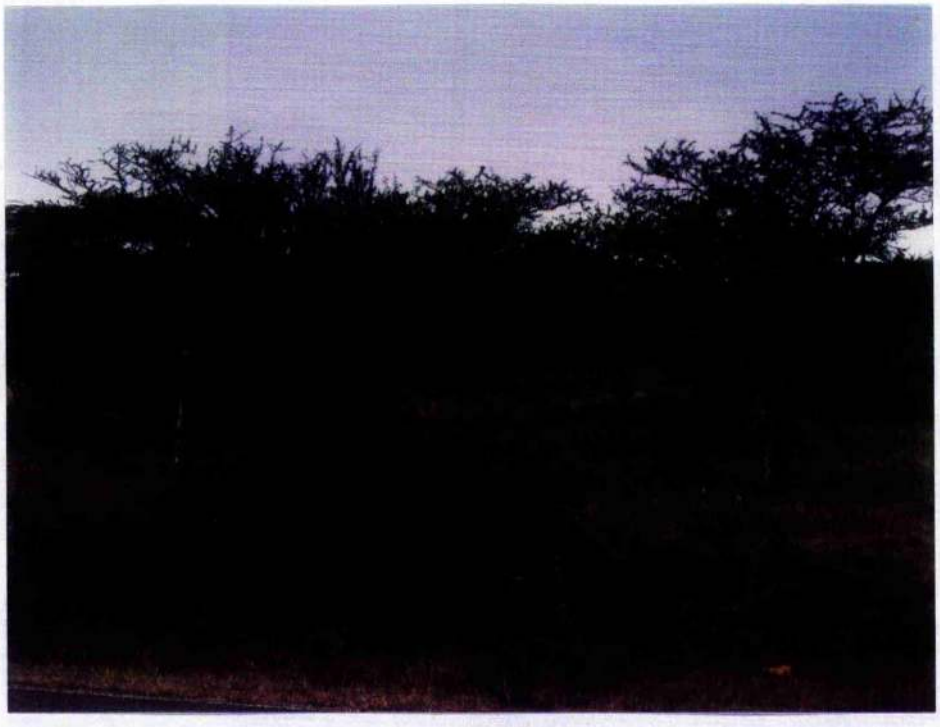


Plate 5.1. This photograph shows the two putative parents of *A. laeta* and *A. laeta* itself. In the foreground as a shrubby bush is *A. mellifera*, behind this to the left is *A. laeta*, and to the right of *A. laeta* is *A. senegal*. This picture was taken near Kajiado, Kenya. This natural population was the second population of *A. laeta*, *A. senegal* and *A. mellifera* sampled (see materials).

11 characters in *A. laeta* are intermediate between *A. senegal* and *A. mellifera*;

9 characters are shared by *A. mellifera* and *A. laeta*;

3 characters are shared by *A. senegal* and *A. laeta*.

El Amin noted that the *A. laeta* he studied had a triploid number of chromosomes, 2n=39. Both *A. senegal* and *A. mellifera* have diploid numbers, 2n=26. He viewed this as a good indication that *A. laeta* was of hybrid origin. However, 2n=39 is not the only chromosome number recorded for *A. laeta*. 2n=26 and 2n=52 (Darlington and Wylie, 1955), (diploid and tetraploid respectively) have been recorded.

El Amin concluded that it was quite possible that *A. laeta* was a hybrid of *A. senegal* and *A. mellifera* and that the closer affinity with *A. mellifera* was probably due to backcrossing with it in the past. He suggested that instability in morphological characters, such as number and arrangement of leaf parts, is not unexpected in a taxon of hybrid origin.

M-H Chevalier (1994) in a study investigating similarities and divergence among 15 *Acacia* species/subspecies, using isozymes as genetic markers concluded that "*A. laeta* has isozymic contributions from both *A. senegal* and *A. mellifera*". He interpreted this result as confirming *A. laeta* as a hybrid between these two species.

5.2 How can molecular tools or methods help us investigate this problem?

Avise (1994) states that "molecular markers provide powerful means for identifying hybrid organisms". Rieseberg and Brunsfeld (1992) add "molecular markers currently provide the best means for analysing ambiguous cases of introgression (or hybridisation)". What features of molecular characters have led the above authors to make such statements? Unlike morphological characters which tend to converge when exposed to similar selective pressures, molecular markers tend to be neutral (Kimura, 1982). Specific

molecular markers can be polarised by comparison to related taxa, allowing one to distinguish between different hypotheses (Rieseberg and Brunsfeld, 1992). Theoretically, an almost unlimited number of independent molecular markers that differentiate the taxa under investigation can be obtained. Finally molecular characters can be found in both the nuclear and cytoplasmic genome, allowing gene flow in both to be monitored.

The chloroplast genome is useful for the study of hybridisation because it exists in multiple nonrecombinant copies. It is also not diluted throughout repeated backcrossing, if inherited maternally or paternally, and can be valuable for tracing the presence of genetic material in backcrossed progeny which might not be apparent in introgressed individuals (Kron *et al.* 1993).

Nuclear markers are of more direct importance to the study of introgression, because nuclear markers are considered to be co-dominantly inherited, any hybrid accession ought to have the additive profile of the two putative parents.

Now let us look at two examples where molecular characters have been used to re-examine or confirm cases where hybridisation or introgression is thought to have occurred.

Rieseberg *et al.* (1990) examined a proposed case of introgression in the genus *Helianthus*. Heiser (1949) suggested that a weedy race of *Helianthus bolanderi* had originated by the introgression of genes from *H.annuus* into a serpentine race of *H.bolanderi*. To determine whether the weedy race actually originated via hybridisation/introgression they analysed allozyme, chloroplast DNA and nuclear-ribosomal DNA variation.

None of the collected evidence supported the proposed introgressive origin of weedy *H.bolanderi*. Given the lack of parallel or convergent mutations in the cpDNA and rDNA phylogenetic trees, and the congruence of these trees with the isozyme (and flavonoid) patterns, and the presence of a unique and divergent chloroplast genome in the weedy race of *H.bolanderi*,

they suggested that the weedy race of *H.bolanderi* was not derived through introgression, as hypothesised by Heiser, but was relatively ancient in origin.

Another example is that of Spooner *et al* .(1991), who re-examined the putative hybrid *Solanum raphanifolium*. Ugent (1970) suggested that S.*raphanifolium* was a recent stabilised diploid hybrid between *S.canasense* and *S.megistacrolobum*. This was supported by: the intermediate morphology of the putative hybrid; its occurrence in weedy disturbed habitats (both parents favour undisturbed locations); and it occurrence in a restricted zone of overlap between the two putative parental species in southern Peru. The study carried out by Spooner *et al.* (1991) tested the hypothesis of hybrid origin for *S.raphanifolium* with cpDNA and rDNA characters. If the hybridisation was relatively recent or ongoing then the expectation would be for the hybrid to possess an identical or nearly identical, maternally inherited, cpDNA pattern from one parent (Hosaka *et al.*, 1984) and an additive biparentally inherited rDNA pattern from both parents. If *S.raphanifolium* were not derived from a recent diploid hybridisation, the chloroplast analysis would place it with either of the other two species (but exhibiting a number of autapomorphies), or perhaps even as a sister group to the other two species; and the rDNA pattern would not be expected to exhibit additivity.

This is in fact what they found. Phylogenetic analysis of both the cpDNA and rDNA indicate that *S.raphanifolium* is not of recent hybrid origin from the two putative parental species. The parental species form a monophyletic clade with *S.raphanifolium* as a sister group. Spooner *et al.*(1991) view it as more likely that *S.raphanifolium* is related to other *Solanum* species that they did not examine.

In this preliminary study both chloroplast and nuclear markers were used to investigate whether *A. senegal* and *A. mellifera* are the parents of *A. laeta*, as suggested by El Amin (1976).

Population	Tree No.	Species	Latitude °S	Longitude °E
1	1	*A. senegal*	2.03.716	36.47.889
	2	*A. mellifera*	2.03.744	36.47.921
	3	*A. laeta*	2.03.704	36.47.922
	4	*A. mellifera*	2.03.721	36.47.974
	5	*A. laeta*	2.03.776	36.47.955
	6	*A. laeta*	2.03.751	36.47.958
2	1	*A. laeta*	2.02.550	36.47.906
	2	*A. mellifera*	2.02.550	36.47.900
	3	*A. senegal*	2.02.530	36.47.936
	4	*A. laeta*	2.02.533	36.47.941
	5	*A. laeta*	2.02.548	36.47.922
	6	*A. senegal*	2.02.474	36.47.959
3	1	*A. laeta*	1.59.384	36.47.716
	2	*A. senegal*	1.59.384	36.47.716
	3	*A. mellifera*	1.59.372	36.47.761
	4	*A. laeta*	1.59.446	36.47.650
	5	*A. laeta*	1.59.358	36.47.730
	6	*A. mellifera*	1.59.358	36.47.730

Table 5.2, showing the origin of the collected taxa.

5.3 Experimental details.

5.3.1 Materials.

The plant material for this study came from leaf material collected from three populations in Kenya. The populations were situated on the A104 between Kajiado and Ilbisil. Each population was separated from the others by about 3 km. From each population 6 trees were sampled - each population sample contained *A. senegal*, *A. mellifera* and *A. laeta*. Leaf material was harvested according to the method of Chase and Hills (1991). Table 5.2 shows the origin of the plant material.

5.3.2 Methods.

The methods used to investigate the parentage of *A. laeta* were very similar to those used in the phylogeny analysis. These can be found in Appendix B. DNA was extracted from the leaves and subjected to a restriction site analysis as described in Chapter 3. The enzymes used were the same 15 as

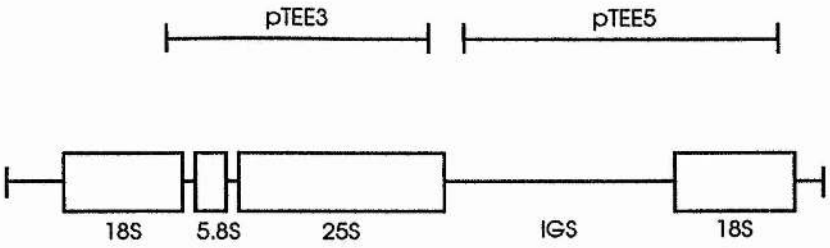


Figure 5.3. The bottom map is of a simple rDNA repeat unit. The 18S, 5.8S and 25S are coding regions within the unit, the IGS is the intergenic spacer. The lines above the map represent the position and the size of the pTEE probes used in this study. The sizes are 5.0kb for pTEE 3 and 6.5kb for pTEE 5.

used for the cpDNA phylogeny of *Acacia*. An additional two probes were used as well as the 9 mung bean probes mentioned earlier. These two probes, pTEE 3 and pTEE 5, were from *Taraxacum officinale* and encoded nuclear ribosomal DNA (rDNA) (King & Schaal, 1990) (see figure 5.3). These two clones represent about 98% of a single *Taraxacum* rDNA repeat. This enabled the study of both chloroplast and nuclear DNA.

Initially a screening of all 150 possible probe/enzyme combinations (PEC) with the two parental taxa took place. This was to identify which PEC differentiated between the two parental taxa, *A. senegal* and *A. mellifera*. The results of this survey then dictated which PEC were used to survey all 18 collected accessions.

5.4 Results.

After surveying 150 PEC, twelve probe/enzyme combinations which differentiated between the two parental taxa, *A. mellifera* and *A. senegal*, were identified. These twelve PEC consisted of seven combinations in which the probes were from cpDNA and five in which the probes were from rDNA. The PECs, mutations and taxa they were present in can be seen in table 5.4.

Having identified and characterised the 12 PEC which differentiated the parental taxa, all 18 accessions collected were analysed with these PEC. The results can be seen in Table 5.5.

			Enzyme	*EcoR* I	*EcoR* I	*EcoR* I	*EcoR* I	*Bcl* I	*Bsc* I	*Nsi* I	*EcoR* I	*Bgl* II	*HinD* III	*Stu* I	*Stu* I
			Probe	MB 11+ MB 12	MB 12	MB 8+9	MB 1	MB 7	MB 1	MB1	pTEE 5	pTEE 3 + pTEE 5	pTEE 3 + pTEE 5	pTEE 3 + pTEE 5	pTEE 3 + pTEE 5
Pop	Tree	Species													
1	1	*senegal*		s	s	s	s	s	s	s	s	s	s	s	s
	2	*mellifera*		m	m	m	m	m	m	m	m	m	m	m	m
	3	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	4	*mellifera*		m	m	m	m	m	m	m	m	m	m	m	m
	5	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	6	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
2	1	*laeta*		m	m	m	m	m	m	m	m	m	m	m	m
	2	*mellifera*		m	m	m	m	m	m	m	m	m	m	m	m
	3	*senegal*		s	s	s	s	s	s	s	s	s	s	s	s
	4	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	5	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	6	*senegal*		s	s	s	s	s	s	s	s	s	s	s	s
3	1	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	2	*senegal*		s	s	s	s	s	s	s	s	s	s	s	s
	3	*mellifera*		m	m	m	m	m	m	m	m	m	m	m	m
	4	*laeta*		m	m	m	m	m	m	m	s + m	m	m	m	m
	5	*laeta*		m	m	m	m	m	m	m	m	m	m	m	m
	6	*mellifera*		m	m	m	m	m	m	m	m	m	m	m	m

Table 5.5. This table shows the results of the full analysis. For each particular probe/enzyme combination the analysed accessions could either show *A.senegal* characters (s), or *A.mellifera* characters (m) , or both (s + m).For specific details of mutations see table 5.4.

Plate 5.2. This is a picture of the autoradiogram for the PEC *Bcl* I / MB7. In lane 1 is *A. senegal* T1P1; lane 2, *A. senegal* T3P2; lane 3, *A. senegal* T6P2; lane 4, *A. senegal* T2P3; lane 5, *A. mellifera* T2P1; lane 6, *A. mellifera* T4P1; lane 7, *A. mellifera* T2P2; lane 8, *A. mellifera* T3P3, lane 9, *A.mellifera*, T6P3, lane 10 *A. laeta* T3P1; lane 11, blank (no DNA visible); lane 12, *A. laeta* T6P1; lane 13, *A. laeta* T1P2; lane 14, *A. laeta* T4P2; lane 15, *A. laeta* T5P2; lane 16, *A. laeta* T1P3; lane 17, *A. laeta* T4P3; lane 18, *A. laeta* T5P3. In lane 10 (*A.laeta*) the smear of fragments >3.6kb are the products of a partial digestion. There is however, no 4.9kb band comparable to *A.senegal* present.

We can see from this that the chloroplast DNA phenotypes of *A. mellifera* and *A. laeta* are similar, with the exception noted above. None of the *A. laeta* accessions examined has the *A. senegal* cpDNA phenotype.

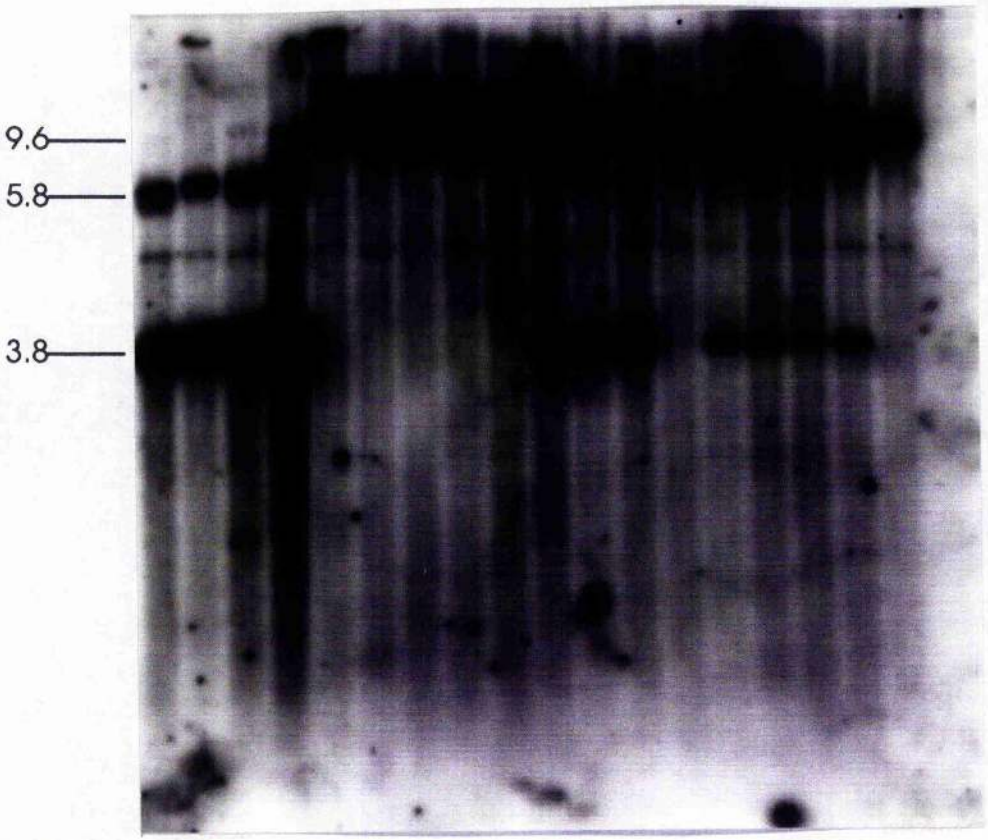


Plate 5.3. This is a picture of the autoradiogram for the PEC *Eco*R I / pTEE 5. The lane order is the same as for plate 5.2, except in lane 11 where *A. laeta* T5P1 is present. We can see in this autoradiogram the presence of *A. senegal* markers in some of the *A. laeta* accessions surveyed. The characters shared between *A. laeta* and *A. senegal* are the bands at 5.8 kb and 3.8 kb. The lanes showing both characters are 10, 11, 12, 14, 15, 16 and 17.

Probe/Enzyme Combination	Site mutation	Bands present in	
		Acacia senegal	*Acacia mellifera*
Bcl I / MB 7	2.7 + 2.2 ↔ 4.9	4.9	2.7 + 2.2
Bsc I / MB 1	2.0 + 0.3 ↔ 2.3	2.0 + (0.3)	2.3
*Eco*R I / MB 1	3.4 + 0.4 ↔ 3.8	3.4 + (0.4)	3.8
*Eco*R I / MB 8+9	2.9 + 0.9 ↔ 3.8	2.9 + 0.9	3.8
*Eco*R I / MB 11 and MB 12	2.9 + 1.2 ↔ 4.1	4.1	2.9 + 1.2
*Eco*R I / MB 12	2.2 + 0.4 ↔ 2.6	2.6	2.2 + (0.4)
Nsi I / MB 1	10.0 + 6.2 ↔ 16.2	10.0 + 6.2	16.2
Bgl II / pTEE 3 and pTEE 5	12.8 + 4.7 ↔ 17.5	17.5	12.8 + 4.7
*Eco*R I / pTEE 5	5.8 + 3.8 ↔ 9.6	5.8 + 3.8	9.6
*Hin*D III / pTEE 3 and pTEE 5	8.9 + 3.9 ↔ 12.8	12.8	8.9 + 3.9
Stu I / pTEE 3 and pTEE 5	3.9 + 7.4 ↔ 11.3	7.4 + 3.9	—[a]
Stu I / pTEE 3 and pTEE 5	5.9 + 5.4 ↔ 11.3	—[a]	5.4 + 5.9

Table 5.4. This table shows the restriction sites that were found to differ between *A. senegal* (tree 2 from population 3) and *A. mellifera* (tree 5, population 3). The first two columns detail the restriction site mutation and the probe/enzyme combination which highlights it. The final two columns show which banding pattern was encountered in the two sample accessions. For example, for the first marker *Bcl* I/MB 7, in *A. senegal* you would find a 4.9kb band, and in *A. mellifera* this band has been cut in to two smaller bands, so you would expect to find bands of 2.7 and 2.2kb. Plate 5.2 illustrates this particular probe/enzyme combination and the differences in banding patterns between *A. senegal* and *A. mellifera*.

[a] In the *Stu* I PEC the 11.3kb band did not exist in either *A. senegal* or *A. mellifera*. This band is putatively present in the ancestor of *A. senegal* and *A. mellifera*, and has undergone a different mutation in the lineages leading to either *A. senegal* or *A. mellifera*.

5.4.1 Chloroplast probe results.

As can be seen from table 5.5 all the *A. laeta* specimens surveyed had chloroplast DNA markers that were identical to those of *A. mellifera*, i.e. no *A. senegal* chloroplast characters were present in any *A. laeta* sample (see plate 5.2, overleaf).

5.4.2 Ribosomal DNA probe results.

Of the five rDNA PEC that distinguished *A. senegal* from *A. mellifera*, four of these could not distinguish *A. laeta* from *A. mellifera*.

However, seven *A. laeta* specimens surveyed with the PEC EcoR I/pTEE 5 showed a marker diagnostic of *A. senegal* (see table 5.5). These samples also displayed the *A. mellifera* marker. The other two *A. laeta* samples displayed just the *A. mellifera* marker. This can be seen in plate 5.3

displayed just the *A. mellifera* marker. This can be seen in plate 5.3. The presence of both markers is suggestive of additive inheritance.

5.5 Discussion.

5.5.1Chloroplast DNA

The presence of only *A. mellifera* chloroplast DNA characters in the *A. laeta* surveyed does not disprove that *A. laeta* is a hybrid, there are many plausible explanations available to account for absence of *A. senegal* chloroplast markers in *A. laeta*. The most likely is that only hybridisation events with *A. mellifera* as the female parent are those which take place or survive. This could occur for many reasons; for example *A. mellifera* pollen may not be able to germinate or perhaps cannot grow enough on *A. senegal* styles, but *A. senegal* pollen can grow on *A. mellifera* styles; or fertilisation takes place but no gametic fusion or endospermic fusion takes place in crosses where *A. senegal* is the female parent.

An alternative hypothesis is that the sample was not large enough. It is possible that hybridisation occurs both ways and no *A. laeta* with *A. senegal* as the female parent was collected, as the sample size of *A. laeta* surveyed was relatively small.

However, at the sites surveyed, there appeared to be more *A. mellifera* trees than *A. senegal*. If this is the case it would be expected that *A. senegal* would be the female parent due to 'pollen swamping'. Despite this, it appears that if crossing occured, it only occured in one direction - with *A. mellifera* as the female parent.

In *Acacia*, chloroplast DNA inheritance has been shown to be biparental (Tilney-Basset, 1978). Tilney-Basset (1978) interpreted Moffet's (1965) observation of hybrid variegation as indicating the existence of at least two types of plastid and the biparental inheritance of these plastids.

This means that in *A. laeta* there is the possibility of *A. senegal* type chloroplast DNA being present, even if *A. mellifera* is always the female parent, as appears to be the case. No *A. senegal* molecular characters were seen in the hybrid specimens, not even in combination with *A. mellifera* characters.

Plastid transmission is best thought of as a continuum rather than an all-or-nothing process (Harris and Ingram, 1992). Although biparental inheritance has been shown to occur, it is probable that comparatively fewer paternal chloroplast genomes are inherited. This means that *A. senegal* type chloroplasts may be present in *A. laeta* but at low concentrations, beyond the resolution of the Southern blot technique, or in chimeric tissue not collected.

5.5.2 r DNA

The ribosomal DNA offers evidence that *A. laeta* could be a hybrid between *A. senegal* and *A. mellifera*. This is shown by the presence of both *A. mellifera* and *A. senegal* markers in several of the hybrid samples. Additive inheritance of parental markers is an indicator of hybridity.

If we look again at plate 5.3, we can see that the bands from *A. mellifera* and *A. senegal* in the *A. laeta* samples are at different intensities. If these were diploid hybrids then we would expect the intensities to be similar. It is possible, though no hard evidence exists, that some of the *A. laeta* collected contain more than one *A. mellifera* genome.

However, before concluding that *A. laeta* is a hybrid two questions must first be answered:

i) Is the marker PEC *Eco*R I/pTEE 5, which has been used to distinguish *A. senegal* from *A. mellifera* unique to *A. senegal*? It is possible that another *Acacia* species could have the same rDNA marker as *A. senegal*. If so it is possible that this species could be a parent of *A. laeta*. This marker could also be present in other individuals of *A. mellifera*, but not in the samples surveyed.

This is possible as sample sizes were relatively small. More research is required to answer this question, such as a broad survey of related species and a wider survey of *A.mellifera*. If another *Acacia* species carries the same marker, then additional markers would have to be found to discriminate between the possible parents.

ii)Why did the other rDNA markers not show additive inheritance as well? The only enzyme that showed additive phenotypes was *Eco*R I. The other enzymes *Bgl* II, *Hin*D III and *Stu* I gave non-additive profiles for all of the hybrid material. The phenotype of the hybrid material surveyed with these enzymes was that of *A. mellifera*.

A. laeta T1P2 and *A. laeta* T5P3 displayed non-additive phenotypes for all of the PEC surveyed. The material for this study was collected from the field, because of this we could not designate the parentage of the *A. laeta* collected. It is therefore possible that they could have been F1 hybrids, F2 hybrids or backcrosses to either parent. These two, T1P2 & T5P3, could thus be F2 hybrids but are more likely to be backcrosses to *A. mellifera*. Through this (these) backcrossing event(s), the rDNA genes of *A. senegal* could have been lost. This could explain why non-additive phenotypes were seen with the PECs using the rDNA probes pTEE 3 and pTEE 5.

An explanation as to why the other *A. laeta* samples showed additive phenotype with the PEC *Eco*R I/pTEE 5 only is harder to explain. However results similar to this have been reported by other authors (Zimmer et al, 1988; Fabijanski et al, 1990; Harris and Ingram, 1992 and Hughes and Harris, 1994).

Zimmer *et al* . (1988) artificially crossed *Zea mays* L. and *Z.luxurians* (DURIEO) BIRD. The F1 progeny of this cross were screened with the restriction enzymes *Eco*R I and *Sst* I. All of the F1 plants had additive *Eco*R I phenotypes, but 4 of the 12 individuals tested had only the *Z.mays Sst* I phenotypes, i.e. some of the F1 hybrids apparently lacked the rDNA genes from *Z.luxurians* when tested with one enzyme, *Sst* I, but not when tested with the *Eco*R I.

Zimmer offered no explanation for these curious results other than a need to know more about the inheritance of rDNA.

Fabijanski *et al.* (1990) reported the presence of non-additive phenotypes in hexaploid *Avena* species using random repeat sequence probes but did not suggest an explanation for their occurrence.

Harris and Ingram (1992) also observed the occurrence of non-additive rDNA phenotypes in one population of *Senecio cambrensis* ROSSER (*S.vulgaris* L. x *S.squalidus* L.). All of the *S.cambrensis* surveyed had additive *Bam*H I phenotypes, but 5 out of the 6 individuals tested had only the *S.vulgaris* phenotype for the enzymes *Eco*R I and *Eco*R V.

Hughes and Harris (1994) again reported similar results, that is the presence of non-additive rDNA phenotypes in putative hybrid accessions between *Leucaena leucocephala* (LAM.) DE WIT subsp. *glabrata* (ROSE) S.ZARATE and *Leucaena esculenta* (MOC. *et* SESSÉ *ex* A.D.C) BENTH. subsp. *esculenta*. Of the eight enzymes used, three (*Apa* I, *Nsi* I, *Stu* I) showed additive phenotypes for all of the accessions and the other five (*Bcl* I, *Bgl* II, *Eco*R I, *Hin*D III and *Sac* I) gave non additive profiles for some, but not all, of the putative hybrid accessions.

The reason for the appearance of non-additive phenotypes in investigations of hybrids such as those outlined above, has not been fully explained. Flavell *et al.* (1986) suggest that results similar to this may be attributed to the sensitivity of certain enzymes to methylated DNA. Methylation of DNA involves the alteration of adenine and thymine residues in prokaryotes and only cytosine in eukaryotes (Campbell, 1991). The product of methylation of cytosine is 5-methylcytosine (Stryer, 1988). For example, the restriction enzyme *Eco*R I will not cut DNA that is methylated because the cutting site cannot be recognised. This is not a problem for cpDNA studies, as cpDNA is not methylated, wheras nuclear DNA can often be methylated. As another possible answer to the non-additive phenotypes encountered in their study

Harris and Ingram suggested that some *S.squalidus* plants at the site in question possessed *S.vulgaris* type rDNA, either as a rare DNA phenotype or to the exclusion of other phenotypes. Following the hybridisation event, rare rDNA phenotypes were amplified in some lines but not in others (Harris and Ingram, 1992).

Both these explanations are possible in the case of *A. laeta*. All of the enzymes selected for the full survey were methylation sensitive. Also due to the relatively small sample size of *A. senegal* it is perhaps possible that some *A. senegal* possesses the rDNA phenotype similar to *A. mellifera* and these *A. senegal* were the parents of the *A. laeta* surveyed.

5.6 Summary

This initial survey of *A. laeta*, a putative hybrid between *A. senegal* and *A. mellifera*, has produced some interesting results.

The chloroplast DNA evidence suggests that *A. mellifera* is always the maternal parent. This is shown by all the *A. laeta* surveyed having *A. mellifera* cpDNA phenotypes.

The rDNA evidence supports a hybrid origin for *A. laeta*. This is shown by several trees of *A. laeta* displaying additive inheritance of the parental phenotypes for the PEC *Eco*R I/pTEE 5. However, in the four other PEC that separated the parental taxa, the putative hybrid gave non-additive profiles resembling the phenotype of *A. mellifera*. This result cannot be fully explained at present.

The presence of non-additive phenotypes in the hybrids does have implications for the use of rDNA restriction fragment analyses in identifying hybrids. Reliance on one or two enzymes in hybrid identification may lead to errors. "The choice of restriction enzyme therefore appears to influence whether a hybrid taxon will produce an additive DNA profile, indicating that a

suite of nuclear taxon-specific markers should be used ..." (Hughes and Harris, 1994).

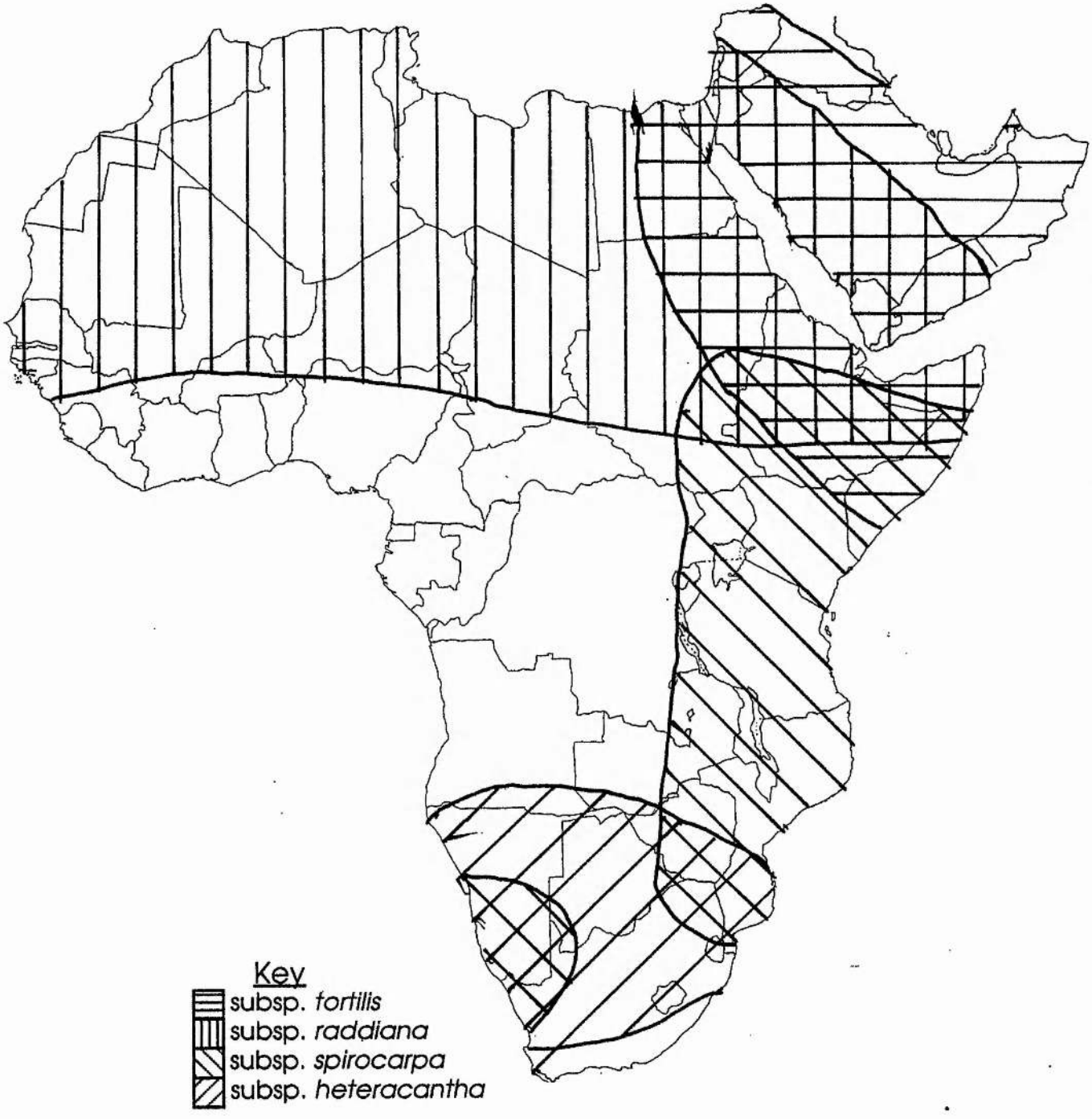
The result gives backing to El Amin (1976) in his suggestion that *A. laeta* is a hybrid between *A. senegal* and *A. mellifera*, though the need for further research as outlined in the discussion should be noted.

Chapter 6
RAPD analysis of *Acacia tortilis.*

6.1 Introduction.

Acacia tortilis (FORSSK.) HAYNE is one of the most widespread *Acacia* species in the savannah and seasonally dry zones of Africa and the Middle East. Over much of its range it is recognised by its characteristic spreading crown, giving rise to its popular name of 'umbrella thorn' (see plate 6.1 following page 145). It is dominant in many savannah communities, providing important sources of fodder and browse for wild and domesticated animals. It is also suitable for stabilising sand dunes (Fagg and Barnes, 1990).

It is a very variable species, containing 6 infraspecific taxa, including 4 subspecies; *Acacia tortilis* subsp. *tortilis*, *Acacia tortilis* subsp. *spirocarpa* (HOCHST. *ex* A.RICH.) BRENAN var. *spirocarpa*, *Acacia tortilis* subsp. *spirocarpa* var. *crinita* CHIOV., *Acacia tortilis* subsp. *heteracantha* (BURCH.) BRENAN, *Acacia tortilis* subsp. *raddiana* (SAVI) BRENAN var. *raddiana* and *Acacia tortilis* subsp. *raddiana* var. *pubescens* A. CHEV. (Brenan, 1983). The subspecific taxa are largely allopatric, overlapping on the borders of their ranges. Subsp. *tortilis* occurs from Sudan to Somalia northwards to Egypt and Israel and extends into Arabia as far as Iran (see map 6.1); subsp. *spirocarpa* is the dominant taxon over most of east Africa from Somalia and Sudan southwards to Namibia and Botswana; subsp. *heteracantha* is restricted to southern Africa, extending as far north as Zimbabwe and Angola; subsp. *raddiana* occurs throughout the countries of northern and western Africa, fringing the Sahara to Sudan and extending through Egypt to Israel, Jordan and Saudi Arabia (Fagg and Barnes, 1990). Some authors consider subsp. *raddiana* a separate species (*A. raddiana* SAVI), but recent revisions have treated it as a subspecies of *A. tortilis.*



Map 6.1 (from Brenan 1983). This map shows the approximate distributions of the four subspecies of *Acacia tortilis* in Africa and Arabia.

A. tortilis varies from multistemmed shrub forms (subsp. *tortilis*, plate 6.2, following pages) to tall distinctive trees of over 20 metres with rounded crowns (subsp. *raddiana*, plate 6.1) or flat-topped crowns (subspp. *heteracantha* and *spirocarpa*, plate 6.3) (Fagg and Barnes, 1990). Although the species is distinctive in the field, differences between the infraspecific taxa are not very clear (Fagg and Barnes, 1990). For a detailed description of the taxa see Ross (1979).

A. tortilis has many uses that benefit rural communities (Fagg and Barnes, 1990). The pods are a highly nutritious fodder, especially at the end of the dry season (see plate 6.4). It provides an important source of shade and its prolific flowering means that it is used as a honey source (plate 6.3). In Sahelian countries it is used extensively as a source of charcoal and firewood, being one of twelve species preferred and used by most rural populations (Von Maydell, 1986, see plate 6.5). Young thorny branches are used extensively for livestock kraals (see plate 6.6) and the longer trunks are used as poles in house construction.

A. tortilis was researched by Olng'otie (1991) who studied the genetic diversity of the species. His study, based on isozyme analysis, revealed a high level of genetic variation, 95% of the loci studied being polymorphic. The majority of the total gene diversity, 82%, was apportioned within the populations studied (Olng'otie, 1991). No attempt to investigate the validity of the infraspecific ranks for *A. tortilis* was attempted by Olng'otie.

This chapter deals with the attempt to clarify whether the genetic difference between subspecies can be recognised using RAPDs. The production of a dendrogram of the subspecies and the identification of taxon specific markers using RAPDs is described.

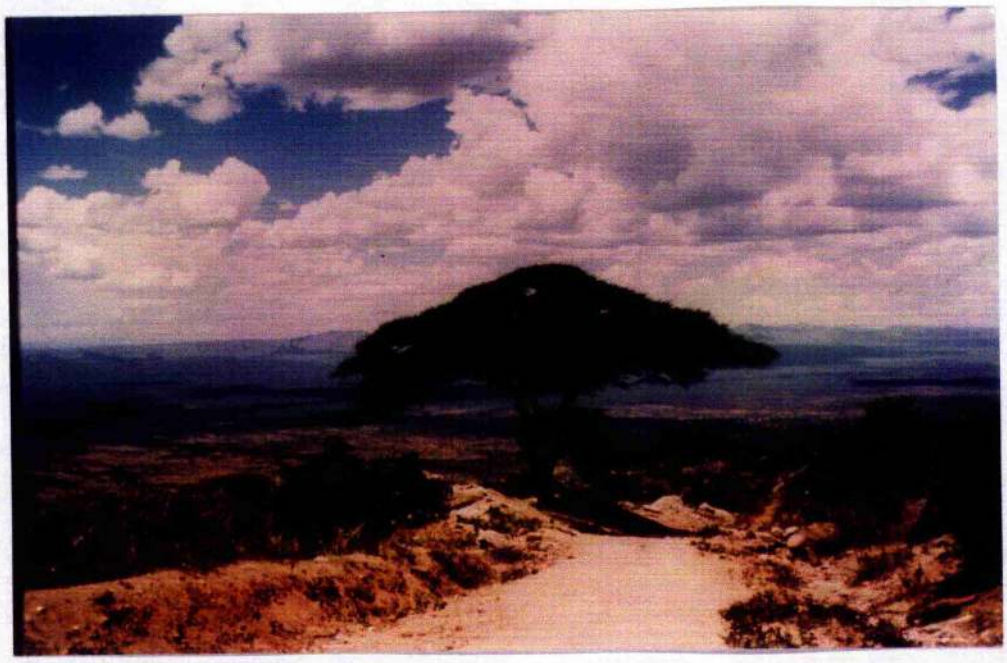
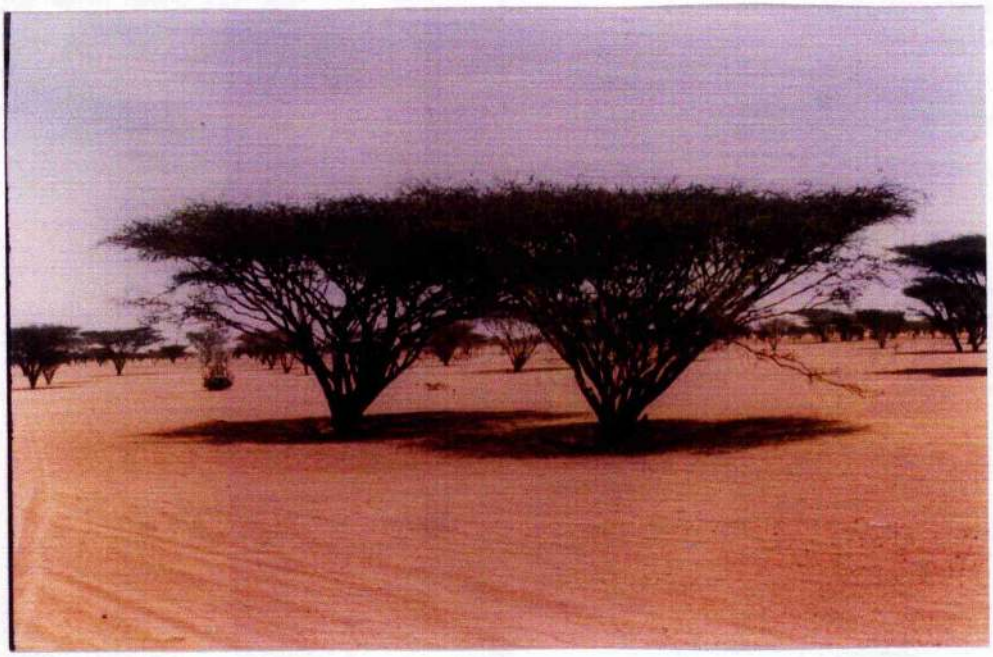


Plate 6.1 ▲. *Acacia tortilis* subsp. *raddiana* growing in the Cherangany Hills, Kenya. Photo from C.Fagg OFI.

Plate 6.2 ▼. *Acacia tortilis* subsp. *tortilis* growing at the Ben Naga Forest Reserve, Sudan. Photo from C Fagg OFI.



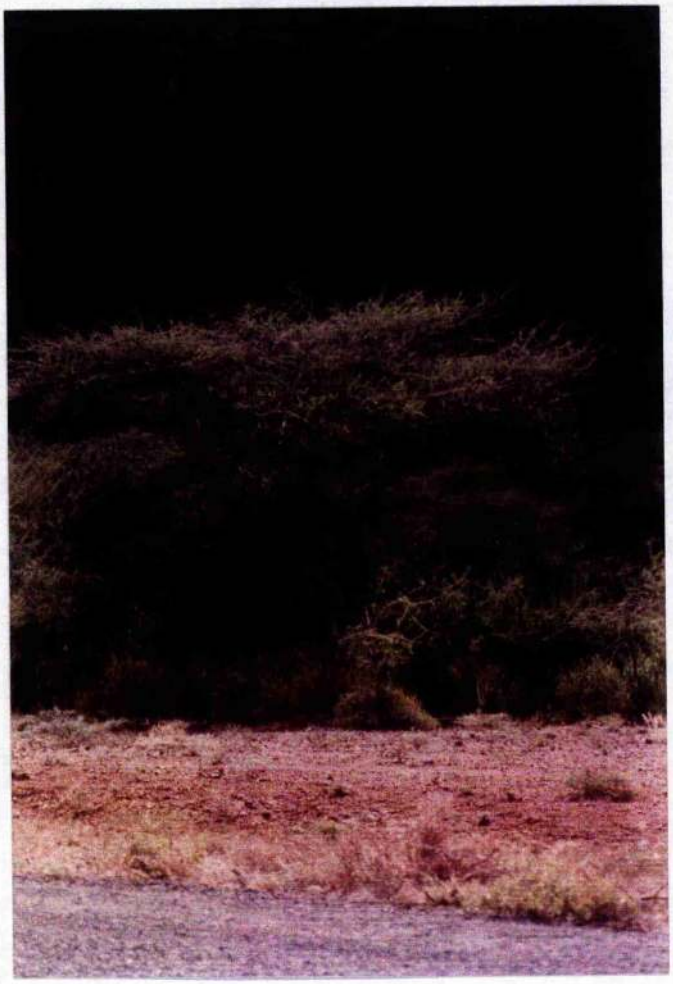


Plate 6.3 ▲. *Acacia tortilis* subsp. *spirocarpa* growing in Kenya. The dark object underneath the crown of the tree is a beehive mounted there by local farmers.

Plate 6.4 ▼. Domesticated goats eating the pods of *Acacia tortilis* subsp.*heteracantha*, Botswana. Photo from C.Fagg OFI.

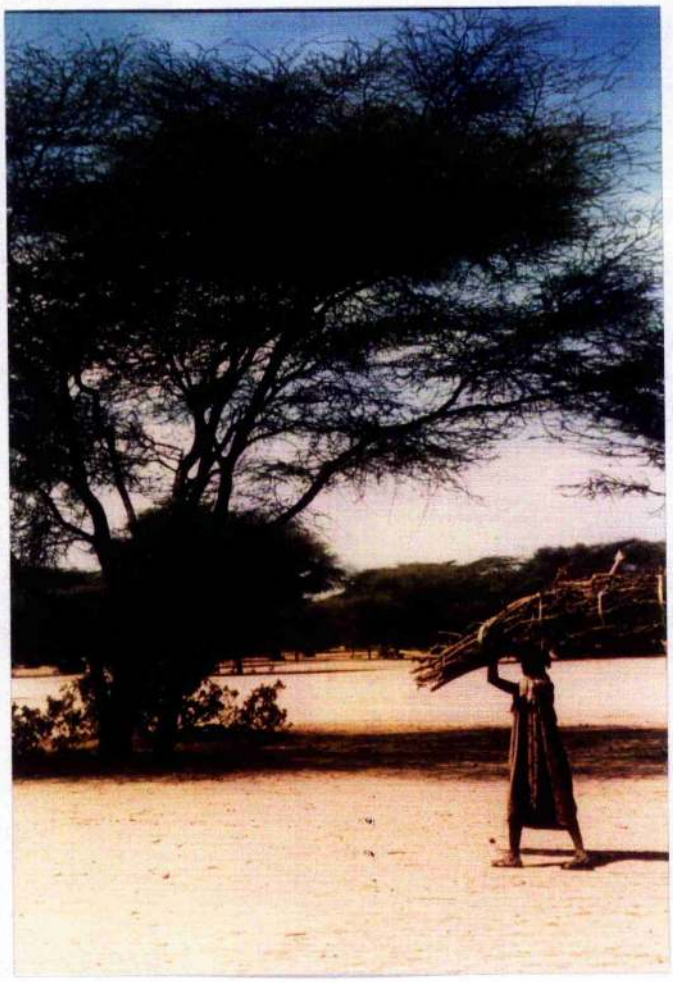


Plate 6.5 ▲. Woman collecting firewood in Turkwell, Kenya. The bundle of wood appeared to be *Acacia tortilis*. Photo from C.Fagg OFI.

Plate 6.6 ▼. Fences of *Acacia tortilis* on degraded land around huts in Njempi Flats, Kenya. Photo from C.Fagg OFI.

6.2 The Polymerase Chain Reaction.

The polymerase chain reaction (PCR), which was developed in the late 1980s, is an *in vitro* method for the synthesis of multiple copies of DNA sequences (Cherfas, 1990). The starting material for a PCR is a gene or segment of DNA. This target sequence is amplified exponentially. This is accomplished in three distinct steps. Initially the double stranded DNA molecule is separated into single strands by heating. Then two small sequences of synthetic DNA each complementing a specific sequence at the end of the target sequence in the single strand DNA serve as primers for DNA polymerase. This starts at each primer and copies the DNA strand. When this is completed, the process begins again with the denaturation step. In each reaction cycle the number of copies of the sequence of interest is doubled, thus producing multiple copies of this sequence.

The major step in the evolution of the PCR process was the discovery of a thermostable DNA polymerase. Before this the Klenow fragment of DNA Polymerase I had usually been used as the DNA polymerase. This had two drawbacks; i) additional enzyme had to be added in every cycle, as the denaturation step inactivated the enzyme and ii) amplification products from mismatched primers were common due to the low stringency conditions caused by having the polymerase phase at 37°C, which is the optimum temperature for the Klenow fragment (Erlich, 1989). Initially the thermostable enzyme *Taq* polymerase was isolated from *Thermus aquaticus*, a bacterium which grows in hot spring water, but recently polymerases have been isolated from a range of thermophilic bacteria e.g. *Thermus thermophilus* and *T.brockianus*. The enzymes differ in their precise tolerance of heat, but on average their DNA polymerase is stable at the denaturation temperature of 94°C, with a half life of approx. 40 min (Gefland, 1989). The optimum temperature for polymerase activity is approx. 72°C, well above the temperature required to prevent primer mismatches

The PCR procedure has been widely discussed. General introductions are those of Erlich (1989) and Innis *et al.* (1990) who cover the whole range of PCR applications. This introduction will consider the impact of the PCR process on studies of plant evolution and diversity only. The studies can roughly be divided into two types; those in which there was prior knowledge of the sequence amplified and those that amplified an unknown sequence. In both types of study a major consideration has been the ease of the PCR protocol. It is easy to use because only small amounts of relatively unpurified DNA per taxon are required for PCR analysis i.e. in the nanogram range as opposed to the 300-500 µg of highly purified DNA needed for a restriction site analysis. This is advantageous when working with taxa where the amount of material is limited e.g. herbarium material or seedlings. The PCR reaction is also relatively quick and simple. Once DNA has been extracted, results can be obtained in one day, rather than the days and weeks required with traditional molecular approaches. Because the technique is relatively simple, large numbers of samples can be analysed at once.

6.2.1 Applications where the PCR product was known.

The restriction site approach to phylogeny reconstruction was taken a step further by Liston (1992), who used evidence from restriction site mapping of a PCR amplified fragment to reconstruct a phylogeny of *Astragalus* species. The amplified region encompassed the chloroplast genes RNA polymerase C1 (rpoC1), RNA polymerase C2 (rpoC2) and the intergenic spacer between the two genes. The PCR product was approx. 4100bp. Liston (1992) attempted to digest this product with 32 different restriction enzymes. Of these, only 23 cut the amplified product. Approximately 144 restriction sites were identified, 37 of which were informative. In addition a 10bp insertion was identified. Phylogenetic analysis of these restriction site characters resulted in a single 41-step phylogenetic tree, which was consistent with previous classifications of

Astragalus (Liston, 1992). Liston (1992) concluded that this study demonstrated the utility of restriction site analysis of PCR amplified cpDNA in the study of plant phylogenetic relationships and molecular evolution.

In addition to phylogenetic reconstruction, the PCR approach has been utilised for cultivar identification. Weining and Langridge (1991) were able to identify different cereal cultivars by their PCR profiles. The primers that they used were from the α-amylase gene and also several Intra Splice Junction (ISJ) primers. Unique banding patterns for several cultivars were observed. Another example is that of D'Ovidio *et al.* (1990) who detected genetic polymorphisms in wheat using the PCR process. As a primer D'Ovidio *et al.* (1990) selected a 20-mer sequence from the terminal sequences of a gamma-gliadin gene. Electrophoretic analysis of the PCR products showed specific bands which revealed both inter- and intra-specific genetic polymorphisms among the examined genotypes. D'Ovidio *et al.* (1990) suggested that this technique could be a very simple and efficient alternative to RFLP markers.

6.2.2 Application where the PCR product is unknown.

This encompasses a distinct field of study. The use of RFLPs as genetic markers has been a powerful technique for the generation of genetic markers. However it has distinct drawbacks. In practice the cost, level of technical skill required and the use of radioisotopes has limited its applications. In addition the level of variation revealed by RFLPs has sometimes not been high enough for them to be used as genetic markers. The main drawback of the PCR process for fingerprinting or identifying genetic markers is the requirement for knowledge of the DNA sequence in the area one wished to amplify, in order to produce specific primers. Williams *et al.* (1990) and Welsh and McClelland (1990) simultaneously described a new technique that they believed would overcome many of these limitations. This method was based

on the typical PCR process except that the primers were of an arbitrary sequence, usually around 10bp in length. The PCR products of these reactions were unknown. The polymorphisms between any two taxa result from the sequence differences in one or both of the primer binding sites and are visible as the presence or absence of a band of DNA (Rafalski and Tingey, 1993). The polymorphisms can also take the form of variation between the primer sequences, such as insertions or deletions. These are recognisable as changes in the size of the PCR product. The polymorphisms are inherited in a Mendelian fashion (Williams *et al.*, 1990) and are dominant markers (Rafalski and Tingey, 1993). The acronym RAPD, for Randomly Amplified Polymorphic DNA, has become the term to describe this technique. The RAPD technique has been successfully used to generate molecular markers for a wide variety of taxonomic and other investigations.

Crawford *et al* . (1993) used RAPD markers to document the origin of the intergeneric hybrid x*Margyracaena skottsbergii* BITTER. The hybrid is endemic to the island of Masatierra in the Juan Fernandez archipelago, and was thought to be a hybrid between the endemic *Margyricarpus digynus* (BITTER) SKOTTSB. and the introduced *Acaena argentea* RUIZ *et* PAVON. Previous studies utilising isozymes (Crawford *et al.*, unpubl. data) had been unable to prove or disprove this hypothesis. Crawford *et al.* (1993) used 13 decamer primers, which produced 18 consistent species specific bands for *A. argentea* and 27 for *M. digynus.* All 45 bands were present in the presumed hybrid x*Margyracaena.* This provided strong evidence (Crawford *et al.*, 1993) that x*Margyracaena skottsbergii* is an intergeneric hybrid between *A.argentea* and *M. digynus.* Both x*Margyracaena* and *M. digynus* are very rare plants and only small amounts of leaf material were available. The RAPD technique was ideal for this study as adequate material for a 'standard' restriction site analysis of cpDNA or rDNA was unlikely to have been obtained (Crawford *et al.*, 1993).

Phylogenetic relationships in the *Festuca/Lolium* complex were investigated using RAPD technology by Stammers *et al.* (1995). Three decamer primers were used to generate RAPD profiles from groups of genotypes of several species. The degree of band sharing between taxa was used to calculate genetic distance (Nei, 1987). These values were then used to calculate a phylogenetic tree. The RAPD generated phylogenetic tree showed a number of "strong similarities" (Stammers *et al.*, 1995) to those generated in previous studies (based on both cpDNA data and morphological and seed protein analyses). However, there were several anomalies.

In the latter part of the paper Stammers *et al.* (1995) addressed the concerns of many workers about the use of RAPD markers for phylogenetic inference. The concerns are threefold: i) the extent to which it is possible to infer homology of bands that show the same electrophoretic mobility; ii) the causes of variation in the fragment mobility and iii) the origin and repetitive status of the sequences under analysis (Stammers *et al.*, 1995). Stammers *et al.* (1995) investigated these concerns by excising RAPD generated bands from the agarose gels used and using them as probes in a Southern hybridisation of replica gel. This enabled them to assess the identity of fragments classified as 'shared'. Of the six bands investigated, four were found to be homologous, i.e. bands that were scored as identical probably had similar sequences. The other two bands were found not to be homologous between the two taxa which shared them, although they had a similar electrophoretic mobility. Stammers et al. (1995) concluded that a substantial proportion of the RAPD bands are likely to correspond to homologous sequences with conserved organisation, but they did not consider the effect non-homology would have on phylogeny reconstruction.

One must question whether 66% can be viewed as a substantial proportion. Moreover, the effect of such a low level of homology must call

into question the validity of the resulting phylogenetic tree. The use of non-homologous characters will certainly obscure many of the relationships between taxa investigated. Most other studies have assumed the homology of co-migrating bands rather than testing for it. Two studies that have tested for homolgy are; Wilkie et al. (1993) who found 100% homology between seven comigrating RAPD bands in *Allium* and Gillies (1994) who estimated the level of homology between comigrating RAPD fragments in *Stylosanthes* to be 58%. These two studies seem to reinforce the belief that homology cannot be assumed. It must be confirmed in each separate investigation.

The final example is that of Chalmers *et al.* (1992). This investigation looked at genetic variation between and within populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. Both *G. sepium* and *G. maculata* are multipurpose trees native to Central America and Mexico. Chalmers *et al.* (1992) suggest that accurate estimates of diversity between- and within-populations is a prerequisite for the optimisation of sampling and breeding strategies. Using eleven decamer primers and 50 samples from 10 populations Chalmers *et al.* detected extensive genetic diversity between the species, providing molecular support for the conclusion that *G. maculata* should be considered a distinct taxon. They did not attempt a phylogeny reconstruction with the data.

Most of the variation in *Gliricidia* occurred between populations. This was contrary to the results of Hamrick and Godt (1990) who suggested that for woody outbreeding plants most diversity exists within populations. Chalmers *et al.*(1992) noted that this suggestion was based on isozyme data, representing only coding regions of the genome and so may not be comparable with RAPD data. They also showed that primers differed in their capacity to detect variation within- and between-populations (Chalmers *et al.*, 1992). Population specific amplification products were identified. Introduced populations of *G. sepium* in Thailand and Venezuela had relatively low levels of polymorphism,

supporting their putative history of domestication and introduction from a limited genetic base. Chalmers *et al.* (1992) concluded that RAPDs provide a cost effective method for the precise and routine evaluation of variability and may be used to identify areas of maximum diversity of the taxa under investigation.

The above examples have shown the range of problems in plant diversity and evolution that can be solved or investigated using RAPD technology. To summarise; RAPDs can help in the identification of hybridity; they can be used to construct phylogenies, if suitable criteria are met such as testing the homology of co-migrating bands; and they can be of use in analysing genetic variability in commercially valuable species.

6.3 Aims.

The aims of this project were twofold, i) to investigate whether a RAPD analysis would discriminate between the taxonomic sub-groupings of *Acacia tortilis*, and ii) to obtain taxon specific and, if possible, accession specific markers.

To the best of our knowledge this is the first time that molecular markers have been used to try to discriminate between the subspp. of *A. tortilis*, an important multi-purpose tree (MPT). The results of this investigation could be of use in breeding programmes, if the distinctiveness of the individual subspecies is revealed. The identification of subspecific or accession specific markers will also be of use in breeding programmes.

The reasons for choosing a RAPD technique for this investigation were twofold:

i) Although it is a relatively recent technique many papers have been published advocating the use of RAPDs in investigating genetic diversity due to its cheapness and ease of use.

Subspecies	Accession Ident No.	Locality	Country
tortilis	§	Negev Desert	Israel
tortilis	1065/82	Ahwar	Yemen
tortilis	A/3	Wadi Ghaba	Saudi Arabia
raddiana	§	Eilat	Israel
raddiana	1402/84	Guidick	Senegal
raddiana	1240/84	Bara	Sudan
raddiana	86/5555.5574	Filiugue	Niger
raddiana	§	Wadi Ghafar	Saudi Arabia
raddiana	72/92	Lamu	Kenya
heteracantha	69/92	Sesfontein	Namibia
heteracantha	30/90	Gaborone	Botswana
heteracantha	29/92	Malapati	Zimbabwe
heteracantha	27/92	Gungundhlovu	South Africa
spirocarpa	70/92	Turkwell	Kenya
spirocarpa	131/91	Chinzombo	Zambia
spirocarpa	15/92	Mtandika	Tanzania
spirocarpa	1340/87	Khartoum	Sudan
spirocarpa	110/87	Bay Region	Somalia
Acacia planifrons	T.A.N.U	Karnataka	India

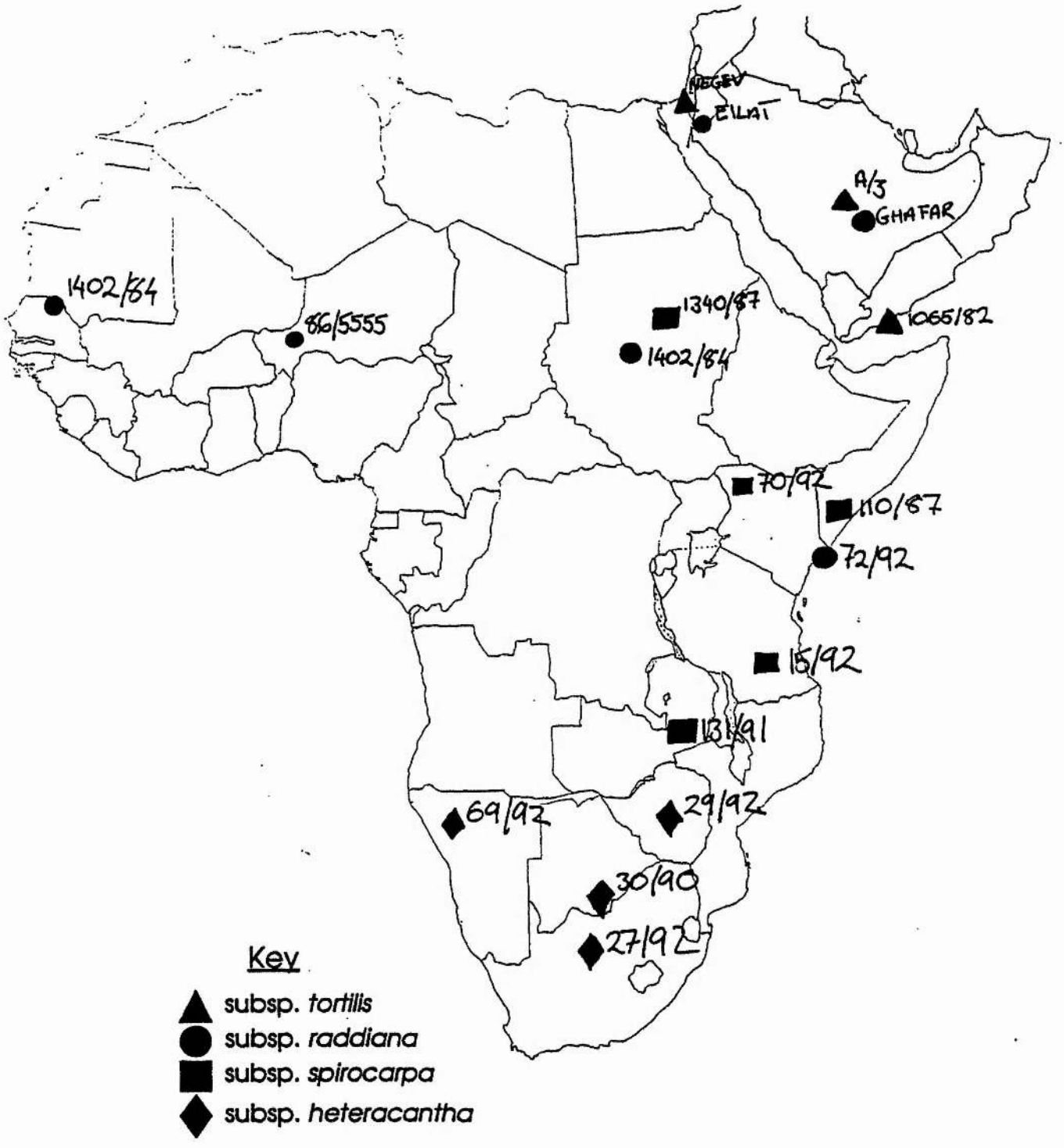
Table 6.1. This table shows the origin of the taxa used in this investigation. All the seeds for this investigation were obtained from the Oxford Forestry Institute. Several accessions were without accession numbers when sent. These are indicated with a §.

ii) The RAPD technique would serve as a pilot study to investigate firsthand the use of RAPDs, and to see how useful they could be in the study of woody Mimosoids, many of which are important MPTs.

6.4 Methods and Materials.

6.4.1 Plant material.

Accessions of *A. tortilis* were selected for the study which represented a sample of the distribution range in each subspecies. The accessions that were selected can be seen in table 6.1 and in map 6.2. The accessions were bulk collections from over 25 single parent identified trees, with a wide spacing between trees. This methodology is further outlined in Hughes (1987). *A. planifrons* was included as an outgroup taxon. This species is very similar to *A.*



Map 6.2. This map shows the approximate origin of the accessions studied. The accession numbers over the symbol identifying the subspecies are taken from table 6.1.

tortilis (pers. com., Chris Fagg OFI), but is only found only in the Indian sub-continent.

6.4.2 The Polymerase Chain Reaction.

The methodology for the extraction of DNA for the PCR analysis and for the PCR reaction itself can be found in Appendix B. Several different DNA extraction procedures were attempted. These included a CTAB extraction as used for DNA extraction in the cpDNA work, a scaled down version of the CTAB extraction procedure in microcentrifuge tubes excluding CsCl purification, and the mini-prep method of Edwards *et al.* (1991). Finally a modification of Edwards *et al.* (1991) was used. To reduce the chances of contamination of the sample with foreign DNA, leaf material was not used. Instead cotyledons were removed from the seeds under sterile conditions. No differences in amplification products between DNA extracted from cotyledons or leaves was observed for the few taxon/primer combinations tested. Initially the method of Edwards *et al.* (1991) did not produce DNA that could be amplified, but because of the ease of the procedure and the low cost of the reagents we persisted with this method and found that by adding two more purification steps, i.e. a phenol precipitation and an ethanol wash, DNA could be amplified.

Oligonucleotide primers.

The choice of primers that are used in a RAPD study is important. Different levels of diversity are revealed by different primers (see Chalmers *et al.*, 1992). Due to limited resources only twenty primers were available for this study , i.e. those in Operon Technologies Kit H (the sequences of these 20 are in Appendix B). Each of these primers was tested. From the 20 primers, 10 were selected. The criteria for selection were successful amplification of DNA from

Primer number.	Sequence (5'-3').
OPH-01	GGTCGGAGGA
OPH-02	TCGGACGTGA
OPH-03	AGACGTCCAC
OPH-05	AGTCGTCCCC
OPH-07	CTGCATCGTG
OPH-12	ACGCGCATGT
OPH-13	GACGCCACAC
OPH-14	ACCAGGTTGG
OPH-16	TCTCAGCTGG
OPH-17	CACTCTCCTC

Table 6.2. This table shows the sequences of the primers that were selected for the investigation. The primer number refers to its Operon Technologies identification number.

the studied taxa and the production of banding patterns that appeared to be scorable. The sequences of these primers can be seen in table 6.2.

Reaction conditions

The reaction conditions can be found in Appendix B. The cycling conditions were selected from published papers in which the stated aims were similar to those of this investigation (Crawford *et al.* 1993; Chalmers *et al.*, 1992; Adams and Demeke 1993; Kazan *et al.*, 1993; Huff *et al.*, 1993).

6.4.3 Data analysis.

Each accession was coded for the presence or absence of specific RAPD bands, and a binary data matrix was constructed. From this matrix pairwise genetic distances were calculated using PAUP 3.1.1 (Swofford, 1993). This method of calculating distances was chosen because it allowed the inclusion of missing data in the data matrix. The distances thus calculated were used in a cluster analysis by the Neighbour-Joining method of Saitou and Nei (1987) using PHYLIP 3.3c (Felsenstein, 1993). The Neighbour-Joining method was used because of its simplicity and because it allowed for unequal rates of change along branches (Avise, 1994).

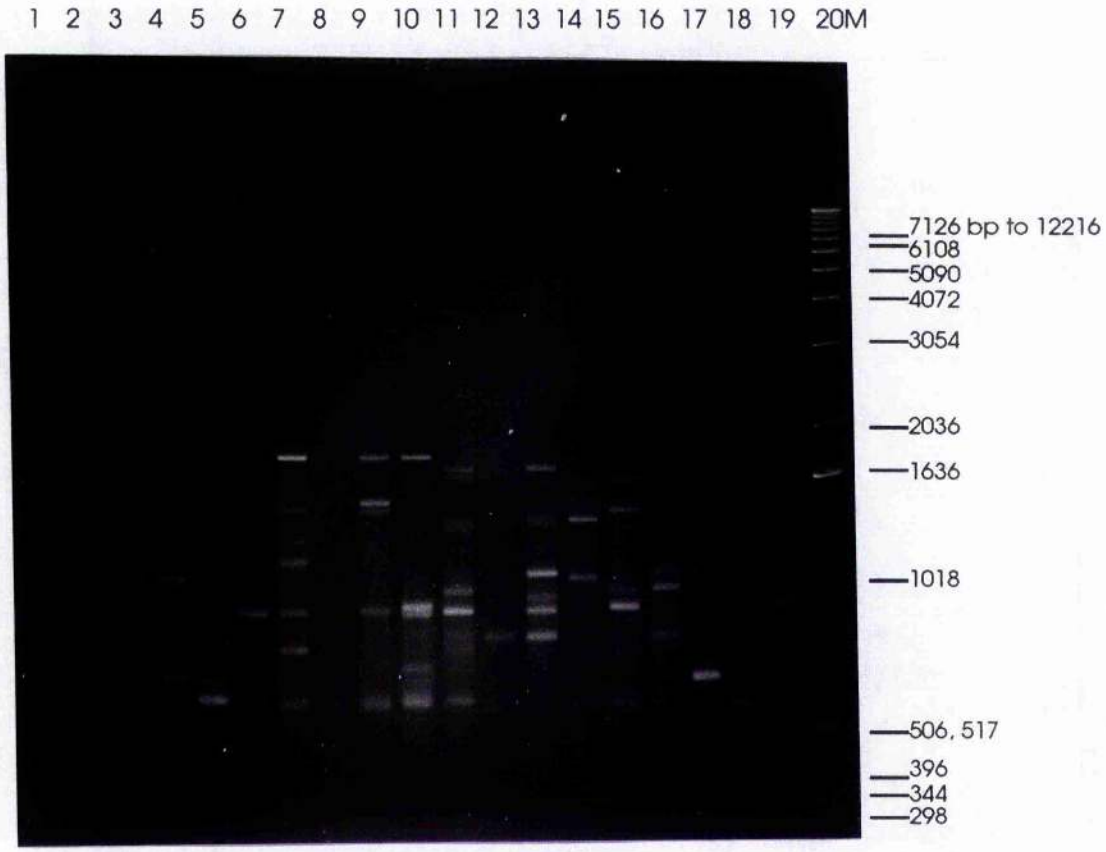


Plate 6.7 This plate shows the RAPD amplification profile for the primer OPH-05 and the *Acacia tortilis* accessions studied. In lane 1 is *Acacia tortilis* subsp.*heteracantha* 30/90, lane 2 subsp. *heteracantha* 69/92, lane 3 subsp. *heteracantha* 27/92, lane 4 subsp. *heteracantha* 29/92, lane 5 subsp. *raddiana* Eilat, lane 6 subsp. *raddiana* Lamu, lane 7 subsp. *raddiana* 86/5555, lane 8 subsp. *raddiana* 1402/84, lane 9 subsp. *raddiana* 1240/84, lane 10 subsp. *raddiana*, lane 11 subsp. *tortilis* Negev, lane 12 subsp. *tortilis* 1065/82, lane 13 subsp. *tortilis* A/3, lane 14 subsp. *spirocarpa* 70/92, lane 15 subsp. *spirocarpa* 110/87, lane 16 subsp. *spirocarpa* 1340/84, lane 17 subsp. *spirocarpa* 15/92, lane 18 subsp. *spirocarpa* 131/91, lane 19 *Acacia planifrons* and in lane 20 the size marker (Gibco BRL 1kb DNA ladder, the vertical ais is numbered in base pairs).

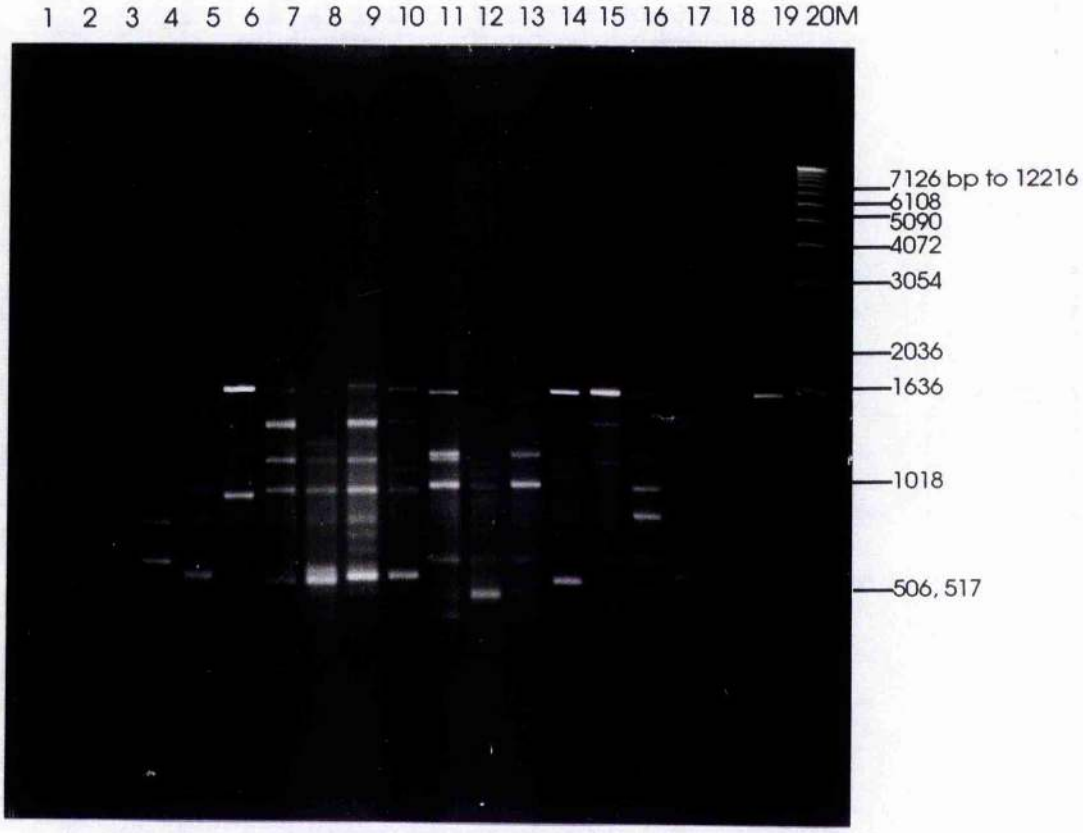


Plate 6.8. This Plate shows the RAPD amplification profile for the primer OPH-17 and the *Acacia tortilis* accessions studied. The lane order and accessions examined are exactly the same as in plate 6.7.

6.5 Results.

The ten primers that were selected amplified 161 bands between them. Of these 161 bands, 125 were informative. Of the rest three bands were constant throughout all the taxa studied and 33 were found in one accession only. The size of the DNA fragments, the primers from which they were amplified and the binary data matrix can be found in Appendix D. Plates 6.7 and 6.8 (following pages) show examples of the RAPD amplification profiles detected. The distance matrix calculated by PAUP from the binary data matrix can be seen in table 6.3.

The distance matrix was then used to calculate a dendrogram using PHYLIP. The dendrogram can be seen in figure 6.4.

Only two taxon specific bands were found, one for subsp. *raddiana* and one for subsp. *tortilis*. These were an amplification product of 1281bp with primer OPH-07 for subsp. *raddiana* and an amplification product of 1031bp with primer OPH-05 for subsp. *tortilis*.

6.6 Discussion.

6.6.1 Relationships of the taxa suggested by the dendrogram.

From the dendrogram (figure 6.4) it is apparent that all the individuals of each subspecies are not clustered together. The dendrogram appears to have two major clusters. The top cluster contains subsp. *tortilis*, subsp. *raddiana* and subsp. *spirocarpa*. The bottom cluster contains subsp. *heteracantha* and subsp. *spirocarpa*. All the subsp. *tortilis* and subsp. *raddiana* accessions are in the top cluster and all the subsp. *heteracantha* accessions are in the bottom cluster. Accessions of subsp. *spirocarpa* can be found in both clusters. Accessions 110/87, 1340/87 and 70/90 are in the top cluster of the dendrogram, and 131/91 and 15/92 in the lower cluster.

Taxa	het 30/90	het 69/92	het 27/92	het 29/92	rad eilat	rad lamu	rad 86/ 5555	rad 1402/ 84	rad 1240/ 84	rad wadi	tort negev	tort 1065/ 82	tort A/3	spiro 70/90	spiro 110/ 87	spiro 1340/ 84	spiro 15/92	spiro 131/ 91	*A.plani -frons*
het 30/90	-																		
het 69/92	0.217	-																	
het 27/92	0.255	0.241	-																
het 29/92	0.255	0.286	0.221	-															
rad eilat	0.304	0.286	0.331	0.335	-														
rad lamu	0.252	0.252	0.291	0.294	0.252	-													
rad 86/5555	0.317	0.298	0.324	0.385	0.211	0.266	-												
rad 1402/84	0.371	0.300	0.347	0.393	0.221	0.311	0.193	-											
rad 1240/84	0.292	0.273	0.331	0.323	0.149	0.266	0.161	0.221	-										
rad wadi	0.324	0.283	0.331	0.345	0.166	0.331	0.186	0.210	0.124	-									
tort negev	0.354	0.335	0.366	0.398	0.348	0.315	0.348	0.279	0.348	0.379	-								
tort 1065/82	0.282	0.275	0.310	0.380	0.338	0.298	0.331	0.240	0.352	0.389	0.211	-							
tort A/3	0.360	0.354	0.345	0.354	0.317	0.294	0.329	0.300	0.39	0.386	0.193	0.246	-						
spiro 70/90	0.348	0.329	0.372	0.379	0.280	0.280	0.342	0.314	0.304	0.372	0.342	0.317	0.298	-					
spiro 110/87	0.329	0.360	0.386	0.373	0.323	0.301	0.335	0.329	0.311	0.345	0.311	0.282	0.317	0.292	-				
spiro 1340/87	0.342	0.335	0.386	0.373	0.311	0.350	0.373	0.307	0.311	0.386	0.224	0.239	0.217	0.267	0.286	-			
spiro 15/92	0.315	0.287	0.291	0.343	0.343	0.329	0.301	0.328	0.329	0.299	0.392	0.347	0.399	0.399	0.406	0.385	-		
spiro 131/91	0.248	0.230	0.303	0.304	0.304	0.308	0.292	0.343	0.292	0.317	0.379	0.338	0.410	0.360	0.391	0.354	0.203	-	
A.planiforns	0.255	0.323	0.331	0.335	0.385	0.287	0.323	0.350	0.360	0.386	0.348	0.310	0.329	0.354	0.311	0.360	0.364	0.317	-

Table 6.3. This is the distance matrix used for the Fitch analysis by PHYLIP. The pairwise genetic distances were calculated by PAUP from a binary data matrix showing the presence or absence of specific RAPD bands. Only the lower half of the data matrix is shown.

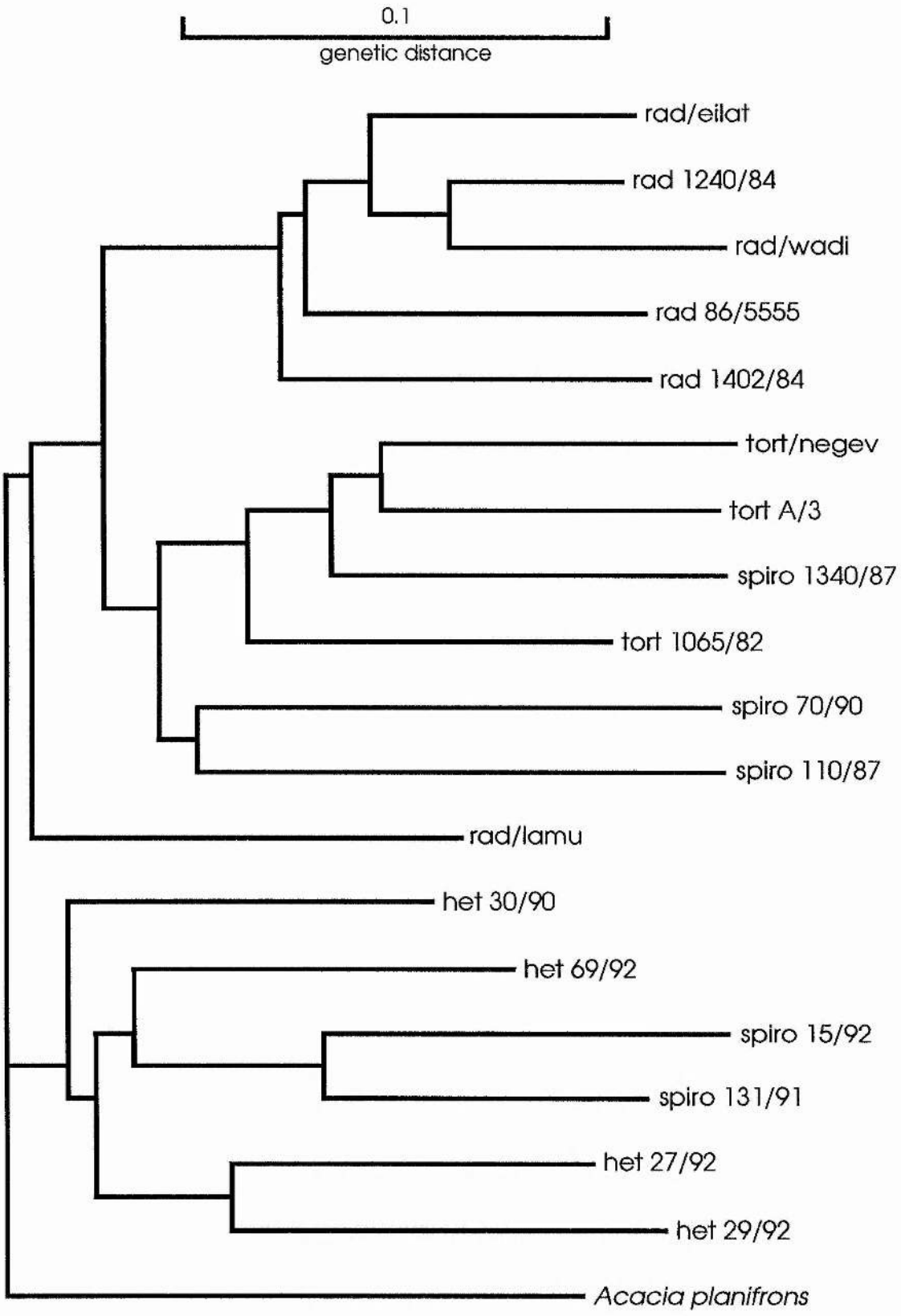


Figure 6.4. This figure shows the dendrogram calculated by PHYLIP. The tree is unrooted, but orientated so that *Acacia planifrons*, the putative outgroup, is basal to the other taxa.

However, if we look beyond the subspecific designation for each taxon and look instead at their geographic origin, a pattern emerges (see figure 6.5). The upper grouping consists entirely of accessions from northern Africa (northwards from Kenya), and the lower grouping consists of accessions from the south of Africa (southwards from Tanzania). The dividing line appears to be the border between the Afroriental domain and the Zambezian domain of the Sudano-Zambezian region (Brenan, 1978)[1]. This approximates to the border between Kenya and Tanzania. The distributions of the subspecies were mentioned in the introduction of this chapter, but are recapitulated here. Subsp. *raddiana* occurs over much of Sahelian Africa extending into Arabia. Subsp. *tortilis* reaches south to Somalia and is centred around the Red Sea extending into Arabia. Subsp. *heteracantha* is limited to the southern tip of Africa and subsp. *spirocarpa* is distributed down the east coast of Africa from Somalia to Mozambique. This information helps us understand the dendrogram. Two subspecies are found in north Africa only, subsp. *tortilis* and subsp. *raddiana*. These are limited to the upper cluster only. One subspecies is found in southern Africa only, i.e. subsp. *heteracantha*. This subspecies is found in the lower cluster only. The final subspecies is subsp. *spirocarpa* which has a distribution down the east coast of Africa from northern Africa to southern Africa. This is the subspecies that is found in both the upper (northern) cluster and the lower (southern) cluster. Accessions of subsp. *spirocarpa* from north Africa are found in the northern grouping of taxa, and accessions of subsp. *spirocarpa* from southern Africa are found in the southern grouping of taxa. This result will be discussed later. First I want to consider the relationships of the taxa within each of the two major clusters as implied by the dendrogram.

---

1 Brenan (1978) quotes from Wickens (1976) who partitioned Africa into phytogeographical regions. The Sudano-Zambezian region covers most of the tropical savannah in Africa. It extends from the Sahel regions of north Africa to the Karoo, Namib and Cape regions of southern Africa. In the middle of Africa it is constricted to east of Lake Victoria. West of this is the tropical Guineo-Congo region.

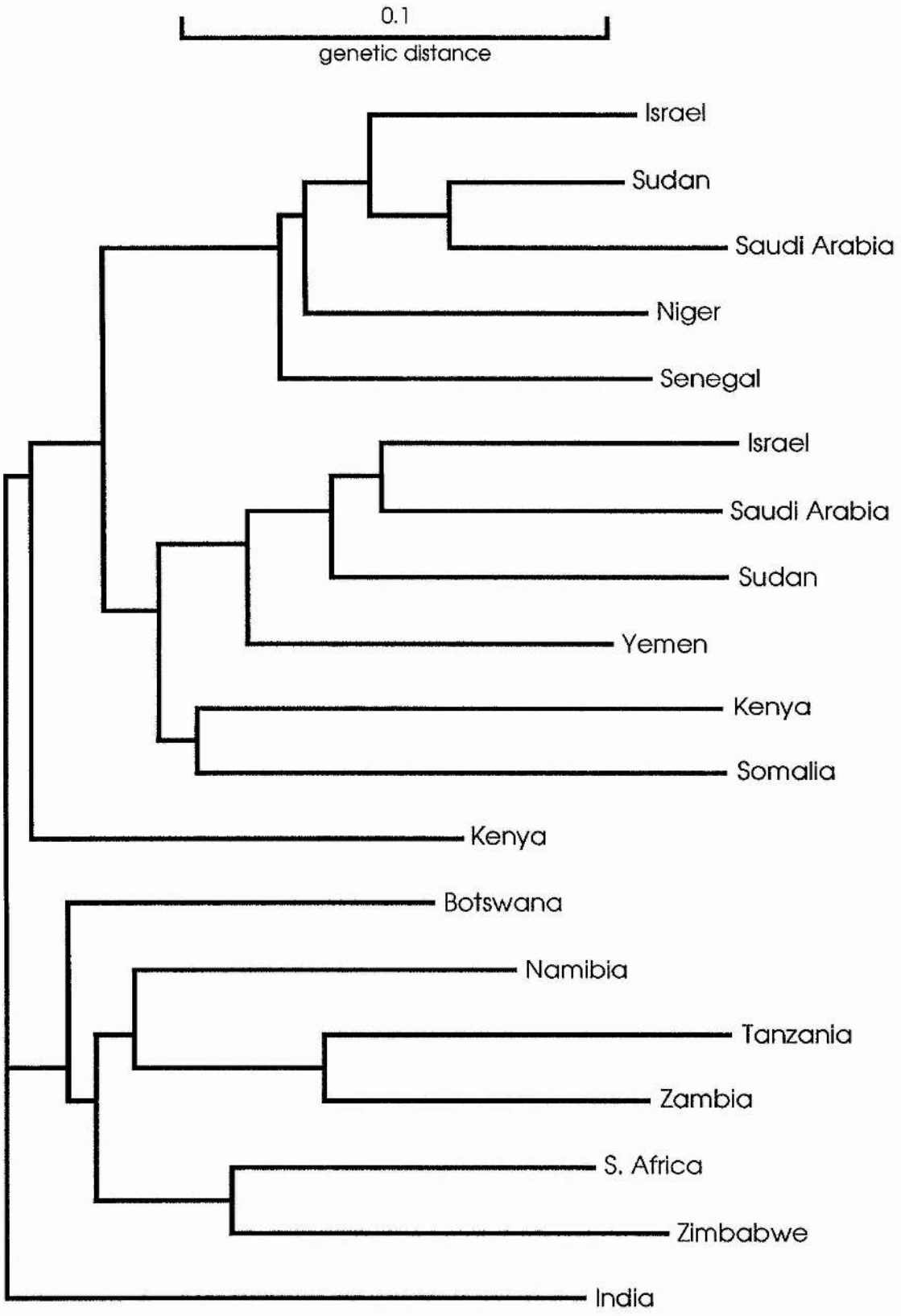


Fig 6.5. In this dendrogram the accession numbers and subspecies designation have been replaced with the geographical origin of that accession.

6.6.2 North African cluster.

In this cluster there appear to be two groupings, one consisting of subspp. *tortilis* and *spirocarpa*, and the other consisting of subsp. *raddiana*. The one accession that does not fit in to either of these two groups is an accession of subsp. *raddiana* from Lamu in Kenya. This accession is from the southern most limit of distribution of subsp. *raddiana* (see map 6.2). There are two possible explanations for its anomalous position. One possibility is that this accession is on the southern limit of subsp. *raddiana* range. It has been isolated as an island population (Lamu is an island off the coast of Kenya), and during this period of isolation this population of subsp. *raddiana* has possibly acquired novel mutations. This would account for its genetic distance from the rest of the north African grouping. Another possibility, which does not exclude the first, is that due to its peripheral position, it has interbred with other subspecies of *A. tortilis*. This again could have led to a differentiated genome, dissimilar to the rest of the north African accessions of *A. tortilis*.

Apart from the aforementioned accession, subspecies *raddiana* appears to be a distinct taxon, with all the accessions from a large geographical range grouping together, even though some of these accessions came from outwith Africa. This would appear to support the suggestion by some authors that subsp. *raddiana* is a distinct species, *A. raddiana*. However, this work suggests that it is grouped within *A. tortilis* and so should not be considered a separate species.

The other group within this northern cluster consists of a mixture of subsp. *tortilis* and subsp. *spirocarpa*. These two subspecies appear to be partly differentiated (see figure 6.6). Accessions 110/87 and 70/90 of subsp. *spirocarpa* are clustered together at the base of the subsp *tortilis/spirocarpa* grouping. The other grouping consists of all of the subsp. *tortilis* accessions plus one subsp. *spirocarpa* accession from Khartoum, Sudan -1340/87. This is nested within the subsp. *tortilis* group. In this subsp. *tortilis* group the closest

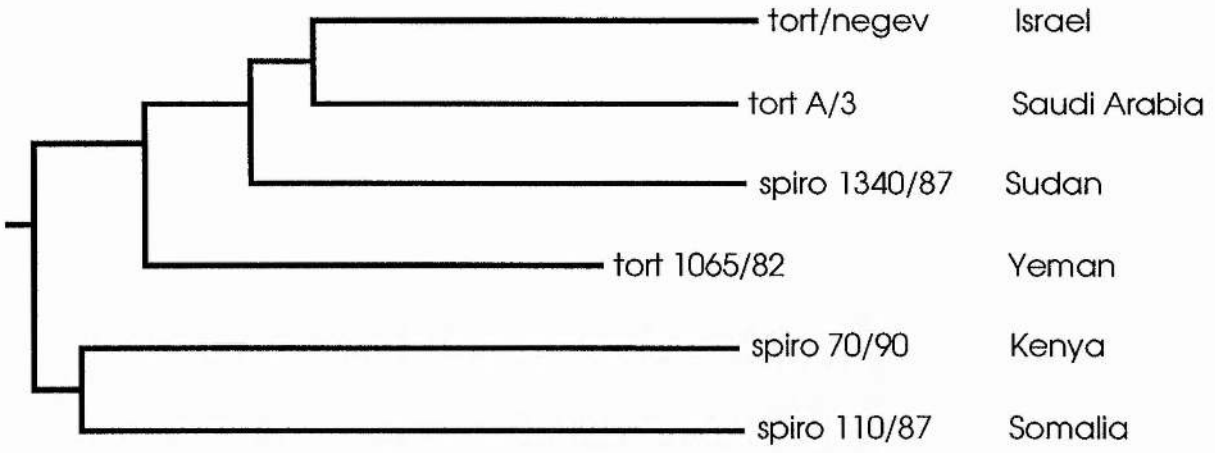


Figure 6.6. Showing the northern African subspp. *tortilis* and *spirocarpa* cluster of the dendrogram only. This cluster is taken from figure 6.4.

related accessions are those from Israel and Saudi Arabia. The subsp. *spirocarpa* accession is next and the subsp. *tortilis* accession from Yemen the least similar accession in this grouping (see figure 6.6).

The appearance of a subsp. *spirocarpa* accession within the subsp. *tortilis* cluster must be further discussed. There are two possible explanations for this observation. Firstly, there is an area of sympatry of subsp. *tortilis* and subsp. *spirocarpa* in the general region of Sudan, Ethiopia and Somalia. This is where the 'aberrant' subsp. *spirocarpa* accession comes from i.e. the northernmost limit of the distribution of subsp. *spirocarpa*. There could have been hybridisation between subsp *tortilis* and subsp. *spirocarpa* in the locality of this accession (1340/87), i.e. the area around Khartoum. The hybridisation event or events could have led to an accession which through introgression has the morphological features of subsp. *spirocarpa* with a genome characteristic of subsp. *tortilis*. An alternative to this hypothesis is that the accession of subsp. *spirocarpa* 1340/87 has been incorrectly identified. This misidentification hypothesis would suggest that accession no. 1340/87 is not subsp. *spirocarpa*, but is an accession of subsp. *tortilis*. If this was correct, then no explanation would be needed for its position on the dendrogram.

At present neither hypothesis can be discounted, and further investigation is required. The first avenue for investigation would be to check the voucher specimen for accession 1340/87 to see if it has been correctly identified. Unfortunately this was not possible in the timescale of this thesis as the seedlot for this accession was collected by the F.A.O. (Food and Agriculture Organisation, Rome, Italy), and the voucher was not readily available. If the accession has been correctly identified then an extended survey of taxa within the zone of sympatry might be a possible line of investigation.

6.6.3 South African cluster.

Again the associations of taxa within this cluster are not simple, this may be due to the small number of accessions studied from Southern Africa. In addition, any conclusions or hypotheses drawn from these accessions are, by virtue of the small sample size, tentative.

Two taxa occur in this, the southern, cluster, i.e. subsp. *heteracantha* and subsp. *spirocarpa*. These are the only two taxa which occur in southern Africa. The four subsp. *heteracantha* accessions appear at the base of this cluster, with the two subsp. *spirocarpa* accessions appearing within it. There does not appear to be any geographic partitioning of accessions within this cluster. The most closely related accession to the two subsp. *spirocarpa* accessions is subsp. *heteracantha* 69/92 from Northern Namibia. In terms of geographic proximity accession 29/92 and possibly 30/90 are nearer, although these represent subsp. *heteracantha*. In addition the accessions closely paired, subsp. *heteracantha* 29/92 and 27/92, are not geographically the nearest. If we look at Map 6.2 accession 30/90 lies approximately between these two accessions.

At present not enough information exists to explain the reasons for the non-homogeneity of the individual subspecies of *A. tortilis* in southern Africa. It

may be the case that in southern Africa the subspecies of *A. tortilis* are not genetically distinct, i.e. if further accessions, particularly of subsp. *spirocarpa*, were examined there might be no overall pattern to the relationships of these two subspecies. Alternatively this investigation may possibly be showing the introgression of subsp. *heteracantha* DNA into subsp. *spirocarpa*. Further investigation into both subspecies is needed to understand the nature of genetic variation in these two subspecies in southern Africa.

6.6.4 The distribution of subsp. *spirocarpa*.

The appearance of accessions of subsp. *spirocarpa* in both the northern and southern groups is a result that still needs an explanation. Another related problem is the question of the reality of the divide between the north and south African populations of *A. tortilis*. Unfortunately with the data I have it is difficult to address this latter problem. The study of a greater number of accessions from around what appears to be the dividing line would be helpful. However, these preliminary results suggest that there is little gene flow across the dividing line between the north African and the south African accessions at present. The basis of this divide is not readily apparent.

Gene flow between the northern and southern populations of subsp. *spirocarpa* has possibly also been very limited. This in turn could have led to the differentiation of the northern and southern populations of subsp. *spirocarpa*, which is the only taxon that straddles the dividing line. This could be the reason why subsp. *spirocarpa* occurs in both the clusters. The RAPD study has uncovered this genetic differentiation, which again needs further investigation.

6.6.5 Is hybridisation or introgression occurring?

Another question raised by this study is the extent of hybridisation or introgression occurring between the subspecies of *A. tortilis*. The subspecies of

A. tortilis are not reproductively isolated from each other, so that when the subspecies occur near each other it is possible that they may interbreed to some degree. The above results are only preliminary, but may possibly indicate that hybridisation or introgression is occurring between some subspecies of *A. tortilis*. The extent and effect of localised introgression requires further research. If evidence of localised introgression between the subspecies is uncovered it would help explain the differentiation between the north African and the south African populations of *A. tortilis*. If introgression is occurring then it is possible that populations of several subspecies in the same geographical locality will be more closely related than morphology might suggest, and further work on introgression is therefore needed.

Introgression is also a possible explanation for the division of subsp. *spirocarpa* between the northern and southern clusters. The accessions I chose from north Africa were in the region of the zone of sympatry between subsp. *spirocarpa* and subsp. *tortilis*. If introgression is occurring then there will be a leakage of subsp. *tortilis* genes into subsp. *spirocarpa* and vice versa. The end result of this could be the differentiation of subsp. *spirocarpa* into two entities, those that have introgressed with subsp. *tortilis* and those that have not. This could have led to the differentiation of a northern and southern race of subsp. *spirocarpa*. This again confirms the need for further work on introgression.

6.7 Summary.

This initial RAPD survey of *Acacia tortilis* in Africa and Arabia has produced some novel insights.

It appears that the populations of *A. tortilis* in Africa and Arabia are separated into two distinct groups: populations in north Africa and populations in south Africa, the dividing line corresponding approximately to the border between Kenya and Tanzania. The basis for this divide is unknown at present,

but appears to follow a boundary defined by Wickens (1976) in his division of Africa into phytogeographical regions. One subspecies, subsp. *spirocarpa*, straddles this boundary between northern and southern populations. This has produced some interesting results. Accessions of subsp. *spirocarpa* from the north of Africa group together with accessions of *A. tortilis* from north Africa. This situation is mirrored in southern Africa, where accessions of subsp. *spirocarpa* group together with accessions of subsp. *heteracantha* from southern Africa. The reasons for this division of subsp. *spirocarpa* are not clear at the moment and further research is required.

This preliminary survey of *A. tortilis* accessions also suggest that hybridisation/introgression is possibly occurring between the subspecies of *A. tortilis*, since introgression is the simplest explanation of the results.

Finally this investigation has demonstrated the utility of RAPD techniques in investigating plant evolution and diversity despite the fact that many of the initial claims for RAPDs are being questioned. This uncomplicated preliminary survey of *A. tortilis* has defined questions concerning the population genetics and dynamics of *A. tortilis* which should be further addressed.

Chapter 7
General Discussion.

The research detailed in this thesis has demonstrated the utility of molecular techniques in systematic and taxonomic investigations in the genus *Acacia*. There were three aims for this thesis, i) to produce a phylogeny of *Acacia* based on chloroplast DNA restriction site variation, ii) to investigate whether *A. laeta* was an interspecific hybrid and iii) to investigate whether RAPD techniques would differentiate the subspecies of *A. tortilis*. Each of these aims was realised to some extent, and the direction of further research indicated.

7.1 Phylogenetic relationships of the genus *Acacia*.

In Chapter 2, the history of the taxonomy of *Acacia* was reviewed. The review details the present conflict concerning the evolutionary relationships of the three subgenera of *Acacia*, i.e. subgenus *Acacia*, subgenus *Aculeiferum* and subgenus *Phyllodineae*. The conflicting hypotheses are based mainly on the morphological characteristics of the taxa. These characters are not as robust as they may seem, many of them being environmentally plastic, and determining homology between characters is difficult. A need was perceived for a classification of the genus which was based on more robust characters. To this end it was decided to construct a phylogeny based on molecular characters. It was hoped that this would enable us to comment on previous classifications and offer another viewpoint on evolutionary relationships in the genus. Previous work by other authors had shown the utility of the chloroplast DNA genome in resolving taxonomic problems of a similar nature. For this reason cpDNA restriction site variation was selected as the character to be used.

Using 15 restriction endonucleases, at least 559 enzyme sites were identified for the 72 taxa that were analysed. Of these 391 were found to differ between the taxa, and the computer program PAUP 3.1.1 was used to analyse the cpDNA restriction site data. The phylogenetic tree derived from these data was not in agreement with previous classifications of the genus.

The cladogram (figure 4.2) suggested that the genus *Acacia* is polyphyletic. The two subgenera, subgenus *Acacia* and subgenus *Aculeiferum* appear to be closely related, possibly derived from the same stock. The third subgenus, *Phyllodineae*, does not appear to be closely related to either of the other two subgenera of *Acacia*. Rather subgenus *Phyllodineae* appears to be related to taxa in the Ingeae, another tribe of the Mimoseae. This contrasts with morphological data which suggests that subgenus *Aculeiferum* and subgenus *Phyllodineae* are very closely related, possibly basal to the Ingeae (Chappill and Maslin, 1995). This classification (Chappill and Maslin, 1995) also suggests that subgenus *Acacia* is closely related to genera within the Ingeae.

Rather than clarifying the relationships of taxa within the Acacieae and the Ingeae, the cpDNA data suggest a novel classification of both tribes. This was unexpected, and rather than simplifying relationships, the cpDNA data have further complicated the picture.

The taxonomic structure suggested by Chappill and Maslin (1995) is, as the authors note, based on a preliminary investigation. Probably due to time limitations, the authors were not able to fully investigate each taxon studied and there is a certain amount of missing data. There were also internal contradictions between the results of their investigations. Until their investigations are fully completed there is little point discussing the relative merits of the two approaches.

Moreover, the possibility exists that both sets of suggested relationships are correct. The cpDNA molecule is not like an inflorescence, i.e. fixed to the plant on which it occurs. It is a molecule, which within defined parameters, is

able to move between taxa. This characteristic and its implications for phylogeny reconstruction are only now being recognised. Many initial studies that used cpDNA as a character reaffirmed previous investigations based on morphological, chemical or karyological characters. However, as time has progressed and the mechanics of cpDNA inheritance are better understood, more investigations are appearing the results of which cannot be simply explained. An example is the work on *Helianthus* by Rieseberg *et al.* (1988).

"It is becoming increasingly apparent" wrote Rieseberg and Soltis (1991) "that DNA phylogenies are often discordant with organismal phylogenies". Rieseberg and Soltis (1991) list contributory factors which they believe are responsible for the discordance. A factor which concerns us here is "cytoplasmic transfer between major evolutionary lines during early stages of their divergence", this according to Rieseberg and Soltis (1991) has the "potential to impact cpDNA phylogeny reconstruction". By this, they mean that introgressive or hybridisation events during early stages of divergence between taxa have the possibility of giving misleading results. In addition to this we must consider a form of lineage sorting of cpDNAs, where taxa may share similar cpDNA molecules but be only distantly related. The apparent close relationship would be due to these two taxa having randomly kept the same cpDNA molecule, which is derived from their distant ancestor. The taxa which are actually closely related to these taxa have randomly lost this cpDNA molecule and therefore appear dissimilar when the cpDNA molecule only is considered.

These factors may be influencing the phylogeny of *Acacia* considered here, and may account for the contrast between the cpDNA derived phylogeny and the morphologically derived phylogeny of Chappill and Maslin (1995). Further comment has to await the completion of the morphological work, as well as further investigations into the morphological implications of the cpDNA work.

The cpDNA data, at their face value, suggest that a new classification of the genus *Acacia* is desirable. The tribe Acacieae would be kept, consisting of the genus *Acacia*, with two subgenera, subgenus *Acacia* and subgenus *Aculeiferum*. This tribe appears intermediate between the tribes Mimoseae and Ingeae, as has been suggested before (Elias, 1981). The other subgenus of *Acacia*, subgenus *Phyllodineae* would have to be transferred to the Ingeae with a new generic name. The monotypic genus *Faidherbia*, rather than being in the Acacieae as has been suggested (Vassal, 1972), appears closely related to the Ingeae, and perhaps should be included in this tribe. These suggestions for the classification of the Acacieae and Ingeae are only outlines for a possible classification. Further work is required to help explain the apparent dichotomy of the morphological and cpDNA data, and until the conflict is resolved no nomenclatural changes are recommended.

From this we can see that the aim of this investigation has been realised. A phylogeny of *Acacia* based on cpDNA characters has been produced, which has enabled comparisons with previous classifications. However, the anomalous suggestions of these results mean that no firm conclusions concerning the evolution of the genus can be reached.

At the beginning of the investigation an aim was to survey the diversity of the African Acacias, from both subgenus *Acacia* and subgenus *Aculeiferum*. To some extent the present results have enabled us to clarify the diversity of these taxa. The results must take into account two factors. Firstly the wide taxonomic range of the accessions used for the cpDNA study may mean that many of the infrasubgeneric relationships could be influenced by homology between characters, i.e. the probability of the same cpDNA restriction site mutation (especially the loss of a restriction site) in two unrelated lineages is high due to the wide taxonomic base of the study. This has the potential to obscure and confound relationships. The second factor is that the variation in cpDNA exhibited at the level of species may not be enough to

elucidate species relationships. Both factors are contributory to some of the problems found in the infrasubgeneric relationships.

In subgenus *Acacia* only the accessions from the New World were fully resolved. Accessions from Africa were unresolved. This was probably due to the lack of variation in cpDNA restriction sites in these taxa. Further investigations may have to use molecular markers which exhibit greater variation at this taxonomic level, if relationships are to be fully resolved. The relationships between accessions from the New World as suggested by the cpDNA data, appear to correlate with relationships suggested by multi-variate analysis of morphological characters. Within the New World, species of *Acacia* can be divided roughly into five groups (Maslin and Stirton, in press). The results of this present investigation give some support to these groupings and also to the suggested contents of these groups.

The relationships of taxa in subgenus *Aculeiferum* were less obvious. The cpDNA data did succeed in resolving taxa but the meaning of the resolution was less clear. The results gave little support for the division of the subgenus into three sections as suggested by Vassal (1972). Section *Filicinae* as proposed by Vassal (1972) was monophyletic, though it appeared to be derived from within one of the other sections. The other sections proposed by Vassal (1972), section *Monacanthea* and section *Aculeiferum*, appear to be inter-related with accessions from section *Monacanthea* appearing within section *Aculeiferum*. The accessions from subgenus *Aculeiferum* did not separate according to their geographic origin either. The lack of resolution of the accessions of subgenus *Aculeiferum* into three sections may indicate that the supraspecific classification of subgenus *Aculeiferum* may need re-evaluating. It is possible that the relationships of taxa within this subgenus have been complicated by homologous mutations in other parts of the cladogram. The relationships of taxa within subgenus *Aculeiferum* thus require further investigation.

Within subgenus *Phyllodineae* the relationships of the accessions were impossible to compare with previous subgeneric classifications of this subgenus. In subgenus *Phyllodineae* there are approximately 900 species, this investigation only looked at six of these. Therefore no conclusions could be drawn from these accessions of subgenus *Phyllodineae*.

Within subgenus *Aculeiferum* there are several taxa which display what are considered to be the ancestral morphological characteristics of the genus (Ross, 1981). Subgenus *Acacia* on the whole appears more advanced (Ross, 1981). The cpDNA data suggests that subgenus *Aculeiferum* was quite well differentiated at the specific level when Gondwanaland fragmented. This can be seen in the cladogram (figure 4.16) where accessions of subgenus *Aculeiferum* do not appear to be grouped geographically. This suggests that subgenus *Aculeiferum* could possibly have existed before the fragmentation of Gondwanaland. Subgenus *Acacia* appears to be divided into two lineages, one in Africa and one in the New World. This may suggest that it is unlikely that subgenus *Acacia* was very differentiated at the time of fragmentation.

Pedley (1986) suggested that the genus *Acacia* had originated at about the time of fragmentation. He based this conclusion on the present day distributions of taxa and the observation that taxa within the genus were not adapted for long distance dispersal. The results presented here with their tentative conclusions provide some support for Pedley's hypothesis, i.e. the genus *Acacia* existed before or during the fragmentation of Gondwanaland. Further investigation is however, required to confirm the speculative suggestions of this present investigation. The question as to which of the subgenera is the ancestral form of *Acacia*, and where the genus originated will also have to await further investigation.

7.2 Is *Acacia laeta* a hybrid?

This question was only partly answered. *A. laeta* was suggested to be a hybrid initially by Aubreville (1950), but the first investigation was that of El Amin (1976). The morphological investigation of El Amin concluded that *A. laeta* was a hybrid between *A. mellifera* and *A. senegal*. This molecular investigation reaffirms the suggestion that *A. laeta* is a hybrid between *A. mellifera* and *A. senegal*, but the possibility that species other than *A. senegal* and *A. mellifera* could be involved still exists.

Two types of markers were used to address the question, cpDNA restriction site markers and ribosomal nuclear DNA (rDNA) restriction site markers. The cpDNA phenotype of the hybrid *A. laeta* surveyed was exactly the same as *A. mellifera*. This suggests that *A. mellifera* is always the maternal parent. This would account for El Amin's (1976) observation that the close affinity of *A. laeta* with *A. mellifera* was due to backcrossing. If mating was taking place at random then one would expect that in an area where one parent was overrepresented, the majority of hybrids between the two would have the underrepresented parent as the maternal parent, in this case *A. senegal*. However, *A. mellifera* is always the maternal parent, in spite of a lack of *A. senegal* specimens in the areas surveyed. This observation suggests that 'pollen load' is not an important factor in determining the maternal and paternal parents of *A. laeta*, but that another unknown factor determines the direction of the cross.

The rDNA evidence gives some support to the hypothesis that *A. mellifera* and *A. senegal* are the parents of *A. laeta*. Several of the hybrid accessions examined have rDNA phenotypes showing additive inheritance of a marker from both putative parents. This suggests that *A. laeta* is a hybrid between the two mentioned parents, but this evidence must be qualified by the knowledge that only one probe/restriction enzyme combination (PEC) gave an additive phenotype. In the other four rDNA PECs that were examined

the phenotype of *A. laeta* was similar to *A. mellifera*. This result, i.e. non-additive hybrid phenotypes in a hybrid, has been found in other investigations of hybrids. For example, Hughes and Harris (1994) reported the presence of non-additive hybrid phenotypes in hybrids between *Leucaena leucocephala* subsp. *glabrata* and *L. esculenta* subsp. *esculenta*. Of the restriction enzymes they used three gave additive phenotypes and five gave non-additive phenotypes for some of the accessions. The precise mechanism behind these results is not known at present. Variable methylation of restriction sites in nuclear DNA and inadequate sampling of genetic variation in the parents have both been suggested as possible causes for the appearance of non-additive phenotypes. Both suggestions plausibly explain the results of this investigation and further studies based on the present results are required.

7.3 RAPD analysis of *A. tortilis*.

The results of this investigation were interesting in two respects. First the relationships suggested by the RAPD data were novel and to some extent unexpected. Secondly the study proved that RAPD techniques are able to give meaningful results and define further issues to investigate.

A. tortilis is a widespread species in Africa, it has many uses which benefit rural communities. It is a very variable species with 4 subspecies, subsp. *tortilis*, subsp. *raddiana*, subsp. *spirocarpa* and subsp. *heteracantha*. The RAPD generated data separated the accessions studied into two distinct groups. These groupings correlate with the geographic position of the accessions, their taxonomic position being of secondary importance. All the north African accessions group together, and all the south African accessions group together. The dividing line appears to follow a phytogeographical boundary suggested by Wickens (1976). The basis for this divide is at present unknown, but this study is not unique in highlighting a boundary between north and south Africa. S.Harris (pers. com., 1995) has also been investigating the genetic

diversity of *A. tortilis* using the enzyme peroxidase in an isozyme study. His results also point to a divide between north and south African populations of *A. tortilis.* In addition, a study of *Faidherbia albida* using the enzyme peroxidase and additional isozymes also highlights a genetic divide between the northern and southern accessions of *F. albida* (S.Harris, pers. com.). These two studies further reinforce the existence of a boundary between north and south Africa that appears to limit gene flow between these two areas. The exact nature of the boundary is not known at present and requires further investigation.

In addition to the highlighted boundary between north and south Africa the RAPD data also suggest that hybridisation and/or introgression may possibly be occurring between the subspecies of *A. tortilis.* At present however, this cannot be confirmed or dismissed. The presence of an accession of subsp. *spirocarpa* in a cluster of subsp. *tortilis* (figure 6.6) and the clustering pattern of subsp. *spirocarpa* and subsp. *heteracantha* in southern Africa need further explanation. All four subspecies are interfertile so it is not unexpected that hybridisation can occur between neighbouring populations of different subspecies. Further investigations to discover whether hybridisation and/or introgression are actually occurring, and if so the extent of them may enable us to further understand the population genetics of *A. tortilis* over the continent of Africa as a whole.

The results of this investigation have proved the utility of RAPD techniques in investigating plant evolution and diversity. This study was relatively cheap, and quick, and succeeded in highlighting areas for further investigation. This is the area in with RAPD techniques may prove most beneficial, i.e. in initial or preliminary studies to define the course of further investigations. The extension of RAPD techniques into in-depth investigations may not be so useful. Apart from the cost factor, which rises exponentially as the number of accessions increases, there are many questions concerning the use of such markers for phylogenetic purposes. Initially there is the problem

of repeatability and optimisation of the PCR reaction. Optimisation of the reaction can add greatly to the costs of the investigation, and a lack of repeatability may compromise the final conclusions of the experiment. The evolution, heritability and transmission of RAPD markers is poorly understood at present, and this should discourage the use of RAPD techniques to investigate problems above the species level. The problems associated with ensuring that only homologous bands are used in the analysis enhance the above problems. These problems appear to be correlated with the scale of the survey or investigation, i.e. in small or initial studies many of the problems need not be addressed as the conclusions of these studies are only preliminary, and define further investigations to be carried out using more rigorous techniques.

7.4 Concluding remarks.

The above investigations have shown that molecular characters have provided novel insights into relationships between taxa at all taxonomic levels, from the relationships of genera in the Acacieae and Ingeae to relationships between the subspecies of *A. tortilis*.

The results reinforce some of the caveats concerning the use of molecular characters, i.e. the importance of not overstating the results using such characters, 'Traditional' characters such as morphology have just as important a role to play in elucidating taxonomic and biosystematic problems, and the interpretation of molecular results must be tempered by a knowledge of their limitations.

Bibliography

ADAMS, M.E. 1967: A study of *Acacia mellifera*, *A. seyal* and *Balanites aegyptiaca* in relation to land clearing. - Journal of Applied Ecology **4**:221-237.

ADAMS, R.P. and DEMEKE, T. 1993: Systematic relationships in *Juniperus* based on random amplified polymorphic DNAs (RAPDs). - Taxon **42**:553-571

ALBERT, V.A, MISHLER, B.D, and CHASE, M.W. 1992: Character state weighting for restriction site data in phylogenetic reconstruction, with an example from chloroplast DNA. - In SOLTIS, P.S., SOLTIS, D.E. and DOYLE, J.J. (eds.): Molecular systematics of plants. - New York: Chapman and Hall.

ANDREWS, E.C. 1914: Development and distribution of the natural order Leguminosae. - Journal of the Proceedings of the Royal Society of New South Wales. **48**:33-407.

ARCE, M.L.R., 1989: Systematic study of the generic patterns in the tribe Ingeae (Leguminosae: Mimosoideae), with emphasis on *Zygia-Caulanthon*. - University of Southampton: PhD. Thesis.

ATCHISON, E. 1948: Studies in the Leguminosae II: cytogeography of *Acacia* (TOURN.) L. - American Journal of Botany **35**:651-655.

AUBREVILLE, A., 1950: *Acacia laeta*. - In Flore forestiere Soudano-Guinéenne - Société d'Editions Géographiques, Paris.

AVISE, J.C., 1994: Molecular Markers, Natural History and Evolution. - New York: Chapman and Hall.

BAUM B.R. and BAILEY, L.G. 1989: Species relationships in the *Hordeum murinum* aggregate viewed from chloroplast DNA restriction patterns. - Theoretical and Applied Genetics 78:311-317

BENTHAM, G., 1840: Contributions towards a Flora of South America - Enumeration of Plants collected by Mr. Schomburgk in British Guiana. - Journal of Botany. (Hooker) **2**:127-146.

— 1841: Observations on the distinctive characters of the Papilionaceae and Caesalpinieae, Suborders of Leguminosae. - Journal of Botany. (Hooker) **3**:125-133.

— 1842a: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **1**:318-392.

— 1842b: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **1**:494-528.

— 1844a: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **3**:82-112.

— 1844b: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **3**:195-226.

— 1845: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **4**:577-622.

— 1846: Notes on Mimoseae, with a synopsis of species. - London Journal of Botany **5**:75-108.

— 1865: Leguminosae. - In BENTHAM, G. and HOOKER, J.D.: Genera Plantarum. - London: Lovell Reeve and Co. and Williams and Norgate.

— 1875: Revision of the Suborder Mimoseae. - Transactions of the Linnean Society London **30**:335-664.

BIRKY, C.W. 1978: Transmission genetics of mitochondria and chloroplasts. - Annual Review of Genetics **12**:471-512.

BRAIN, P., 1987: Immunology and phylogeny: a preliminary study of *Acacia*. - South African Journal of Science **83**:422-427.

— 1990: Immunology and phylogeny II: further studies on *Acacia*. - Suid Afrikaanse Ttoskifvir Weinnskap **86**:195-199.
BREHM, B. AND OWNBEY, M. 1965: Variation in the chromatographic patterns in the *Tragopogon dubius-pratensis-porrifolius* complex (Compositae). - American Journal of Botany **52**:811-818.
BREMER, B. 1991: Restriction data from chloroplast DNA: is there only one accurate way of scoring?. - Plant Systematics and Evolution **175**:39-54.
— and JANSEN, R.K. 1991: Comparative restriction site mapping of chloroplast DNA implies new phylogenetic relationships within *Rubiaceae*. - American Journal of Botany 78:198-213.
BRENAN, J.P.M., 1959: Leguminosae. Mimosoideae. - In HUBBARD, C.E. and MILNE-REDHEAD, E., (Eds.): Flora of Tropical East Africa. - London: Crown Agents for Overseas Governments and Administration.
— 1978: Some aspects of the phytogeography of tropical Africa. - Annals of the Missouri Botanic Garden **65**:437-478.
— 1983: Manual on the taxonomy of *Acacia* species: present taxonomy of four species of *Acacia* (*A.albida, A.senegal, A.nilotica, A.tortilis*). - Rome: FAO.
BRITTON, N.L. and ROSE, J.N., 1928: North American Flora. Vol. 23. - New York: New York Botanical Garden.
BROWN, W.L., 1960: Ants, Acacias and browsing animals. - Ecology **41**:587-592
BURKART, A., 1952: Las Leguminosas Argentinas Silvestres y Cultivades. - ed.**2**. - Buenos Aires.
CAMIN, J.H. and SOKAL R.R. 1965: A method for deducing branching sequences in phylogeny. - Evolution **19**:311-326.
CAMPBELL, M.K 1991: Biochemistry. - Saunders College Publishing, Fort Worth.
CAPUTO, P., STEVENSON, D.W. and WURTZEL, E.T. 1991: A phylogenetic analysis of American Zamiaceae (Cycadales) using chloroplast DNA restriction fragment length polymorphisms. - Brittonia **73**:135-145.
CHALMERS, K.J., WAUGH, R., SPRENT, J.I., SIMONS, A.J. and POWELL, W. 1992: Detection of genetic variation between and within populations of *Gliricidia sepium* and *G.maculata* using RAPD markers. - Heredity **69**:465-472.
CHAPPILL, J.A. and MASLIN, B.R. 1995: A phylogenetic assessment of Tribe Acacieae. - In CRISP, M. and DOYLE, J.J. (eds.): Advances in Legume Systematics 7: Phylogeny. - Royal Botanic Gardens Kew.
CHASE, M.W. and HILLS, H.H., 1991: Silica gel: An ideal material for field preservation of leaf samples for DNA studies. - Taxon **40**:215-220.
CHERFAS, J. 1990: Genes unlimited. - New Scientist **14 April**:29-33.
CHEVALIER, M-H., 1993: Genetic diversity in the genus *Acacia*: phylogenetic relationships among species. - Paper presented at the Fourth Congress of the European Society for Evolutionary Biology: Montpellier, France.
CLARKE, H.D., SIEGLER, D.S. and EBINGER, J.E. 1990: *Acacia constricta* (Fabaceae: Mimosoideae) and related species from south-western U.S. and Mexico. - American Journal of Botany **77**:305-315.
CLEGG, M.T. and ZURAWSKI, G. 1992: Chloroplast DNA and the study of plant phylogeny: present status and future prospects. - In SOLTIS, P.S., SOLTIS, D.E. and DOYLE, J.J. (eds.): Molecular systematics of plants. - New York: Chapman and Hall.
CONN, E.E., SEIGLER, D.S., MASLIN, B.R. and DUNN, J., 1989: Cyanogenesis in *Acacia* subgenus *Aculeiferum*. - Phytochemistry **28**:817-820.
CRAWFORD, D.J., BRAUNER, S., COSNER, M.B and STUESSY, T.F. 1993: Use of RAPD markers to document the origin of the intergeneric hybrid *xMargyracaena*

skottsbergii (Rosaceae) on the Juan Fernandez Islands. - American Journal of Botany **80**:89-92.
DARLINGTON, C.D. and WYLIE, A.P., 1955: Chromosome Atlas of Flowering Plants. - London: George Allen and Unwin.
D'OVIDIO, R., TANZERELLA, O.A. and PORCEDDU, E. 1990: Rapid and efficient detection of genetic polymorphism in wheat through amplification by polymerase chain reaction. - Plant Molecular Biology **15**:169-171.
DOWLING, T.E., MORITZ, C. and PALMER, J.D. 1990: Chapter 8 Nucleic Acids II, restriction site analysis. - In HILLIS, D.M. and MORITZ, C. (eds.): Molecular Systematics. - Mass.: Sinauer.
DOYLE, J.J., DOYLE, J.L. and BROWN, A.H.D. 1990: Chloroplast DNA polymorphism and phylogeny in the B genome of Glycine subgenus Glycine (Leguminosae). - American Journal of Botany **77**:772-782.
DOYLE, J.J., LAVIN, M. and BRUNEAU, A. 1992: Contributions of molecular data to Papilionoid legume systematics. - In SOLTIS, P.S., SOLTIS, D.E. and DOYLE, J.J., (Eds.): Molecular systematics of plants, pp. 151-176. - London: Chapman and Hall.
EDWARDS, K., JOHNSTONE, C. and THOMPSON, C. 1991: A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. - Nucleic Acids Research **19**:1349.
EHRENDORFER, F., 1954: Phylogeny and evolutionary mechanisms in *Lepto-Galium*. - 8e Congr. Inernat. Botany. Paris.
EL AMIN, H.M., 1976: *Acacia laeta* R.BR. *ex* BENTH., considered as a species of hybrid origin. - Sudan Silva 21: 14-23.
ERLICH, H.A. (ed.) 1989: PCR technology: principles and applications for DNA amplification. - New York: Stockton.
FABIJANSKI, S., FEDAK, G., ARMSTRONG, K. and ALTOSAAR, I., 1990: A repeated sequence probe for the C genome in *Avena* (oats). - Theoretical and Applied Genetics **79**:1-7.
FAGG, C. and BARNES, R.D. 1990: African Acacias: Study and acquisition of the genetic resources. - Final Report ODA research scheme R.4348, Oxford Forestry Institute, University of Oxford, Oxford.
FARRIS, J.S. 1970: Methods for computing Wagner trees. - Systematic Zoology **34**:21-34
— 1972: Estimating phylogenetic trees from distance matrices. - American Naturalist **106**:645-668.
— 1977: Phylogenetic analysis under Dollo's law. - Systematic Zoology **26**:77-88.
— 1989: The retention index and the rescaled consistency index. - Cladistics **5**:417-419.
FELSENSTEIN, J. 1983: Parsimony in systematics: biological and statistical issues. - Annual Review of Ecology and Systematics **14**:313-333.
— 1985: Confidence limits on phylogenies: an approach using the bootstrap. - Evolution **39**:783-791.
— 1992: Phylogenies from restriction sites: a maximum likelihood approach. - Evolution **46**:159-173.
— 1993: PHYLIP (Phylogeny Inference Package) Version 3.5c. - Dept. of Genetics, University of Washington, Seattle.
FINK, W.L. 1986: Microcomputers and phylogenetic analysis. - Science **234**:1135-1139.
FITCH, W.M. 1971: Toward defining the course of evolution: minimal change for a specific tree topology. - Systematic Zoology **20**:406-416.
— AND MARGOLIASH 1967: Construction of phylogenetic trees. - Science **155**:279-284.

FLAVELL, R.B., O'DELL, M., THOMPSON, W.F., VINCENTZ, M., SARDANA, R. and BARKER, R.F., 1986: The differential expression of ribosomal DNA genes. - Philosophical Transactions of the Royal Society London B **314**: 385-397.
FOWLER, H.W. and FOWLER, F.G. (ed.s) 1964: The consise Oxford dictionary of current English, Fifth Edition. - Oxford, Oxford University Press.
GEESINK, R. 1981: Tephrosieae. - In POLHILL, R.M. and RAVEN, P.H. (eds.) : Advances in Legume systematics Part 1. - Royal Botanic Gardens, Kew.
GEFLAND, D. 1989: *Taq* DNA polymerase. - In ERLICH, H.A. (ed.) : PCR technology: principles and applications for DNA amplification. - New York: Stockton.
GERLACH, W.L. and BEDBROOK, J.R. 1979: Cloning and characterisation of ribosomal RNA genes from wheat and barley. - Nucleic Acids Research **7**:1869-1885.
GILLIES, A.C.M. 1994: Molecular analysis of genetic diversity and evolutionary relationships in the tropical legume genus *Stylosanthes* (AUBL.) SW. - University of St. Andrews: PhD. thesis.
GUINET, P., 1967: Parallélisme entre les caractéres du pollen dans les groupes Piptadeinées et Acaciées des Mimosacées. - Review of Paleobotany and Palynology **3**:151-153.
— 1969: Les Mimosacées, étude de palynologie fondamentale, corrélations, évolution. - Trav. Sec. Sci. Techn. Pondichery. **9**:1-293.
— 1990: The genus *Acacia* (Leguminosae, Mimosoideae): its affinities as borne out by its pollen characters. - Plant Systematics and Evolution Suppl. 5: 81-90.
— and VASSAL, J., 1978: Hypotheses on the differentiation of the major groups of the genus *Acacia* (Leguminosae). - Kew Bull. **32**:509-527.
HAGOS, M., SAMUELSON, G., KENNE, L. and MODAWI, B.W. 1987: Isolation of smooth muscle relaxing 1,3,diaryl-propan-2-ol derivatives from *Acacia tortilis*. - Planta Médica **53**(1):27-31.
HAMRICK, J.L., and GODT, M.J.W. 1990: Allozyme diversity in plant species. - In BROWN, A.H.D, CLEGG, M.T. KAHLER, A.L. and WEIR, B.S. (eds.) : Plant population genetics, breeding and genetic resources. - Sinauer Associates Inc.
HARRIS, S.A. 1990: Molecular Systematics in some members of the genus *Senecio* L. (Asteraceae). - University of St. Andrews PhD. thesis.
— and INGRAM, R., 1991: Chloroplast DNA and biosystematics: The effects of intraspecific diversity and plastid transmission. - Taxon **40**:393-412.
— and —, 1992; Molecular systematics of the genus *Senecio* L.: hybridisation in a British polyploid complex. - Heredity **69**: 1-10.
—, — and ABBOTT, R.J. 1993: A molecular study of genetic resources in the genus *Leucaena* BENTH. - Final report, ODA research scheme R.4727
— and ROBINSON, J. 1994: Preservation of tropical plant material for molecular analyses. - Proceedings of the Second International DNA-Bank Net Conference, St.Louis.
HEISER, C.B., 1949: Study in the evolution of the Sunflower species *Helianthus annuus* and *H.bolanderi*. - University of California Publications in Botany **23**:157-196.
HILLIS, D.M., HUELSENBECK, J.P. and CUNNINGHAM, C.W. 1994: Application and accuracy of molecular phylogenies. - Science **264**:671-677.
—, LARSON, A., DAVIS, S.K. and ZIMMER E.A. 1990: Chapter 9 Nucleic Acids III: Sequencing. - In HILLIS, D.M. and MORITZ, C. (eds.): Molecular Systematics. - Mass.: Sinauer.
— and MORITZ, C. 1990: Chapter 12, An overview of the applications of molecular systematics. - In HILLIS, D.M. and MORITZ, C. (eds.): Molecular Systematics. - Mass.: Sinauer.

HIRATSUKA, J., SHIMADA, H., WHITTIER, R., ISHIBASHI, T., SAKOMOTO, M., MORI, M., KONDO, C., HONJI, Y., SUN, C.-R., MENG, B.-Y., LI, Y.-Q, KANO, A., NISHIZAWA, Y., HIRAI, A., SHINOZAKI, K. and SUGIURA, M. 1989: The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. - Molecular and General Genetics. **217**:185-194.

HOSAKA, K. and HANNEMAN, R.E. 1988: The origin of the cultivated tetraploid potato based on chloroplast DNA. - Theoretical and Applied Genetics **76**:333-340

— , OGIHARA, Y., MATSUBAYASHI, M. and TSUNEWAKI, K., 1984: Phylogenetic relationship between tuberous *Solanum* species as revealed by restriction endonuclease analysis of chloroplast DNA. - Japanese Journal of Genetics **59**:349-369.

HUFF, D.R., PEAKALL, R. and SMOUSE, P.E. 1993: RAPD variation within and among natural populations of outcrossing buffalograss (*Buchloë dactyloides* (NUTT.) ENGELM.). - Theoretical and Applied Genetics **86**: 927-934

HUGHES, C.E. 1987: Biological considerations in designing a seed collection strategy for *Gliricidia sepium* (JACQ.) WALP. (Leguminosae). - Commonwealth Forestry Review **66**(1):31-48.

— and HARRIS, S.A., 1994: The characterisation and identification of a naturally occurring hybrid in the genus *Leucaena* (Leguminosae: Mimosoideae). - Plant Systematics and Evolution **192**:177-197

INNIS, M.A., GEFLAND, D.H., SNINSKY, J.J. and WHITE, T.J. (eds.) 1990: PCR protocols : a guide to methods and applications. - San Diego: Academic Press.

JACKSON, J.K. and PEAKE, 1955: Forestry Research in the Sudan. - Memoirs of the Forest Division No.**7**.

JANZEN, D.H. 1966: Coevolution of mutualism between ants and acacias in Central America. - Evolution **20**:249-275.

— 1974: Swollen thorn Acacias of central America. - Smithsonian Contrib. Botany. **13**:1-131.

KANIS, A., 1986: Taxonomic changes in Australian Mimosaceae. - Australian Systematic Botany Society Newsletter **48**:9-13

KAYONGO MALE, H AND FIELD, C.R. 1983: Feed quality and utilisation by cattle grazing natural pasture in the range areas of northern Kenya. - In LUSINGI, W. (ed.) - IPAL Report **A5**:230-245.

KAZAN, K., MANNERS, J.M. and CAMERON, D.F. 1993: Genetic variation in agronomically important species of *Stylosanthes* determined using random amplified polymorphic DNA markers. - Theoretical and Applied Genetics **85**:882-888.

KELLOGG, E. and CAMPBELL, C. 1987: Phylogenetic analysis of the Gramineae. - In SODERSTROM, T., HILU, K.W., CAMPBELL, C.S and BARKWORTH, M.E. (eds.): Grass systematics and evolution. - Smithsonian Institution Press, Washington.

KIMURA, M., 1982: Molecular Evolution, Protein Polymorphism and Evolutionary Theory. - Tokyo: Japanese Scientific Societies Press.

KING, L.M. and SCHAAL, B.A., 1990: Genotypic variation within asexual lineages of *Taraxacum officinale*. - Proceedings of the National Academy of Sciences U.S.A. **87**:998-1002.

KLUGE, A.G. and FARRIS, J.S. 1969: Quantitative phyletics and the evolution of anurans. - Systematic Zoology **18**:1-32

KRON, K.A., GAWEN, L.M. and CHASE, M.W., 1993: Evidence for introgression in Azaleas (*Rhododendron*; Ericaceae): chloroplast DNA and morphological variation in a hybrid swarm on Stone Mountain, Georgia. - American Journal of Botany **80**:1095-1099.

LAVIN, M. 1987: A cladistic analysis of the tribe Robinieae (Papilionoideae; Leguminosae). - In STIRTON, C.H. (ed.): Advances in Legume Systematics, Part 3. - Royal Botanic Gardens, Kew.
— and DOYLE, J.J. 1991: Tribal relationships of *Sphinctospermum* (Leguminosae): integration of traditional and chloroplast DNA data. - Systematic Botany **16**:162-172.
LEE, Y.S., SIEGLER, D.S. and EBINGER, J.E. 1989: *Acacia rigidula* (Fabaceae) and related species in Mexico and Texas. - Systematic Botany **14**:91-100.
LISTON, A. 1992: Variation in the chloroplast genes *rpo*C1 and *rpo*C2 of the genus *Astragalus* (Fabaceae): evidence from the restriction site mapping of a PCR amplified fragment. - American Journal of Botany **79**:953-961.
LUMARET, R., BOWMAN, C.M. and DYER, T.A. 1989: Autopolyploidy in *Dactylis glomerata* L.: further evidence from studies of chloroplast DNA variation. - Theoretical and Applied Genetics **78**:393-399.
MASLIN B.R., 1987: Report of discussions at I.G.S.M. Meeting No. 6 on the proposal to split *Acacia*. - Bulletin of the International Group for the Study of Mimosoideae **15**:108-118.
— 1988: Should *Acacia* be divided? - Bulletin of the International Group for the Study of Mimosoideae **16**:54-76.
— and STIRTON, C.H., in press: Generic and infra-generic classification in *Acacia* (Leguminosae: Mimosoideae): a list of critical species on which to build a comparative data set. - Bulletin of the International Group for the Study of Mimosoideae.
MILLER, P., 1754: Gardeners Dictionary. - abridg. ed. 4.
MOFFET, A.A., 1965: Genetic Studies in acacias III. Chlorosis in interspecific hybrids. - Heredity **20**:609-620.
MORETTI, A., CAPUTO, P., COZZOLINO, S., DE LUCA, P., GAUDIO., L., SINISCALCO-GIGLIANO, G. and STEVENSON, D.W. 1993: A phylogenetic analysis of *Dioon* (Zamiaceae). - American Journal of Botany **80**:204-214
MOORE, D.M., 1968: The karyotype in Taxonomy. - In HEYWOOD, V.H. (ed.): Modern Methods in Plant Taxonomy. - London : Edward Arnold.
NEI, M. 1987: Molecular Evolutionary Genetics. - Columbia University Press, New York.
NEILSEN, I., 1981: Ingeae. - In POLHILL, R.M. and RAVEN, P.H., (eds.): Advances in Legume Systematics part **1**: 173-190. - London: Royal Botanic Gardens, Kew.
NEW, T.R. 1984: A biology of Acacias. - Oxford University Press.
NEWMAN, I.V. 1932: Studies in the Australian acacias. - Journal of the Linnean Society of Botany **49**:133-171
OHYAMA, K., FUKUZAWA, H., KOHCHI, T., SHIRAI, H., SANO, S., UMESONO, K., SHIKI, Y., TAKEUCHI, M., CHANG, Z., AOTA, S., INOKUCHI, H. and OZEKI, H. 1986: Chloroplast gene organisation deduced from the complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. - Nature **322**:572-574.
OLNG'OTIE, P.A.S 1991: *Acacia tortilis* (FORSK.) HAYNE; A study of genetic structure and breeding systems. - University of Oxford: D. Phil Thesis.
OWNBEY, M. 1950: Natural hybridisation and amphiploidy in the genus *Tragopogon*. - American Journal of Botany **37**:487-499.
OWNBEY, M. and MCCOLLUM, G. 1953: Cytoplasmic inheritance and reciprocal amphiploidy in *Tragopogon*. - American Journal of Botany **40**:788-796.
OWNBEY, M. and MCCOLLUM, G. 1954: The chromosomes of *Tragopogon*. - American Journal of Botany **56**:7-21.

PALMER, J.D. 1985: Evolution of chloroplast and mitochondrial DNA in plants and algae. -In MACINTYRE, R.J. (Ed.): Monographs in evolutionary biology; molecular evolutionary genetics. -New York: Plenum Press.
— 1987: Chloroplast DNA evolution and the biosystematic uses of chloroplast DNA variation. - American Naturalist **130**:456-529.
— , JANSEN, R.K., MICHAELS, H.J., CHASE, M.W. and MANHART, J.R. 1988: Chloroplast DNA variation and plant phylogeny. - Annals of the Missouri Botanical Garden **75**:1180-1206.
— and THOMPSON, W.F. 1981: Rearrangements in the chloroplast genomes of mung bean and pea. - Proceedings of the National Academy of Sciences U.S.A. **78**:5533-5537.
PEDLEY, L. 1978: A revision of *Acacia* MILL. in Queensland. - Austrobaileya **1**:75-234.
— 1981: Classification of acacias. - Bulletin of the International Group for the Study of Mimosoideae **9**:42-48.
— 1986: Derivation and dispersal of *Acacia* (Leguminosae), with particular reference to Australia, and the recognition of *Senegalia* and *Racosperma*. - Botanical Journal of the Linnean Society **92**:219-254.
— 1987: In defence of *Racosperma*. - Bulletin of the International Group for the Study of Mimosoideae **15**:123-129.
PETTIGREW, C.J. and WATSON, L. 1975: On the classification of Australian Acacias. - Australian Journal of Botany **23**:833-847.
POLHILL, R.M. and SOUSA, M. 1981: Robinieae. - In POLHILL, R.M. and RAVEN, P.H. (eds.) : Advances in Legume systematics Part 1. - Royal Botanic Gardens, Kew.
RAFALSKI, J.A. and TINGEY S.V. 1993: Genetic diagnostics in plant breeding: RAPDS, microsatellites and machines. - Trends in Genetics **9**(8):275-279.
RAVEN, P.H. and AXELROD, D.L. 1974: Anigiosperm biogeography and past continental movements. - Annals of the Missouri Botanical Garden **61**:539-673.
RIESEBERG, L.H. , BECKSTROM-STERNBERG, S. and DOAN, K. 1990: *Helianthus annuus* ssp. *texanus* has chloroplast DNA and nuclear ribosomal genes of *Helianthus debilis* ssp. *cucumerifolius*. - Proceedings of the National Academy of Sciences of the USA **87**:593-597.
— and BRUNSFELD, S.J., 1992: Molecular evidence and plant introgression. - In SOLTIS, P.S., SOLTIS, D.E. and DOYLE, J.J., (Eds.): Molecular systematics of plants, pp. 151-176. - London: Chapman and Hall.
— and SOLTIS, D.E., 1991: Phylogenetic consequences of cytoplasmic gene flow in plants. - Evolutionary Trends in Plants **5**(1);65-84.
—, SOLTIS, D.E. and PALMER, J.D., 1988: A molecular re-examination of introgression between *Helianthus annuus* and *H.bolanderi* (Compositae). - Evolution **42**:227-238
ROBBERTSE, P.J. 1974: The genus *Acacia* in South Africa. II. With special reference to the morphology of the flower and inflorescence. - Phytomorphology **24**:1-15.
ROOSE, M.L. AND GOTTLIEB, L.D. 1976: Genetic and biochemical consequences of polyploidy in *Tragopogon*. - Evolution **30**:818-830.
ROSS, J.H., 1973: Towards a classification of the African Acacias. - Bothalia **11**:107-113
— 1979: A conspectus of the African *Acacia* species. - Memoirs of the Botanical Survey of S. Africa **44**:1-155.
— 1981: An analysis of the African *Acacia* species: their distribution, possible origins and relationships. - Bothalia **13**:389-413.

RYDBERG, P.A. 1924: Genera of North American Fabaceae II. - American Journal of Botany **11**:470-482.
SAITOU, N. and NEI, M. 1987: The neighbour joining method: a new method for reconstructing phylogenetic trees. - Molecular and Biological Evolution **4**:406-425.
SALE, G.N. 1948: Note on sand dune fixation in Palestine. - Empire Forestry Review **27**:60-61.
SAMBROOK, J., FRITSCH, E.F. and MANIATIS, T. 1989: Molecular Cloning: A laboratory manual, 2nd edition. - Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.
SANDERSON, M.J. and DONOGHUE, M.J. 1989: Patterns of variation in levels of homoplasy. - Evolution **43**:1781-1795.
SHINOZAKI, K., OHME, M., TANAKE, M., WAKUSUGI, T., HAYASHIDA, N., MATSUBAYASHI, T., ZAITA, N., CHUNWONGSE, J., OBOKATA, J., YAMAGUCHI-SHINOZAKI, K., OHTON, C., TORAZAWA, K.,MENG, B.U., SUGITA, M., DENO, H., KAMOGASHIRA, T., YAMADA, K., KUSUDA, J., TAKAIWA, F., KATO, A., TOHDOH, N., SHIMADA, H. AND SUGUIRA, M. 1986: The complete nucleotide sequence of the tobacco chloroplast genome. - Plant Molecular Biology Reports **5**:2043-2049.
SHREVE, F. 1951: Vegetation of the Sonoran Desert. - Publication no. 591, Carnegie Inst. Washington.
SIMMONS, A.J. 1992: Genetic Improvement of non-industrial trees. - Agroforestry Systematics **18**:197-212.
SMITH, J. 1949: Distribution of tree species in Sudan in relation to rainfall and soil texture. - Forestry Bulletin of the Agricultural Publication Committee no. 4.
SNEATH, P.H.A. and SOKAL, R.R. 1973: Numerical Taxonomy. - San Francisco: W.H.Freeman and Co.
SOLTIS, D.E. and SOLTIS, P.S. 1989: Allopolyploid speciation in *Tragopogon*: Insights from chloroplast DNA. - American Journal of Botany **76**:1119-1124.
—, — and MILLIGAN, B.G. 1992: Intraspecific chloroplast variation: systematic and phylogenetic implications. - In SOLTIS, P.S., SOLTIS, D.E. and DOYLE, J.J., (Eds.): Molecular systematics of plants, pp. 151-176. - London: Chapman and Hall.
—, — and NESS, B.D. 1989a :Chloroplast DNA variation and multiple origins of autopolyploidy in *Heuchera micracantha*. - Evolution 43;650-656.
—, —, RANKER ,T.A. and NESS, B.D. 1989b: Chloroplast DNA variation in a wild plant, *Tolmiea menziesii*. - Genetics **121**:819-826.
SOUSA, M. and DE SOUSA, M.P. 1981: New World Lonchocarpinae. - In POLHILL, R.M. and RAVEN, P.H. (eds.) : Advances in Legume systematics Part 1. - Royal Botanic Gardens, Kew.
SOUTHERN, E.M. 1975: Detection of specific sequences and DNA fragments separated by gel electrophoresis. - Journal of Molecular Biology **98**:503-517.
SPOONER, D.M., SYTSMA, K.J. and SMITH, J.F., 1991: A molecular re-examination of diploid hybrid speciation of *Solanum raphanifolium*. - Evolution **45**:757-764.
STAMMERS, M. HARRIS, J., EVANS, G.M., HAYWARD, M.D. and FORSTER, J.W. 1995: Use of random PCR (RAPD) technology to analyse phylogenetic relationships in the *Lolium/Festuca* complex. - Heredity **74**:19-27.
STEBBINS, G.L., 1971: Chromosomal evolution in higher plants. - Edward Arnold, London.
— 1974: Flowering plants. Evolution above the species level. - Edward Arnold, London.
STRYER, L. 1988: Biochemistry 3rd. ed. - W.H. Freeman and Co., New York.

SWOFFORD, D.L. 1993: PAUP - Phylogenetic analysis using parsimony , vers. 3.1.1. - Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.
— and BEGLE D.P. 1993: Users' Manual for PAUP ver. 3.1 - Smithsonian Institution.
— and OLSEN, G.J. 1990: Phylogeny Reconstruction. - In HILLIS, D.M. and MORITZ, C. (eds.): Molecular Systematics. - Mass. : Sinauer.
SYTSMA, K.J, and GOTTLIEB, L.D. 1986: Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). - Evolution **40**:1248-1261.
TILNEY-BASSET, R.A.E. 1978: The inheritance and genetic behaviour of plastids. - In KIRK, J.T.O and TILNEY-BASSET, R.A.E., (Eds.): The Plastids. - Amsterdam: Elsevier.
TINDALE, M.D. and ROUX, D.G. 1975: Phytochemical studies on the heartwoods and barks of African and Australian species of *Acacia*. - Boissera **24a**:299-305.
UGENT, D., 1970: *Solanum raphanifolium*, a Peruvian wild potato species of hybrid origin. - Botanical Gazette (Crawfordsville) **131**:225-233.
VASSAL, J., 1972: Apport des recherches ontogéniques et seminologiques à l'étude morphologique, taxonomique et phylogénique du genre *Acacia*. - Bull. Soc. Hist. Nat. Toulouse **108**:105-247.
VAN DER JOOSTE, W. 1965: Experiments on the control of *Acacia mearnsii*. - South African Journal of Agricultural Science **8**:1165-1166.
VESEY-FITZGERALD, D.F. 1974: The changing state of *Acacia xanthophloea* groves in Arusha National Park, Tanzania. - Biological Conservation **6**:40-47.
VON MAYDELL, H.J. 1986: Trees and shrubs of the Sahel: their characteristics and uses. - Schriftenreihe der GTZ, no. **196**
WICKENS, G.E. 1976: The flora of Jebel Marra (Sudan Republic) and its geographical affinities. - Kew Bulletin Add. ser. **5**:1-368.
WEINING, S. and LANGRIDGE, P. 1991: Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. - Theoretical and Applied Genetics **82**:209-216.
WELSH, J. and MCCLELLAND, M. 1990: Fingerprinting genomes using PCR with arbitrary primers. - Nucleic Acids Research **18**: 7213-7218.
WENDEL, J.F. and ALBERT, V.A. 1992: Phylogenetics of the cotton genus (*Gossypium* L.): character-state weighted parsimony analysis of chloroplast DNA restriction site data and its systematic and biogeographic implications. - Systematic Botany **17**:115-143.
WERGER, M.J.A. 1977: Effects of game and domestic livestock on vegetation in east and southern Africa. - In KRAUSE, W. (ed.): Handbook of Vegetation Science vol. XIII. - W.Junk : The Hague.
WHITTEMORE, A.T. and SCHAAL, B.A. 1991: Interspecific gene flow in sympatric oaks. - Proceedings of the National Academy of Sciences of the USA **88**:2540-2544.
WILKIE; S.E., ISAAC, P.G. and SLATER, R.J. 1993: Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. - Theoretical and Applied Genetics **86**:497-504.
WILLIAMS, C.B. 1964: Patterns in the balance of nature. - London: Academic Press.
WILLIAMS, J.G.K., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A. and TINGEY, S.V. 1990: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research **18**:6531-6535.
WOLFE, K.H., LI, W.-H. and SHARP, P.M. 1987: Rates of nucleotide substitution vary greatly amongst plant mitochondrial, chloroplast, and nuclear DNAs. - Proceedings of the National Academy of Science, New York **84**:9054-9058.

WOOD, C.E. 1949: The American barbistyled species of *Tephrosia* (Leguminosae). - Contributions of the Gray Herbarium **170**:193-384.

ZIMMER, E.A., JUPE, E.R. and WALBOT, V., 1988: Ribosomal gene structure, variation and inheritance in maize and its ancestors. - Genetics **120**:1125-1136.

ZUCKERKANDL, E. and PAULING, L., 1965: Evolutionary divergence and convergence in proteins. - In BRYSON, V. and VOGEL, H.J. (Eds.): Evolving genes and proteins. - New York Academic Press.

Appendix A:
Material Used In This Study.

Material Used In This Study.

Species	Identification Number.[1]	Type[2]	Origin[3]	Geographic locality[4]
Acacia abyssinica HOCHST. *ex* BENTH.	39/90	Seed	OFI	Zimbabwe
Acacia alata R.BR.	341-83-04103	Dried	Kew living	(Australia)
Acacia amentacea DC.	CEH 1602	Dried	OFI	Mexico
Acacia angustissima (MILL.) KUNTZE	37/88	Seed	OFI	Guatemala
Acacia arenaria SCHINZ	20/90	Seed	OFI	Zimbabwe
Acacia ataxacantha DC.	F622	Dried	C.Fagg	Kenya
Acacia breviscapa HARMS	F624	Dried	C.Fagg	Kenya
Acacia burkei BENTH	D7	Dried	C.Fagg	(HNBG)
Acacia caffra (THUNB.) WILLD.	D8	Dried	C.Fagg	(HNBG)
Acacia caven (MOL.) MOL	455-65-4501	Dried	Kew living	(New World)
Acacia chamelensis L RICO	RIWI 832B	Dried	L.Rico	Mexico
Acacia chariessa MILNE-REDH.	D9	Dried	C.Fagg	(HNBG)
Acacia choriophylla BENTH.	LR 1122	Dried	L.Rico	Cuba
Acacia cucuyo R.BARNEBY & T.A.ZANONI	LR 1156	Dried	L.Rico	Cuba
Acacia deamon EKMAN & URBAN	LR 1080	Dried	L.Rico	Cuba
Acacia dolichocephala HARMS	F634	Dried	C.Fagg	Kenya
Acacia drepanolobium HARMS *ex* SJÖSTEDT	12/92	Seed	OFI	Tanzania
Acacia erioloba E. MAY	23/88	Seed	OFI	Zimbabwe
Acacia exuvialis VERDOON	D4	Dried	C.Fagg	(HNBG)
Acacia farnesiana (L.) WILLD.	9/83	Seed	OFI	Guatemala
Acacia galpinii BURTT DAVY	D1	Dried	C.Fagg	(HNBG)
Acacia gaumeri S.F.BLAKE	LR 1161	Dried	L.Rico	Mexico
Acacia glomerosa BENTH.	LR/SVC 1005B	Dried	L.Rico	Mexico
Acacia greggii A.GREY	161-80-01439	Dried	Kew Living	(New World)
Acacia hebeclada DC subsp.*chobiensis* (O.B. MILLER) SCHREIBER	D2	Dried	C.Fagg	(HNBG)
Acacia karroo HAYNE	45/90	Seed	OFI	Zimbabwe

---

[1] Identification Numbers refers to an identification number for the accession used.An identification number of the form XX/XX e.g. 39/90 relates to an OFI collection. Where there are initials before the number these relate to a herbarium specimen. If the identification number is of the form DX, these accessions were collected by C.Fagg (OFI) from the National Botanic Gardens, Harare, Zimbabwe (HNBG).

[2] Type means how the accession was used or collected. Seed means that the accession was collected as seed, it was then grown in a greenhouse and fresh leaf material was collected as was needed. Dried means that the material was collected as dried material either in the field or from a duplicate herbarium specimen.

[3]Origin refers to the origin of the material, i.e. which organisation or person the material was acquired from.

[4]Geographic locality refers, where applicable, to the country of origin of the material used.

Material Used In This Study.

Acacia karroo HAYNE	21/90	Seed	OFI	Zimbabwe
Acacia koa A.GREY	178-90-01142	Dried	Kew Living	(Hawaii)
Acacia leuderitzii ENGL. var. *leuderitzii*	D10	Dried	C.Fagg	(HNBG)
Acacia macracantha HUMB. & BONPL. *ex* WILLD.	LR 1124	Dried	L.Rico	Cuba
Acacia mammifera SCHLECHT	GONE 1243	Dried	L.Rico	Mexico
Acacia mearnsii DE WILLD.	F638	Dried	C.Fagg	Kenya
Acacia melanoxylon R.BR.		Dried	Kew living	(Australia)
Acacia mellifera (VAHL) BENTH. subsp. *detinens* (BURCH.) BRENAN	D6	Dried	C.Fagg	(HNBG)
Acacia montigena BRENAN & EXELL	F621	Dried	C.Fagg	Kenya
Acacia nigrescens OLIV.	D5	Dried	C.Fagg	(HNBG)
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *kraussiana* (BENTH.) BRENAN	17/90	Seed	OFI	Zimbabwe
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *kraussiana* (BENTH.) BRENAN	31/90	Seed	OFI	Botswana
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *kraussiana* (BENTH.) BRENAN	48/90	Seed	OFI	Zimbabwe
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *leiocarpa* BRENAN	82/90	Seed	OFI	Kenya
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *subalata* (VATKE) BRENAN	81/90	Seed	OFI	Kenya
Acacia nilotica (L.) WILLD *ex* DEL. subsp. *tomentosa* (BENTH.) BRENAN	70/90	Seed	OFI	Cape Verde
Acacia paradoxa SCHLDL.	000-73-18422	Dried	Kew living	(Australia)
Acacia pennatula (SCHLECT. & CHAM.) BENTH.	67/87	Seed	OFI	Honduras
Acacia persiciflora PAX	F 640	Dried	OFI	Kenya
Acacia planifrons WIGHT & ARN.		Dried	C.Fagg	India
Acacia polycantha WILLD. subsp. *campylacantha* (HOCHST. *ex* A.RICH) BRENAN	11/92	Seed	OFI	Tanzania
Acacia pringlei ROSE	CEH1647	Dried	OFI	Mexico
Acacia pycnantha BENTH.	635-88-05356	Dried	Kew living	(Australia)
Acacia rhemanniana SCHINZ	D3	Dried	C.Fagg	(HNBG)
Acacia riparia KUNTH	LR 1152	Dried	L.Rico	Mexico
Acacia robynsiana MERXM. & SCHREIBER	F+M 5	Dried	C.Fagg	Namibia
Acacia roigii LEON	LR 1152	Dried	L.Rico	Cuba
Acacia rosei STANDLEY	RICO 832	Dried	L.Rico	??????????
Acacia schaffneri (S.WATSON) F.J.HERM.	CEH1632	Dried	OFI	Mexico
Acacia sieberana DC.	71/90	Seed	OFI	Zimbabwe
Acacia sieberana DC. var. *woodii* (BURTT DAVY) KEAY & BRENAN	1/89	Seed	OFI	Zimbabwe
Acacia senegal (L.) WILLD. var. *leiorhachis* BRENAN	38/90	Seed	OFI	Botswana
Acacia senegal (L.) WILLD var. *senegal*	17/92	Seed	OFI	Tanzania
Acacia senegal (L.) WILLD var. *senegal*	13/92	Seed	OFI	Tanzania
Acacia sericea MART. & GAL.	TORR 9582	Dried	L.Rico	Mexico
Acacia seyal DEL. var. *fistula* (SCHWEINF) OLIV.	77/92	Seed	OFI	Malawi
Acacia seyal DEL. var. *seyal*	86/90	Seed	OFI	Kenya

Material Used In This Study.

Acacia subangulata	CEH 1585	Dried	OFI	Mexico
Acacia tequilana	LR 1201	Dried	L.Rico	Mexico
Acacia tortilis (FORSSK.) HAYNE subsp. *heteracantha* (BURCH.) BRENAN	12/90	Seed	OFI	Zimbabwe
Acacia tortilis (FORSSK.) HAYNE subsp. *heteracantha* (BURCH.) BRENAN	13/90	Seed	OFI	Zimbabwe
Acacia tortilis (FORSSK.) HAYNE subsp. *heteracantha* (BURCH.) BRENAN	30/90	Seed	OFI	Botswana
Acacia tortilis (FORSSK.) HAYNE subsp. *heteracantha* (BURCH.) BRENAN	32/90	Seed	OFI	Botswana
Acacia tortilis (FORSSK.) HAYNE subsp. *raddiana* (SAVI) BRENAN	1402/84	Seed	OFI	Senegal
Acacia tortilis (FORSSK.) HAYNE subsp. *spirocarpa* (HOCHST. ex A.RICH) BRENAN	18/90	Seed	OFI	Zimbabwe
Acacia tortilis (FORSSK.) HAYNE subsp. *spirocarpa* (HOCHST. ex A.RICH) BRENAN	85/90	Seed	OFI	Kenya
Acacia tortilis (FORSSK.) HAYNE subsp. *spirocarpa* (HOCHST. ex A.RICH) BRENAN	52/90	Seed	OFI	Zimbabwe
Acacia xanthophloea BENTH.	87/90	Seed	OFI	Malawi
Albizia harveyi FOURN.	D12	Dried	C.Fagg	(HNBG)
Albizia saman (JACQ.) MERR.	69/87	Seed	OFI	Honduras
Albizia schimperana OLIV.	D13	Dried	C.Fagg	(HNBG)
Albizia tomentosa (MICHELI) STANDL.	40/87	Seed	OFI	Mexico
Albizia versicolor WELW. *ex* OLIV.	D11	Dried	C.Fagg	(HNBG)
Calliandra calothyrus	9/89	Seed	OFI	Guatemala
Enterolobium cyclocarpum (JACQ.) CRISEB	15/86	Seed	OFI	Honduras
Faidherbia albida (DEL.) A.CHEV.	26/88	Seed	OFI	Gambia
Faidherbia albida (DEL.) A.CHEV.	25/88	Seed	OFI	Sudan
Faidherbia albida (DEL.) A.CHEV.	74/88	Seed	OFI	Ghana
Faidherbia albida (DEL.) A.CHEV.	75/90	Seed	OFI	Zambia
Faidherbia albida (DEL.) A.CHEV.	76/90	Seed	OFI	Zambia
Faidherbia albida (DEL.) A.CHEV.	84/90	Seed	OFI	Namibia
Faidherbia albida (DEL.) A.CHEV.	23/90	Seed	OFI	Zimbabwe
Faidherbia albida (DEL.) A.CHEV.	47/90	Seed	OFI	Zimbabwe
Piptadenia viridiflora	116/92	Seed	OFI	Mexico
Pithecellobium dulce (ROXB.) BENTH.	63/87	Seed	OFI	Honduras
Prosopis juliflora (SW.) DC.	CEH 1703	Dried	OFI	Guatemala

Appendix B:
Methods

More generic methods than the ones listed below are described in Dowling *et al.*, 1990 and Sambrook *et al.*, 1989. An excellent review of the methodologies, including explanations and histories of the techniques is that of Avise (1994).

B1-Cultivation of material.

Seeds were scarified, by nicking the testa opposite the embryo end with a razor blade prior to sowing. The seeds were sown in a mixture of two parts Levingtons Professional compost (M2) to one part fine gravel. Chalk was added to the mixture at approximately 5g/l. Once the cotyledons had emerged the seedlings were inoculated with a *Rhiziobia* mixture (see section B2) and coarse gravel was placed on the surface of the soil to prevent moss growth. All plants were grown in a glasshouse at 28°C and illuminated by a 16 hr. day. The plants were fed with plant food every two months and were watered as required.

B2-Collection of material for DNA analysis.

Fresh leaf material was collected from the greenhouse grown plants and stored overnight at 4°C to destarch the leaves. Excess material was stored at -20°C for up to 18 months.

Material was collected from the field in three different ways.

a) Contact paper drying:- this consisted of placing fresh leaves between sheets of newsprint in a plant press and changing the papers daily until the specimens were dry.

b) Silica gel drying:- this consisted of placing the young leaves in a snap top plastic with self-indicating silica gel according to the method of Chase and Hills (1991). Approximately 25g of silica gel was used per gram of fresh leaf tissue. Once material was thoroughly dry the leaves were placed in snap-top plastic

bags with approximately 2g of fresh silica gel. The used silica gel was recycled for use again.

c) Corrugate drying:- this involved sandwiching the leaf material between two layers of blotting paper and separating samples with aluminium corrugates. Stacked samples were clamped together and dried overnight, either over a kerosene burner or in a drying cabinet.

Healthy young leaf material free from visible signs of insect and fungal damage was collected. When the samples were returned to the laboratory they were stored at -20°C until used.

B3-Inoculation of Plants with *Rhizobia* cultures.

Seedlings grown in the greenhouse were inoculated with *Rhizobium* cultures to achieve root nodulation. The *Rhizobium* cultures were kindly supplied by Professor Janet Sprent (University of Dundee).

B3.1-Preparation of Inoculation culture and inoculation of seedlings.

A liquid culture of *Rhizobia* was used to inoculate the seedlings, this was prepared by growing up individually the 8 strains provided by Janet Sprent(see table B3.1), in Rhizobia nutrient broth for 2 days at 28°C on an orbital shaker. Then strains were mixed together and approximately 1 ml of the resulting culture was applied to the base of the seedlings.

B3.2-Maintenance of Rhizobium culture.

The cultures were maintained on *Rhizobia* nutrient agar slopes and subcultured periodically.

Strain Identification no.	Species strain isolated from
DUS 308	Acacia arenaria
DUS 48	Acacia nilotica
DUS 297	Acacia nilotica
DUS 316	Acacia tortilis spirocarpa
DUS 17	Acacia xanthophloea
DUS 318	Acacia erioloba
DUS 318	Acacia senegal rostrata
2S Silva	Acacia auriculiformis

Table B3.1. Details of Rhizobium strains used to nodulate seedlings.

B4-DNA extraction

Total DNA was extracted from the leaves according to a modified version of Doyle & Doyle (1987). The protocol is rapid and gives relatively good yields of high molecular weight DNA, however to carry out restriction analyses further purification steps are needed.

1. The tissue was macerated, and its DNA extracted as follows; 2XCTAB extraction buffer, containing 0.2% ß-mercaptoethanol, was preheated to 65°C in a water bath. Approximately 0.5g of fresh leaf material or 0.3g of dried material was ground to a powder in a pestle and mortar in liquid nitrogen with the aid of alumina powder. The tissue macerate was allowed to thaw slightly and then about 2.5 ml of extraction buffer was added to the mortar and the macerate ground to a fine paste. The remaining extraction buffer was added and mixed thoroughly with the paste. 20-30ml of extraction buffer to 1 gram of tissue (fresh weight) was used. The macerate was transferred to centrifuge tubes and then incubated at 65°C for 30 minutes. The samples were shaken every 10 minutes to resuspend the macerate. The centrifuge tubes were then removed from the water bath and allowed to cool to room temperature.

2. The samples were then crudely purified using chloroform extractions. 2 vol. of 'wet' chloroform was added to each tube and mixed gently into the extract. The sample(s) were centrifuged at c. 3000 rpm in a bench top centrifuge for 10

minutes. The uppermost aqueous layer was removed and transferred to a clean centrifuge tube. Again 2 vol. of 'wet' chloroform was mixed into each tube and centrifuged for 10 minutes.

3. The DNA was precipitated by transferring the aqueous layer to another clean centrifuge tube and adding ice cold propan-2-ol to the top of the tube (this was at least twice the vol. of the extract). The tube was then gently inverted to precipitate the nucleic acids. Usually the tube was left overnight at -20°C to increase the precipitate. The tubes were centrifuged for 10 minutes to pellet the DNA and then the supernatant poured away. Wash buffer was then added to the centrifuge tube and the DNA pellet gently resuspended. This was left for at least 1 hour. The DNA was then pelleted again in a centrifuge and the supernatant poured off. The tube was then left inverted for the DNA pellet to dry. Sometimes this process was speeded up by placing the centrifuge tube in a vacuum desiccator for 10 minutes. When the pellet was finally dry it was resuspended in 1 ml of TE (Tris-EDTA buffer, see reagents at the end of this Appendix).

The extracted DNA is now ready for purification although it was standard practise to check the concentration and intactness of the extracted DNA before proceeding with purification. (see section B6)

B5-DNA purification.

During the course of this project two different purification methods were used. The first method tested was DEAE-sephacel chromatography (described in Harris 1990), this is based on purification by ion exchange. Although this method is simple, cheap and quick, it is not very reliable. Often 100% loss of the DNA sample being purified occurred or the sample was too heavily sheared to be used. Another method of purification was tested in order to overcome the limitations of DEAE-sephacel chromatography, namely caesium chloride density gradient ultracentrifugation. I found this procedure routinely gave adequate yields of high

molecular weight DNA that could be cut by restriction enzymes. Below is the method for caesium chloride density gradient ultracentrifugation.

B5.1-Caesium chloride density gradient ultracentrifugation

1. 7.5g of ceasium chloride (CsCl) was measured into a ten ml measuring cylinder. To this was added the DNA sample, 100µl of ethidium bromide, and TE to 10ml. The cylinder was then inverted to mix the contents and dissolve the CsCl. If the volume had dropped below 10ml TE was added to bring the volume back up to 10ml. This was repeated for each DNA sample. In addition a cylinder without DNA was made up for use later in topping up the ultracentrifuge tubes and balancing the tubes.

2. The samples were then transferred to ultracentrifuge tubes and the tubes topped up with the spare solution. They were then degassed by placing the tubes in a vacuum desiccator for 10 minutes. Finally the tubes were balanced to 1mg and sealed.

3. Ultracentrifugation took place in Sorval OTD65B centrifuge with a fixed angle Sorval T865.1 rotor. The samples were centrifuged for 24 hours at 53K rpm and 20°C.

4. After centrifugation the DNA was removed from the gradient with a wide bore (19G) hypodermic needle (see figure B5.1). The ethidium bromide was then removed from the samples, by mixing the sample with TE saturated butan-1-ol until the sample was clear (ethidium bromide is preferentially miscible in butan-1-ol).

5. Each sample was diluted with an equal volume of TE, and had 2 vols of propan-2-ol (room temperature) and 1 ml of 7.5M ammonium acetate added to precipitate the DNA. The samples were left overnight at room temperature and then centrifuged to pellet the DNA sample. The supernatant was poured away and the DNA dried and resuspended in 1 ml TE. Two 500µl aliquots of the DNA solution were placed into microcentrifuge tubes and reprecipitated with propan-

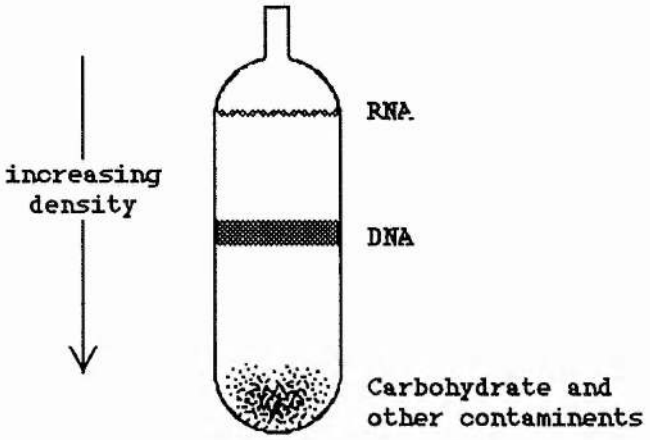


Figure B5.1. Illustration of how the crude components of the DNA extraction are separated after a caesium chloride density gradient centrifugation. The bands visualised under ultraviolet light and the band of DNA removed with a hypodermic needle.

2-ol. The DNA was pelleted for a second time, dried and resuspended in a volume of 100µl. All aliquots of the same sample were then recombined.

The DNA was now ready for restriction digestion after its concentration had been determined.

B6-DNA concentration determination

The concentration of the extracted DNA could be determined in either of two ways; measurement of the absorbance of the sample at 260 nm (1AU = 50µg double stranded DNA) or visual determination by running the sample on an agarose gel. The latter method was used as it also enables an estimate of the molecular weight of the DNA to be made.

Ethidium bromide, an aromatic organic compound, intercalates into double stranded DNA and when illuminated with ultraviolet light fluoresces visible light. The intensity of fluorescence is correlated with the amount of double stranded DNA present. This property of ethidium bromide allows the determination of the concentration of sample DNA by comparing its fluorescence with that of a standard DNA solution.

Electrophoresis of the DNA sample also enabled an estimate of the molecular weight of the DNA to be made. It was important that only high molecular weight DNA was used for the restriction digests. This ensures that bands on the subsequent autoradiographs are 'tight' and easily measured.

1. A 0.8% agarose gel was prepared (see section B8 for preparation of agarose gels).
2. The DNA samples were loaded onto the gel with the standard, 125ng calf thymus DNA present in one of the lanes. The gel was run at 40mA until the tracer dye had moved approximately 5cm and photographed on a ultraviolet light transilluminator .

B7-Restriction Digestion.

Once the DNA sample had been purified it was digested by selected enzymes. Initially a series of enzymes was tested to find which reliably cut, from these 15 were selected for the study(see Table B7).

The amount of DNA digested was in the region of 300-750ng, the exact figure was determined by the amount of purified DNA available. Less than 300ng of DNA made it difficult to complete all the probings and more than 750ng required additional restriction enzyme. The amount of DNA for one round of digests was constant.

1. The digests consisted of DNA sample,3µl, 10X digestion buffer (supplied with the enzyme), 10U enzyme (1µl) and distilled water to 30µl. Digestion took place overnight at the required temperature (usually 37°C).
2. Following overnight digestion the reaction was terminated by the addition of 1/10th volume (3µl) 'stop' buffer

B8-Agarose gel electrophoresis.

After digestion of the DNA samples the fragments produced were separated by gel electrophoresis. This separates the fragments according to their size . As mentioned earlier agarose gels were also used to determine concentration and check on molecular weight.

The concentration of agarose in the gels depended on the purpose of the gels. For concentration gels 0.8% agarose (w/v) gels were used. For separating restriction digests 1.0% agarose gels were necessary and for separating PCR products 1.5% agarose gels were used. As a general rule low percentage gels

Restriction Enzyme	Recognition site	Restriction Enzyme	Recognition site
Apa I	↓ 5′-G GGCC C-3′ 3′-C CCGG G-5′ ↑	*Nsi* I	↓ 5′-A TGCA T-3′ 3′-T ACGT A-5′ ↑
*Bam*H I	↓ 5′-G GATC C-3′ 3′-C CTAG G-5′ ↑	*Pst* I	↓ 5′-C TGCA G-3′ 3′-G ACGT C-5′ ↑
Bcl I	↓ 5-′T GATC A-3′ 3-′A CTAG T-5′ ↑	*Pvu* II	↓ 5′-CAG CTG-3′ 3′-GTC GAC-5′ ↑
Bgl II	↓ 5′-A GATC T-3′ 3′-T CTAG A-5′ ↑	*Sma* I	↓ 5′-CCC GGG-3′ 3′-GGG CCC-5′ ↑
Cla I	↓ 5′-AT CG AT-3′ 3′-TA GC TA-5′ ↑	*Sst* I	↓ 5′-G AGCT C-3′ 3′-C TCGA G-5′ ↑
*Eco*R I	↓ 5′-G AATT C-3′ 3′-C TTAA G-5′ ↑	*Stu* I	↓ 5′-AGG CCT-3′ 3′-TCC GGA-5′ ↑
*Eco*R V	↓ 5′-GAT ATC-3′ 3′-CTA TAG-5′ ↑	*Xho* I	↓ 5′-C TCGA G-3′ 3′-G AGCT C-5′ ↑
Hind III	↓ 5′-A AGCT T-3′ 3′-T TCGA A-5′ ↑		

Table B7. Details of Restriction enzymes used in this study.

(<1%) separate large fragments of DNA optimally and higher concentrations (>1%) separate small DNA fragments optimally.

Before all samples were loaded onto the gels 1/10th volume of stop buffer was added. In addition to terminating any enzyme reactions in the sample the stop buffer also acted as a sinking agent facilitating easy loading of the samples. Also bromophenol blue was added to stop buffer, this acts as a tracking dye enabling electrophoresis to be monitored.

B8.1-Preparation and electrophoresis of gels.

1. The agarose was dissolved in 1X SEB by heating over a Bunsen burner with constant mixing. After the solution had come to the boil the heat was reduced until the solution was at a 'roiling boil'. It was left for 5 minutes to ensure the agarose had completely dissolved.
2. The gel solution was moved to an orbital shaker to cool and 0.5µg/ml of ethidium bromide added. When the gel solution had cooled to approximately 60°C (hand-hot), it was poured into the gel mould, and the comb positioned.
3. When set the gel was covered in 1X SEB, the gel was submerged by approximately 5mm of buffer. The comb was then removed.
4. The DNA samples were then loaded into the wells and a current applied. For restriction digests the gel was run at a constant current of 60mA until the tracer dye had migrated ~15cm (approximately 26 hours). The gel was then transferred to a UV transilluminator and photographed.

B9-Southern Techniques

There are several methods of visualising the DNA fragments. If the amount of DNA is high and from a single source (e.g. chloroplast DNA) ethidium bromide staining will suffice. If the amount of DNA is low then the fragments can be visualised after end labelling them isotopically (e.g.with ^{32}P).

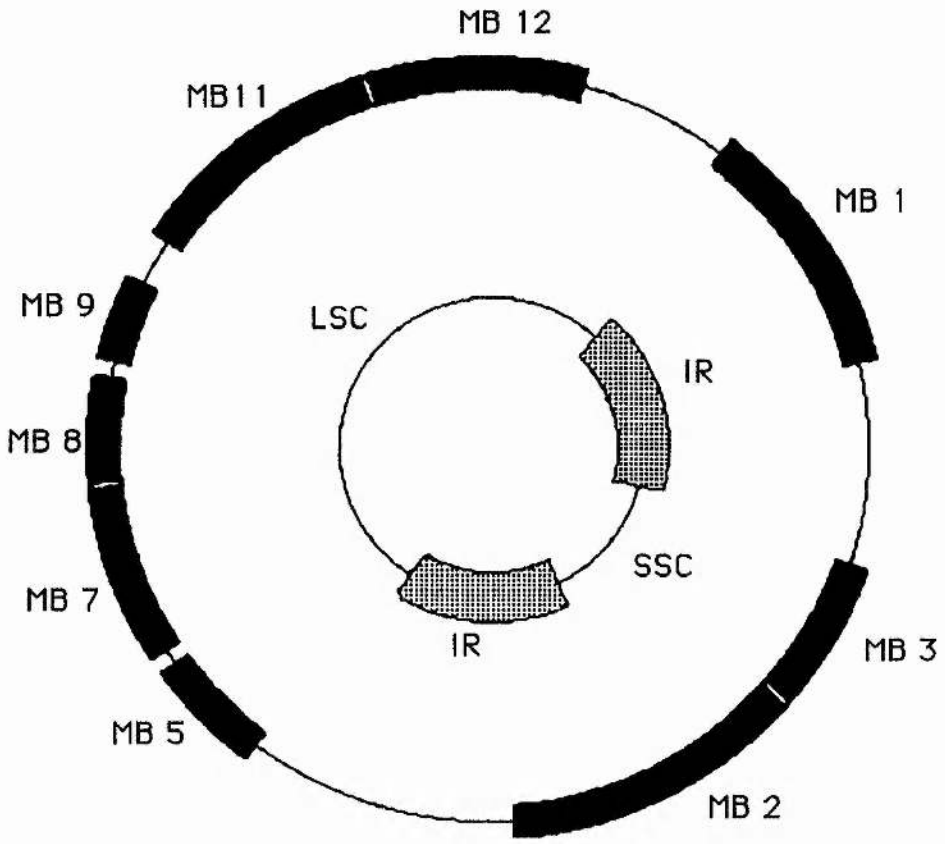


Figure B9. The inner circle represents a typical plant cp DNA molecule, IR are the invert repeats, LSC the large single copy region and SSC the small single copy region. The outer circle details the position of the MB (mung bean) chloroplast probes relative to the major structural features of the chloroplast DNA molecule.

The DNA used in the present work was heterogeneous i.e. total DNA, so the above methods could not be used to visualise the DNA fragments. Instead the technique of Southern hybridisation was used (Southern, 1975). This technique involves denaturing the DNA in the agarose gel and transferring it as single strands by capillary action to a nylon membrane.

The nylon membrane is then incubated with radioactively labelled probe DNA. This probe DNA, usually from a known region of the genome, binds to homologous sequences of DNA on the membrane. The position on the membrane where the probe has bound, can be visualised by exposing an X-ray film over the membrane. This technique allows visualisation of specific DNA fragments from the thousands of other DNA fragments that have migrated through the gel. The

preparation of the probe DNA is described in section B10. Details of the probe DNA I used are shown in figure B9 (and in table B10).

B9.1 Southern Blotting

1. After the gel had been photographed, it was immersed in denaturation buffer for 30 minutes and agitated on an orbital shaker. The gel was then briefly washed in distilled water and immersed in neutralisation buffer for 30 minutes.
2. While the two washes outlined above were taking place, the blotting apparatus was assembled (see figure B9), using 20X SSC as a transfer buffer.
3. The gel was rinsed, trimmed, inverted and placed on the blotting apparatus. The rest of the blot was assembled, and transfer occurred overnight.
4. The blot was disassembled and the top right-hand corner of the membrane removed to orientate the membrane. The membrane was then briefly washed in 2X SSC and left to air dry.
5. The DNA was fixed to the membrane by irradiating the membrane with ultraviolet light for 4 minutes. The membrane was then stored in a dry place until needed.

B9.2-Membrane prehybridisation and hybridisation.

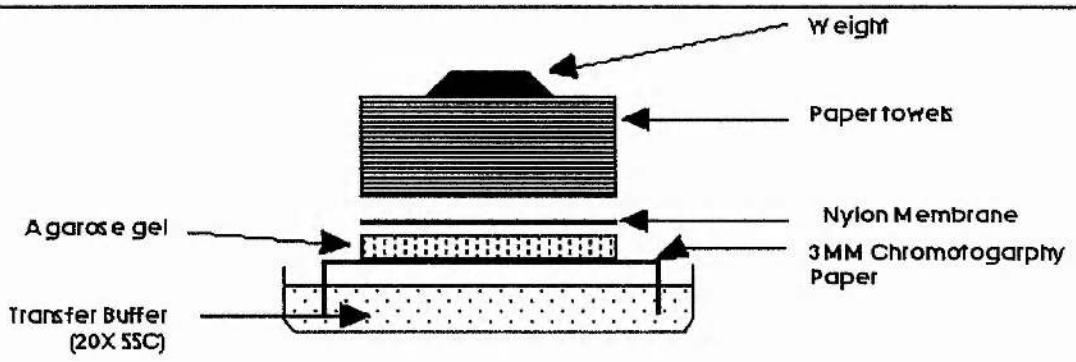


Figure B9. Details of the Southern blotting apparatus. The agarose gel was placed on chromatography paper wicks. On top of the gel the nylon membrane was placed and then paper towels. The towels draw the transfer buffer and DNA through the agarose gel. The buffer is soaked up by the towels, but the DNA is held on the nylon membrane.

Depending on the number of membranes being hybridised, they were hybridised in either a plastic bag or in a hybridisation oven. For one or two membranes a hybridisation oven was used, in excess of two the plastic bag method was used.

Oven hybridisation

1. The membrane was placed inside a hybridisation tube with the DNA side facing inwards. 10 ml of hybridisation buffer was added to the tube.

2. The membrane was prehybridised at 65°C for approximately 6 hours. 3. Denatured labelled probe DNA (see section 9.3) was added to the hybridisation buffer and the membrane and probe left to hybridise overnight at 65°C.

4. The membrane was removed from the tube and washed at the required stringency. This depended on the probe being used. See table 9.2 for details of the washes.

5. After washing excess moisture was removed from the membrane, and it was placed in an autoradiography cassette. X-ray film was overlaid and the cassette left at -70°C to expose.

Plastic Bagging Hybridisation.

1. The membranes were arranged inside a plastic bag approximately 12 times the size of the membranes.

2. 10 ml hybridisation buffer for each of the membranes was added to the bag, i.e for 4 membranes 40 ml of hybridisation buffer was added.

3. As much air as possible was removed from the bag and it was sealed. The bag was incubated at 65 °C for approximately 6 hours in a shaking water bath.

4. Denatured labelled probe DNA (see section 9.3) was injected into the bag and

Wash no.	chloroplast DNA probes	ribosomal DNA probes
1	2X SSC, 0.5 % SDS at room temperature for 30 minutes.	Rinse membranes in 2X SSC.
2	As for wash 1	1X SSC, 0.1% SDS at 65°C for 20 minutes.
3	2X SSC, 0.5 % SDS at 65°C for 30 minutes.	0.3X SSC, 0.1% SDS at 65°c for 25 minutes.

Table 9.2. Details of the washes used to remove all but the desired probe DNA.

the bag resealed. The membranes and probe were left to hybridise overnight at 65°C.

5. The membranes were removed from the tube and washed at the required stringency. This depended on the probe being used. See table 9.2 for details of the washes.

6. After washing excess moisture was removed from the membrane, and it was placed in an autoradiography cassette. X-ray film was overlaid and the cassette left at -70°C to expose.

B9.3-Probe labelling.

The quantities below are those for 1 membrane only. For more than one membrane the quantities were multiplied accordingly.

1. The probe DNA was denatured by mixing 12.6µl distilled water and 60ng of probe DNA in a large microcentrifuge tube, and placing the tube in boiling water for 10 minutes.

2. The tube was rapidly cooled on ice and 5µl HEPES (pH 6.6), 5µl DTM, 1.4µl OL, 1µl bovine serum albumin and 2.5U DNA polymerase large fragment (aka Klenow fragment) added.

3. Finally 10µCi of α-^{32}P-dCTP was added, and the labelling mix incubated for ~ 5 hours at room temperature.

4. Before use in the southern hybridisation the probe DNA was heat denatured by placing the labelling tube in boiling water for 10 minutes.

B10-Microbiological techniques.

The chloroplast and ribosomal DNA probes mentioned earlier were maintained in plasmid vectors and cloned using a bacterial host, in this case *E.coli*. This section covers the methods involved in inserting the plasmid into the

Probe	Size (kb)	Vector	Antibiotic resistance	Cloning site	Source (see below)
MB 1	16.2	pBR322	Tetracycline	*Pst* I	1
MB 2	18.8	pBR322	Tetracycline	*Pst* I	1
MB 3	9.7	pBR322	Tetracycline	*Pst* I	1
MB 5	7.5	pBR322	Tetracycline	*Pst* I	1
MB 7	11.1	pBR322	Tetracycline	*Pst* I	1
MB 8	7.0	pBR322	Tetracycline	*Pst* I	1
MB 9	5.6	pBR322	Tetracycline	*Pst* I	1
MB 11	16.5	pBR322	Ampicillin	*Sal* I	1
MB 12	13.3	pBR322	Ampicillin	*Sal* I	1
pTA71	9.1	pUC19	Tetracycline	*Eco*R I	2
pTEE3	5.0	pUC19	Ampicillin	*Eco*R I	3
pTEE5	6.5	pUC19	Ampicillin	*Eco*R I	3

Table B10. Details of probes used. Bacterial host for all plasmids is *E.coli* strain DH5α. Sources were as follows; 1. Palmer & Thompson (1981), 2. Gerlach & Bedbrook (1979), 3. King & Schaal (1990).

host, maintaining and growing the host and the subsequent extraction of plasmid DNA containing the probe

The probes used were gifts from Jeff Doyle (*Vigna* chloroplast clones), Mike O'Dell (*Triticum* rDNA clones) and Lynn King (*Taraxacum* rDNA clones).The probes were already in plasmids. However, they had to be inserted into the host bacteria (transformation). Before the host could be transformed, it had to be made competent i.e. ready to take up plasmids. Table B10 details several features of these probes.

B10.1-Production of competent cells.

1. A culture of *E.coli* strain DH5α was grown up overnight at 37°C in a Macartney bottle using nutrient broth (Oxoid).

2. 1 ml of the culture was placed in 100 ml of nutrient broth and grown at 37°C in a shaking incubator until the culture had an optical density of 0.5-0.6 at 600nm (approximately 3 hours).

3. 20 mls of the culture were then transferred to a centrifuge tube and left on ice for 15 minutes. This was then centrifuged for 10 minutes at 2,000 rpm and 4°C. The supernatant was discarded.

4. The pellet of bacterial cells was resuspended in 20 mls of 0.1M magnesium chloride and the centrifuged again for 10 minutes at 2,000 rpm and 4°C. The supernatant was discarded.

5. Finally the cells were resuspended in 4ml of 100mM calcium chloride and left for 30 minutes to 12 hours. The cells were now ready to take up plasmids.

For long term storage of competent cells, 500µl of the cell solution was gently mixed with 500µl sterile glycerol. The resulting solution was flash frozen in liquid nitrogen and kept at -70°C until needed.

B10.1-Transformation of competent cells.

1. Competent cells were used directly after production as described above, or from storage at -70°C. If cells were used from storage they were thawed completely on ice and then left for 1 hour.

2. 24µl TMC, 1µl plasmid DNA solution (containing 1ng-10ng DNA) and 50µl of the competent cell culture were placed into a sterile microcentrifuge tube. The solution was gently mixed and then left on ice for 30 minutes.

3. The microcentrifuge tube was placed in a water bath at 42°C for 2 minutes, then rapidly cooled on ice.

4. 1 ml of nutrient broth prewarmed to 37°C was added to the solution and this was incubated at 37°C for one hour.

5. 200µl of this solution was plated out onto nutrient agar containing the appropriate antibiotic (Ampicillin resistant cultures- 50µg Ampicillin/ml and Tetracycline cultures-12.5µg tetracycline/ml). The plates were then left to dry, sealed, inverted and incubated over night at 37°C.

6. Control plates were also produced concurrently with the transformation procedure using untransformed competent cells. One plate with only nutrient

agar was used to test the viability of the competent cells, and another plate with the selective antibiotic added to the agar was used to check for contamination. No bacteria should grow on the plate with the antibiotic whereas the plate with just the nutrient agar should have a good growth of bacteria.

7. Colonies that grew overnight were those of bacteria with the desired plasmid in them. These are then grown up in bulk for the production of the desired probe.

B10.3-Probe preparation.

B10.3.1-Over night cultures.

1. The appropriate antibiotic was added to 10ml of nutrient broth , and the bottle mixed thoroughly. The bottle was then inoculated with bacteria containing the desired probe i.e. either with 50µl of frozen culture in glycerol or a colony stab from the plates prepared in section B10.1.

2. The inoculated bottles (usually 4 bottles per probe were prepared) were then placed in a shaking incubator overnight at 37°C and at ~200 rpm.

3. Before extraction of plasmid DNA, 500µl of the overnight culture was mixed with 500µl of sterile glycerol to provide a new stock culture of bacteria/plasmid. This was stored at -20°C or -70°C.

B10.3.2-Extraction of Plasmid DNA

1. The overnight culture was centrifuged at 3,000-4,000 rpm for 10 minutes to pellet the bacteria. The supernatant was poured off and the bacteria resuspended in 100µl of 25 % sucrose in 50mM Tris-base ph8.0.

2. The cell suspension was transferred to a 1.5 ml microcentrifuge tube and 600µl of MSTET and 14µl of lysozyme (40µg/ml dissolved in 25 % sucrose 50mM Tris-base ph8.0) added to the tube which was then vortexed. Freshly prepared lysozyme was used.

3. The tube was placed in boiling water for 1 minute and then cooled rapidly on ice. It was then centrifuged for 45 minute at 13,000 rpm.

4. The pellet of cellular debris was removed with a sterile toothpick.

5. 200µl of Phenol/0.8% hydroxyquinoline was added to the tube and gently mixed to an emulsion. This was centrifuged for 10 minutes at 13,000 rpm and 600µl of the aqueous layer transferred to a clean microcentrifuge tube.

6. To this was added 60 µl of 7.5M ammonium acetate and 1 ml ice-cold propan-2-ol. The tube was left at -20°c for at least 1 hour, leaving overnight increased the yield of DNA significantly.

7. The tube was centrifuged for 15 minutes at 13,000 rpm to pellet the DNA. The supernatant was discarded and the DNA pellet dried.

8. The DNA was resuspended in 200µl TE and 5µl of RNase(10mg/ml) added. The tube was incubated at 37°C for no more than 30 minutes.

9. 3 phenol extractions followed by 2 'wet' chloroform extractions were used to purify the DNA.

10. After the final extraction the aqueous layer was transferred to a microcentrifuge tube and $^{1}/_{10}^{th}$ vol. of 7.5M ammonium acetate and 1 ml ice-cold propan-2-ol added.

11. The DNA was precipitated DNA at -20°C for at least 30 minutes. Finally this was centrifuged for 30 minute at 13,000 rpm, the supernatant discarded and the DNA pellet dried. It was resuspended DNA in 50µl TE.

B10.3.3-Digesting Plasmid to release probe.

1. The amount of plasmid DNA extracted was determined by running a concentration gel. For every µg of plasmid DNA 10 U of restriction enzyme was required.

2. The digest was prepared with the appropriate enzyme (see table B10) and left to digest overnight.

3. A final purification with chloroform was performed and the DNA resuspended in 50µl of TE.

4. Whether digestion had been successful and the concentration of probe if it was determined by running an aliquot of the sample on a concentration gel. The

concentration of the probe was adjusted to either 30µg/µl or 60µg/µl, and it was stored at -20°C until it was needed.

B11-Polymerase Chain Reaction experiments

The polymerase chain reaction (PCR) was used to produce randomly amplified polymorphic DNAs (RAPDs).

B11.1-DNA extraction:

This extraction protocol is a modified from that of Edwards *et al.*, 1991.

1. The testa of the seed was removed using sterile nail clippers. The cotyledons were placed in a 1.5ml microcentrifuge tube and 100µl of extraction buffer was added.

2. The cotyledons were crushed in the buffer using a microcentrifuge tube pestle (Scotlab) until the sample was homogeneous. Then an additional 900µl of extraction buffer was added, and the sample vortexed to ensure even mixing.

3. The samples were then centrifuged for 5 minutes at 13,00 rpm in a benchtop centrifuge to pellet the cellular debris. As much of the supernatant as possible was removed and transferred to a clean microcentrifuge tube.

4. To this was added 400µl of phenol/0.8% hydroxyquinone and the sample vortexed gently. This was centrifuged for 5 minutes at 13,000 rpm.

5. Again as much of the supernatant was removed as was possible and this was transferred to a clean microcentrifuge tube.

6. 400µl of wet chloroform was added to this and then the tube was vortexed gently. This was centrifuged for 5 minutes at 13,000 rpm.

7. 600µl of the supernatant was removed and transferred to a clean microcentrifuge tube. To this was added 60µl of 7.5M ammonium acetate and 1ml of ice-cold propan-2-ol. The tube was then gently inverted to precipitate the DNA.

8. The DNA was pelleted by centrifugation for 10 minutes at 13,00 rpm.

9. The supernatant was gently poured off and 1ml of Wash Buffer was added. The tube was gently vortexed to un-pellet the DNA and then left for 1 hour.

10. The DNA was pelleted by centrifugation for 10 minutes at 13,000 rpm and the supernatant gently poured off. The DNA pellet was then left to dry. When dry 100µl of HPLC grade water was added and the sample left overnight at 4°C for the DNA to resuspend.

11. The concentration and the intactness of the sample were checked on a concentration gel (see section B6). Finally the concentration of the sample was adjusted to 10ng/µl.

B11.2-The PCR reaction

1. The reaction was conducted in 500µl sterile microcentrifuge tubes.

2. The reaction components were, 0.1mM dATP, 0.1mM dCTP, 0.1mM dGTP, 0.1mM dTTP, Dynazyme buffer, 0.1µM OPH Primer (see table B11.2 for details of primers used), 1U Dynazyme (Finnymes Oy) and 50ng sample DNA. Finally sterile HPLC grade water was added to bring the volume of the reaction mix up to 50µl.

3. The reaction mixture was then gently vortexed to mix the reagents, and then the mixture spun down and overlaid with 50µl of silicone fluid (BDH).

4. Finally the outside of the tube was smeared with silicone fluid, to optimise heat transfer, and the tubes were placed in the thermal cycler.

Primer number.	Sequence (5'-3').
OPH-01	GGTCGGAGGA
OPH-02	TCGGACGTGA
OPH-03	AGACGTCCAC
OPH-04	GGAAGTCGCC
OPH-05	AGTCGTCCCC
OPH-06	ACGCATCGCA
OPH-07	CTGCATCGTG
OPH-08	GAAACACCCC
OPH-09	TGTAGCTGGG
OPH-10	CCTACGTCAG
OPH-11	CTTCCGCAGT
OPH-12	ACGCGCATGT
OPH-13	GACGCCACAC
OPH-14	ACCAGGTTGG
OPH-15	AATGGCGCAG
OPH-16	TCTCAGCTGG
OPH-17	CACTCTCCTC
OPH-18	GAATCGGCCA
OPH-19	CTGACCAGCC
OPH-20	GGGAGACATC

Table B11.2. Name and sequence of the Operon Technologies primers used in the RAPD study.

5. The samples were subjected to the following reaction conditions; an initial denaturation step of 1.5 minutes at 94°C; followed by 35 cycles of, 1 minute at 94°C, 2 minutes at 35°C and 2 minutes at 72°C; 7 minutes at 72°C (to ensure complete extension of fragments). The samples were then held at 10°C.

6. 25µl of the reaction mix was removed for analysis, and the remaining product stored at -20°C. 2.5µl of stop buffer was added to the removed aliquot, and the sample electrophoresed in an 1.5% agarose gel to visualise the DNA fragments.

7. The method of preparing an agarose gel is covered in section B8, however this protocol was modified for RAPD gels. No ethidium bromide was added to the agarose when preparing the gel, instead the gel was stained in a solution of ethidium bromide (1µg/ml) for 15 minutes following electrophoresis. The gel was the destained in distilled water for 20 minutes following staining to reduce background fluorescence.

Reagent Recipes

Rhizobium Nutrient Broth.

0.01% K_2HPO_4 (w/v)
0.04% KH_2PO_4 (w/v)
0.02% $MgSO4 \bullet 7H_20$ (w/v)
0.01% NaCl (w/v)
0.4g/L Yeast Extract (Oxoid)

For agar slopes 15g/L agar was added

2X CATB extraction buffer

2% Hexadecyltrimethylammonium bromide (aka CTAB) (w/v)
100mM Tris(hydroxymethyl)methylamine
20 mM Ethylenediaminetetra-acetic acid, di sodium salt
1.4M Na Cl
1% Polyvinylpyrrolidone 40-T (w/v)
pH 8.0

'Wet' chloroform

24:1 choroform:octan-1-ol

Wash buffer

76% Ethanol
10 mM Ammonium acetate

TE

10mM Tris(hydroxymethyl)methylamine
1mM Ethylenediaminetetra-acetic acid, di sodium salt
pH7.6

Stop Buffer

0.125M Ethylenediaminetetra-acetic acid, di sodium salt
50% glycerol (v/v)
0.1% Sodium dodecyl sulphate
1mg/ml Bromophenol blue

1X SEB (standard electrophoresis buffer)

0.04M Tris(hydroxymethyl)methylamine
0.02M Sodium acetate trihydrate
1mM Ethylenediaminetetra-acetic acid, di sodium salt
pH 7.85

Southern denaturation buffer

1.5M NaCl
0.5M NaOH

Southern neutralisation buffer

1.5M NaCl
0.5M Tris(hydroxymethyl)methylamine
1mM Ethylenediaminetetra-acetic acid, di sodium salt
pH7.2

20X SSC (saline sodium citrate)

3M NaCl
0.3M Trisodium citrate

Hybridisation buffer

0.6M NaCl
10mM Piperazine-NN'-bis-2-ethanesulphonic acid (aka PIPES) (ph6.8)
1mM EDTA-Na_2 (ph8.5)
10X Modified Denhardt's solution
Before the hybridisation buffer was added to the membranes 10μg/mL of heat denatured sonicated salmon testes DNA was added to the hybridisation buffer

100X Modified Denhardt's solution

2% Bovine skin gelatine type B
2% Ficoll® 400
2% Polyvinylpyrollidone-360
10% Sodium dodecyl sulphate
0.5% Tetrasodium pyrrophosphate

DTM

100μM dATP (2'-deoxyadenosine-5'-triphosphate, disodium)
100μM dGTP (2'- deoxyguanosine-5'-triphosphate, trisodium)
100μM dTTP (Thymidine-5'-triphosphate, trisodium)
DTM was made up in TM

TM

250mM Tris(hydroxymethyl)methylamine (pH8.0)
25mM MgCl
50mM 2-Mercaptoethanol

OL

1mM Tris(hydroxymethyl)methylamine
1mM Ethylenediaminetetra-acetic acid, di sodium salt (pH7.5)
90 OD units/mL hexaoligodeoxyribonucleotide (aka

TMC

10mM Tris(hydroxymethyl)methylamine
10mM MgCl
10mM CaCl

ph7.5

MSTET

5% Triton® X-100 (v/v)
50mM Tris(hydroxymethyl)methylamine
50mM Ethylenediaminetetra-acetic acid, di sodium salt
5% Sucrose

PCR extraction Buffer

200mM Tris(hydroxymethyl)methylamine
250mM NaCl
25mM Ethylenediaminetetra-acetic acid, di sodium salt
0.5% Sodium dodecyl sulphate
pH7.5

Appendix C
cpDNA data

Character No.	Enzyme	Probe	Site 0	↔	Change 1
1	*Xho* I	MB 3	19 + 11.5	↔	30.5
2	*Xho* I	MB 5	10.7 + 1.5	↔	12.2
3	*Xho* I	MB 7	7.4	↔	1.6 + 5.8
4	*Xho* I	MB 7	3.4 + 0.4	↔	3.8
5	*Xho* I	MB 12	22 + 18	↔	32
6	*Xho* I	MB 12	18.1	↔	7.9 + 10.2
7	*Xho* I	MB 1	8.8	↔	7.6 + 0.8
8	*Xho* I	MB 1	8.8	↔	6.5 + 2.3
9	*Xho* I	MB 1	8.8	↔	4.4 + 4.4
10	*Pst* I	MB 3	Large Fragment (Lf)	↔	12.5 + Lf
11	*Pst* I	MB 3	2.0	↔	x + y
12	*Pst* I	MB 5	Lf	↔	6.7 + x
13	*Pst* I	MB 7	Lf	↔	15 + 4.5
14	*Pst* I	MB 8+9,11	15	↔	6.1 + 8.7
15	*Pst* I	MB 8+9,11	8.7	↔	6.2 + 2.5
16	*Pst* I	MB 11	5.8	↔	4.3 + 1.5
17	*Pst* I	MB 12	Lf	↔	6.4 + x
18	*Pst* I	MB 12	Lf	↔	9.4 + x
19	*Sma* I	MB 3	17.5	↔	12.2 + 5.3
20	*Sma* I	MB 2	1.3	↔	x + y
21	*Sma* I	MB 2	2.4 + 1.6	↔	4.0
22	*Sma* I	MB 5	2.6 + 0.4	↔	3.0
23	*Sma* I	MB 5	2.6 + 1.6	↔	4.2
24	*Stu* I	MB 3	7.8 + 7.8	↔	15.6
25	*Stu* I	MB 3	7.8 + 2.1	↔	9.9
26	*Stu* I	MB 3	16 + 11.5	↔	27.5
27	*Stu* I	MB 2	6.4 + 6.4	↔	12.8
28	*Stu* I	MB 7	9.3 + 2.3	↔	11.6
29	*Stu* I	MB 7	11.6 + 8	↔	19.6
30	*Stu* I	MB 7	9.3 + 7.0	↔	16.3
31	*Stu* I	MB 7	12	↔	5.4 + 6.6
32	*Stu* I	MB 12	Lf	↔	13.5 + Lf
33	*Stu* I	MB 1	6.8 + 3.7	↔	10.5
34	*Stu* I	MB 1	5.5	↔	3.2 + 2.3
35	*Stu* I	MB 1	2.3	↔	2.0 + 0.3
36	*Bcl* I	MB 3	4.7 + 0.3	↔	5.0
37	*Bcl* I	MB 3	4.8	↔	2.9 + 1.9
38	*Bcl* I	MB 3, 2	23.0	↔	10.0 + 13.0
39	*Bcl* I	MB 2	12.0 + 7.0	↔	19.0
40	*Bcl* I	MB 2	12.0 + 1.0	↔	13.0
41	*Bcl* I	MB 5	2.3 + 2.2	↔	4.5
42	*Bcl* I	MB 5	1.8 + 0.2	↔	2.0
43	*Bcl* I	MB 5	8.0	↔	6.5 + 1.5
44	*Bcl* I	MB 5	1.5 + 1.8	↔	3.3
45	*Bcl* I	MB 7	1.9 + 2.7	↔	4.5
46	*Bcl* I	MB 7	2.7	↔	2.5 + 0.2
47	*Bcl* I	MB 7	2.7 + 2.2	↔	4.9
48	*Bcl* I	MB 7	2.2 + 1.8	↔	4.0
49	*Bcl* I	MB 7	1.8	↔	1.5 + 0.3
50	*Bcl* I	MB 7	1.8	↔	1.7 + 0.1
51	*Bcl* I	MB 7	1.8 + 3.8	↔	5.6
52	*Bcl* I	MB 7	2.8	↔	2.3 + 0.5
53	*Bcl* I	MB 8+9	12.5	↔	10.0 + 2.5
54	*Bcl* I	MB 11	4.7	↔	4.3 + 0.4
55	*Bcl* I	MB 11	4.7 + 3.5	↔	8.4
56	*Bcl* I	MB 11,12	12.5 + 1.6	↔	14.1
57	*Bcl* I	MB 11,12	12.5	↔	10.3 + 2.2
58	*Bcl* I	MB 11,12	12.5	↔	6.9 + 6.1
59	*Bcl* I	MB 1	3.5	↔	3.3 + 0.2
60	*Bcl* I	MB 1	3.5 + 1.3	↔	4.8

cpDNA data

61	*Bcl* I	MB 1	4.4	↔	4.0 + 0.4
62	*Bcl* I	MB 1	4.0	↔	3.6 + 0.4
63	*Pvu* II	MB 5	Lf	↔	13 + Lf
64	*Pvu* II	MB 5	Lf	↔	10.3 + Lf
65	*Pvu* II	MB 5	10.3	↔	7.7 + 2.6
66	*Pvu* II	MB 7	4.8	↔	4.5 + 0.3
67	*Pvu* II	MB 7	4.8 + 0.4	↔	5.2
68	*Pvu* II	MB 7,8+9	9.0	↔	7.4 + 1.6
69	*Pvu* II	MB 7,8+9	9.0	↔	2.7 + 6.3
70	*Pvu* II	MB 8+9	12.9	↔	8.2 + 4.7
71	*Pvu* II	MB 12	Lf	↔	10.0 + Lf
72	*Pvu* II	MB 12	Lf	↔	7.9 + Lf
73	*Pvu* II	MB 12	Lf	↔	7.4 + Lf
74	*EcoR* V	MB 3	Lf	↔	Lf + 1.3
75	*EcoR* V	MB 3	Lf	↔	8.5 + Lf
76	*EcoR* V	MB 3	Lf	↔	10.8 + Lf
77	*EcoR* V	MB 3	x	↔	x + 8.5
78	*EcoR* V	MB 3	Lf	↔	13 + 9.1
79	*EcoR* V	MB 3	Lf	↔	6.1 + 14
80	*EcoR* V	MB 5	3.1	↔	2.3 + 0.8
81	*EcoR* V	MB 5	3.1 + 0.4	↔	3.5
82	*EcoR* V	MB 5	3.5 + 1.1	↔	4.6
83	*EcoR* V	MB 7.	Lf	↔	2.4 + Lf
84	*EcoR* V	MB 7	Lf	↔	10.0 + 9.5
85	*EcoR* V	MB 7	9.7	↔	6.2 + 3.5
86	*EcoR* V	MB 8+9	5.6 + 1.2	↔	6.8
87	*EcoR* V	MB 11	6.0	↔	4.0 + 2.0
88	*EcoR* V	MB 11	6.0	↔	3.3 + 3.7
89	*EcoR* V	MB 11	6.0	↔	4.5 + 1.5
90	*EcoR* V	MB 11	6.1 + 3.9	↔	10.0
91	*EcoR* V	MB 11	3.9	↔	3.4 + 0.5
92	*EcoR* V	MB 12	6.5 +1.2	↔	7.9
93	*EcoR* V	MB 12	6.0	↔	3.5 + 2.5
94	*EcoR* V	MB 12	6.5	↔	4.3 + 1.2
95	*EcoR* V	MB 12	10.4	↔	4.5 + 5.9
96	*EcoR* V	MB 12	4.5	↔	3.1 + 1.4
97	*EcoR* V	MB 12	10.4	↔	9 + 1.4
98	*EcoR* V	MB 1	2.4	↔	2.0 + 2.4
99	*EcoR* V	MB 1	2.5	↔	2.3 + 0.3
100	*Nsi* I	MB 3	6.0	↔	3.6 + 2.4
101	*Nsi* I	MB 3	4.3	↔	1.3 + 3.4
102	*Nsi* I	MB 3	4.3 + 2.1	↔	6.4
103	*Nsi* I	MB 3	5.7	↔	2.4 + 3.3
104	*Nsi* I	MB 3	4.3 + 0.7	↔	5.0
105	*Nsi* I	MB 5	4.4 + 1.5	↔	5.9
106	*Nsi* I	MB 5	5.5 + 0.8	↔	6.3
107	*Nsi* I	MB 5	5.2 + 4.4	↔	9.6
108	*Nsi* I	MB 5	2.1 + 0.3	↔	2.4
109	*Nsi* I	MB 5	2.1	↔	1.9 + 0.2
110	*Nsi* I	MB 7	5.0 + 0.7	↔	5.7
111	*Nsi* I	MB 7	3.2 + 1.0	↔	4.2
112	*Nsi* I	MB 7	3.2	↔	2.5 + 0.7
113	*Nsi* I	MB 8+9	3.6	↔	x + y
114	*Nsi* I	MB 8+9	5.6 + 0.3	↔	5.9
115	*Nsi* I	MB 8+9	5.6 + 5.9	↔	11.5
116	*Nsi* I	MB 8+9,11	11	↔	4 + 7
117	*Nsi* I	MB 11,12	4.7 + 1.8	↔	6.5
118	*Nsi* I	MB 12	4.7 + 5.9	↔	10.6
119	*Nsi* I	MB 12	4.7 + 0.4	↔	5.1
120	*Hin*D III	MB 3	11.0 + 5.0	↔	16.0
121	*Hin*D III	MB 3	15.0 + 11.0	↔	26.0(Lf)
122	*Hin*D III	MB 3	5.0 + 0.5	↔	5.5
123	*Hin*D III	MB 3	5.5 + 0.2	↔	5.7

124	*Hin*D III	MB 5	2.0 + 0.4	↔	2.4
125	*Hin*D III	MB 5	2.0 + 1.2	↔	3.2
126	*Hin*D III	MB 5	1.7 + 1.2	↔	2.9
127	*Hin*D III	MB 5,7	13.0	↔	3.2 + 9.8
128	*Hin*D III	MB 5,7	13.0 + 7.5	↔	20.5
129	*Hin*D III	MB 7	7.5	↔	2.9 + 2.6
130	*Hin*D III	MB 11	8.0 + 3.1	↔	11.1
131	*Hin*D III	MB 11	3.1	↔	x + y
132	*Hin*D III	MB 12	8.0	↔	5.3 + 2.7
133	*Hin*D III	MB 12	8.0	↔	4.5 + 3.5
134	*Hin*D III	MB 12	8.0 + 2.0	↔	10.0
135	*Hin*D III	MB 12	8.0	↔	6.3 + 1.7
136	*Hin*D III	MB 12	8.0	↔	7.5 + 0.5
137	*Hin*D III	MB 12	8.0	↔	7.0 + 1.0
138	*Hin*D III	MB 12	13.0 + x	↔	Lf
139	*Apa* I	MB 3	20.0	↔	12.0 + 8.0
140	*Apa* I	MB 3	7.7 + 0.3	↔	8.0
141	*Apa* I	MB 3	7.7	↔	6.7 + 1.0
142	*Apa* I	MB 3	7.7	↔	5.7 + 2.0
143	*Apa* I	MB 2	3.3 + x	↔	Lf
144	*Apa* I	MB 2	3.3 + 0.1	↔	3.4
145	*Apa* I	MB 5	9.8 + x	↔	Lf
146	*Apa* I	MB 5	4.8 + 4.8	↔	9.6
147	*Eco*R I	MB 3	3.8	↔	3.2 + 0.6
148	*Eco*R I	MB 3	2.2	↔	2.1 + 0.1
149	*Eco*R I	MB 3	2.2	↔	x + y
150	*Eco*R I	MB 3	1.3	↔	x + y
151	*Eco*R I	MB 5	Lf	↔	4.8 + x
152	*Eco*R I	MB 5	Lf	↔	1.9 + x
153	*Eco*R I	MB 7	1.5	↔	x + y
154	*Eco*R I	MB 7	2.5 + 1.1	↔	3.6
155	*Eco*R I	MB 7	4.3	↔	2.2 + 2.1
156	*Eco*R I	MB 7	2.1	↔	1.9 + 0.2
157	*Eco*R I	MB 7	2.1 + 0.6	↔	2.7
158	*Eco*R I	MB 7	1.7 + 0.8	↔	2.5
159	*Eco*R I	MB 7	1.7	↔	x + y
160	*Eco*R I	MB 7	3.6 + 5.7	↔	9.6
161	*Eco*R I	MB 7	3.6	↔	1.9 + 1.7
162	*Eco*R I	MB 7	5.7	↔	5.0 + 0.7
163	*Eco*R I	MB 7,8+9	5.4 + 5	↔	10.4
164	*Eco*R I	MB 8+9	3.8 + 0.3	↔	4.1
165	*Eco*R I	MB 8+9	3.8	↔	2.9 + 0.9
166	*Eco*R I	MB 8+9	x + y	↔	2.9
167	*Bam*H I	MB 3	1.8 + 0.9	↔	2.7
168	*Bam*H I	MB 3	2.3	↔	2.1 + 0.2
169	*Bam*H I	MB 3	3.4	↔	3.1 + 0.3
170	*Bam*H I	MB 3	3.4	↔	3.0 + 0.4
171	*Bam*H I	MB 2	3 + 0.8	↔	3.8
172	*Bam*H I	MB 2	5	↔	3.8 + 1.2
173	*Bam*H I	MB 5	4.6	↔	4.3 + 0.3
174	*Bam*H I	MB 5	4.6 + x	↔	7.8
175	*Bam*H I	MB 5	4.6 + x	↔	6.0
176	*Bam*H I	MB 5	3.3 + 1.2	↔	4.5
177	*Bam*H I	MB 5	Lf	↔	4.5 + 16.0
178	*Bam*H I	MB 5	13.3	↔	9.0 + 4.3
179	*Bam*H I	MB 5	9.0 + 2.5	↔	11.5
180	*Bam*H I	MB 1	4.2 + 3.0	↔	7.2
181	*Bam*H I	MB 1	9.0	↔	7.5 + 1.5
182	*Bam*H I	MB 1	4.4 + 0.2	↔	4.6
183	*Bam*H I	MB 1	4.4	↔	4.0 + 0.4
184	*Bam*H I	MB 1	4.4 + 1.5	↔	5.9
185	*Bam*H I	MB 1	3.1	↔	2.7 + 0.4
186	*Cla* I	MB 3	1.6	↔	0.9 + 0.7

187	*Cla* I	MB 3	4.0 + 2.6	↔	6.6
188	*Cla* I	MB 3	4.0	↔	3.5 + 0.5
189	*Cla* I	MB 3	4.0	↔	1.8 + 2.2
190	*Cla* I	MB 3	5.3	↔	4.3 + 1.3
191	*Cla* I	MB 2	1.7 + 0.1	↔	1.8
192	*Cla* I	MB 2	1.5	↔	x + y
193	*Cla* I	MB 5	8.5	↔	5.5 + 3
194	*Cla* I	MB 5	8.5	↔	3.7 + 4.8
195	*Cla* I	MB 5	2.9	↔	1.4 + 1.5
196	*Cla* I	MB 5	2.9	↔	2.4 + 0.5
197	*Cla* I	MB 1	1.4 + 1.9	↔	3.3
198	*Cla* I	MB 1	2.3 + 0.1	↔	2.4
199	*Cla* I	MB 1	2.3	↔	2.0 + 0.3
200	*Cla* I	MB 1	4.8	↔	4.7 + 0.1
201	*Cla* I	MB 1	7.7	↔	5.5 + 2.2
202	*Cla* I	MB 1	8.2	↔	3.9 + 4.2
203	*Bgl* II	MB 3	3.7 + 0.6	↔	4.3
204	*Bgl* II	MB 3	3.4 + 5.1	↔	8.5
205	*Bgl* II	MB 2	Lf	↔	5.3 + x
206	*Bgl* II	MB 2	Lf	↔	7.2 + x
207	*Bgl* II	MB 2	Lf	↔	9.2 + x
208	*Bgl* II	MB 5	5.1 + 1.1	↔	6.2
209	*Bgl* II	MB 5	5.1	↔	3.5 + 1.6
210	*Bgl* II	MB 5	5.1	↔	1.9 + 3.2
211	*Bgl* II	MB 7	3.9 + 2.3	↔	6.2
212	*Bgl* II	MB 7	6.2 + 1.8	↔	8.0
213	*Bgl* II	MB 7	2.3 + 1.8	↔	3.9
214	*Bgl* II	MB 7	3.9 + 1.8	↔	5.7
215	*Bgl* II	MB 7	3.9	↔	3.5 + 0.4
216	*Bgl* II	MB 7	1.8	↔	1.5 + 0.3
217	*Bgl* II	MB 7	1 + 0.3	↔	1.3
218	*Bgl* II	MB 1	3.5	↔	1.9 + 1.8
219	*Bgl* II	MB 1	1.9 + x	↔	2.5
220	*Bgl* II	MB 1	5.2 + x	↔	6.7
221	*Bgl* II	MB 1	1.5	↔	x + y
222	*Bgl* II	MB 1	2.8 + 0.4	↔	3.2
223	*Apa* I	MB 8+9,11	8.5	↔	8.2 + 0.3
224	*Apa* I	MB 12	Lf	↔	9.9 + x
225	*Apa* I	MB 12	Lf	↔	11.3 + x
226	*Apa* I	MB 12	Lf	↔	5.3 + x
227	*Apa* I	MB 12	5.3 + 0.4	↔	5.7
228	*Cla* I	MB 7	3.2 + 0.9	↔	4.1
229	*Cla* I	MB 7	3.2	↔	2.2 + 1.0
230	*Cla* I	MB 7	1.2	↔	x + y
231	*Cla* I	MB 7	5.1	↔	1.3 + 3.8
232	*Cla* I	MB 7	11.0	↔	5.0 + 7.0
233	*Cla* I	MB 8+9	11.0 + 5.0	↔	16.0
234	*Cla* I	MB 8+9	5.0	↔	3.7 + 1.3
235	*Cla* I	MB 8+9	5.0 + 2.3	↔	7.3
236	*Cla* I	MB 8+9	3.6	↔	2.9 + 0.7
237	*Cla* I	MB 8+9	3.6 + 0.4	↔	4.0
238	*Cla* I	MB 8+9	3.6	↔	3.1 + 0.5
239	*Cla* I	MB 11	3.2 + 0.7	↔	3.9
240	*Cla* I	MB 11	3.2	↔	2.9 + 0.3
241	*Cla* I	MB 11	1.3 + 0.1	↔	1.4
242	*Cla* I	MB 11	5.0 + 0.4	↔	5.4
243	*Cla* I	MB 11	3.5 + 5.0	↔	8.5
244	*Cla* I	MB 11	3.5 + 2.5	↔	6.0
245	*BamH* I	MB 8+9	x + y	↔	1.3
246	*BamH* I	MB 8+9	2.6	↔	x + y
247	*BamH* I	MB 8+9	3.5 + 0.3	↔	3.8
248	*BamH* I	MB 11	6.0	↔	3.0 + 3.0
249	*BamH* I	MB 11	4.0 + 2.0	↔	6.0

250	*BamH* I	MB 11	x + y	↔	1.6
251	*BamH* I	MB 11	3.4	↔	3.2 + 0.2
252	*BamH* I	MB 11	3.4	↔	2.3 + 1.1
253	*BamH* I	MB 11	4.4	↔	2.1 + 2.3
254	*BamH* I	MB 12	3.2 + 1.1	↔	4.3
255	*BamH* I	MB 12	3.2 + 0.6	↔	3.8
256	*BamH* I	MB 12	5.0 + 0.6	↔	5.6
257	*BamH* I	MB 12	5.6	↔	2.8 + 2.8
258	*BamH* I	MB 12	5.6	↔	3.6 + 2.0
259	*BamH* I	MB 12	5.0	↔	4.9 + 0.1
260	*EcoR* I	MB 1	3.4	↔	3.2 + 0.2
261	*EcoR* I	MB 1	3.4 + 0.4	↔	3.8
262	*EcoR* I	MB 1	1.5	↔	x + y
263	*EcoR* I	MB 1	3.8	↔	2.5 + 1.3
264	*EcoR* I	MB 1	3.8 + 1.1	↔	4.9
265	*Bgl* II	MB 8+9	6.0 + 2.9	↔	8.9
266	*Sst* I	MB 5	12 + x	↔	Lf
267	*Sst* I	MB 5	12	↔	3.8 + 8.2
268	*Sst* I	MB 7	16.0	↔	7.0 + 9.0
269	*Sst* I	MB 8+9	4.6	↔	2.8 + 1.8
270	*Sst* I	MB 8+9	Lf	↔	1.5 + x
271	*Sst* I	MB 11	20.0 + 2.0	↔	22.0
272	*Sst* I	MB 11	5.3 + 1.3	↔	6.6
273	*Sst* I	MB 11	6.6 + 5.3	↔	11.9
274	*Sst* I	MB 11	5.3	↔	3.9 + 1.7
275	*Sst* I	MB 11	1.3	↔	0.8 + 0.5
276	*Sst* I	MB 11,12	9.0	↔	7.5 + 1.5
277	*Sst* I	MB 12	Lf	↔	4.0 + Lf
278	*Xho* I	MB 2	3.3 + 0.7	↔	4.1
279	*Xho* I	MB 2	13.7 + 3.3	↔	17.0
280	*Xho* I	MB 5	10.3	↔	8.5 + 1.8
281	*Xho* I	MB 7	7. 9 + 3.8	↔	12.1
282	*Xho* I	MB 12	x + y	↔	1.0
283	*Xho* I	MB 12	16.5	↔	10.1 + 6.4
284	*Xho* I	MB 12	16	↔	2.8 + 13.7
285	*Xho* I	MB 1	15 + 3	↔	18
286	*Xho* I	MB 1	15	↔	2.6 + 13.4
287	*Xho* I	MB 1	10.3	↔	0.8 + 9.5
288	*Pst* I	MB 3	14	↔	9.7 + 3.6
289	*Pst* I	MB 5	6.7 + 1.3	↔	8.0
290	*Pst* I	MB 8+9	7.2	↔	6.3 + 0.9
291	*Pst* I	MB 11	5.8	↔	5.3 + 0.5
292	*Pst* I	MB 12	Lf	↔	8.0 + x
293	*Sma* I	MB 8+9	Lf	↔	13.7 + x
294	*Stu* I	MB 7	10	↔	6.5 + 3.5
295	*Stu* I	MB 7	3.5	↔	3.0 + 0.5
296	*Stu* I	MB 8+9	Lf	↔	5.8 + x
297	*Bcl* I	MB 5	Lf + 8	↔	20
298	*Bcl* I	MB 7	2.2	↔	1.1 + 1.1
299	*Bcl* I	MB 7	2.6	↔	1.5 + 1.1
300	*Bcl* I	MB 7	2.2 + 1.1	↔	3.3
301	*Bcl* I	MB 8+9	15.5 + 3.1	↔	18.6
302	*Bcl* I	MB 11	4.4	↔	2.4 + 2.0
303	*Bcl* I	MB 11,12	13 + 8		21
304	*Bcl* I	MB 1	4.0 + 1.0	↔	5.0
305	*Bcl* I	MB 1	4.2	↔	3.9 + 0.3
306	*Bcl* I	MB 1	3.5	↔	2.6 + 0.9
307	*Bcl* I	MB 1	x + y	↔	1.9
308	*Bcl* I	MB 3	4.3 + 5.7	↔	10.0
309	*Pvu* II	MB 5	Lf	↔	5.8 + x
310	*Pvu* II	MB 8+9,11	12 + 11	↔	23
311	*Pvu* II	MB 11,12	Lf	↔	14.6 + 3.8
312	*Pvu* II	MB 12	Lf	↔	5 + x

313	*Apa* I	MB 11	19.8	↔	10.4 + 9.4
314	*Apa* I	MB 12	5.3	↔	4.5 + 0.8
315	*Sst* I	MB 5	10.0	↔	8.0 + 2.0
316	*Sst* I	MB 11	x + y	↔	2.0
317	*Sst* I	MB 11	5.3	↔	4.3 + 1.0
318	*Sst* I	MB 11	5.3	↔	4.7 + 0.6
319	*Sst* I	MB 11	6.6	↔	4.7 + 1.9
320	*Sst* I	MB 11	6.6	↔	3.6 + 2.9
321	*EcoR* V	MB 3	13.0	↔	4.0 + x
322	*EcoR* V	MB 3	Lf	↔	13.0 + x
323	*EcoR* V	MB 3	Lf	↔	8.1 + x
324	*EcoR* V	MB 1	x + 2.4	↔	2.7
325	*EcoR* V	MB 1	x + 2.4	↔	3.5
326	*Nsi* I	MB 3	4.3 + x	↔	4.6
327	*Nsi* I	MB 3	6.0 + 2.1	↔	8.1
328	*HinD* III	MB 5,7	13.0	↔	6.0 + 7.0
329	*HinD* III	MB 8+9	14.8	↔	5.8 + 9.0
330	*HinD* III	MB 11	2.9 + 0.8	↔	3.7
331	*HinD* III	MB 11	3.1	↔	2.5 + 0.6
332	*HinD* III	MB 1	5.0 + 7.0	↔	12.0
333	*HinD* III	MB 1	1.7 + 0.5	↔	2.2
334	*EcoR* I	MB 2	1.9 + 4.0	↔	5.9
335	*EcoR* I	MB 2	4.0	↔	1.1 + 2.9
336	*EcoR* I	MB 5	4.0	↔	3.7 + 0.3
337	*EcoR* I	MB 5	4.0	↔	2.5 + 1.5
338	*EcoR* I	MB 5	4.0	↔	2.3 + 1.7
339	*EcoR* I	MB 7	3.0	↔	2.8 + 0.2
340	*EcoR* I	MB 7	3.0	↔	2.6 + 0.4
341	*EcoR* I	MB 7	3.0	↔	2.3 + 0.7
342	*EcoR* I	MB 7	11.6	↔	2.7 + 8.9
343	*EcoR* I	MB 8+9	5.2	↔	4.6 + 0.6
344	*EcoR* I	MB 8+9	5.2	↔	4.3 + 0.9
345	*EcoR* I	MB 8+9	1.9	↔	1.8 + 0.1
346	*EcoR* I	MB 1	3.8	↔	3.5 + 0.3
347	*EcoR* I	MB 1	0.3 + 3.8	↔	4.1
348	*BamH* I	MB 3	1.8	↔	1.6 + 0.2
349	*BamH* I	MB 2	2.0	↔	1.7 + 0.3
350	*BamH* I	MB 5	4.6	↔	3.5 + 1.1
351	*BamH* I	MB 5	3.3	↔	3.0 + 0.2
352	*BamH* I	MB 7	8.6	↔	6.2 + 2.3
353	*BamH* I	MB 7	8.6	↔	4.5 + 4.1
354	*BamH* I	MB 7	4.3 + 2.6	↔	6.9
355	*BamH* I	MB 7	4.5	↔	3.9 + 0.6
356	*BamH* I	MB 11	12.0	↔	10.0 + 2.2
357	*BamH* I	MB 11	2.6	↔	2.5 + 0.1
358	*BamH* I	MB 11	6.0 + 2.5	↔	8.5
359	*BamH* I	MB 11	2.6	↔	2.0 + 0.6
360	*BamH* I	MB 11	X + Y	↔	1.6
361	*BamH* I	MB 1	4.4 + 0.7	↔	5.1
362	*Cla* I	MB 3	2.5 + 0.3	↔	2.8
363	*Cla* I	MB 3,2	0.8 + 1.6	↔	2.4
364	*Cla* I	MB 3	6.0	↔	5.4 + 0.6
365	*Cla* I	MB 7	3.2	↔	1.5 + 1.7
366	*Cla* I	MB 7	11.5	↔	10.0 + 1.5
367	*Cla* I	MB 7	10.0	↔	6.5 + 3.5
368	*Cla* I	MB 7	3.5 + 1.5	↔	5.0
369	*Cla* I	MB 7	5.6 + 4.7	↔	10.3
370	*Cla* I	MB 7	10.5	↔	2.3 + 8.3
371	*Cla* I	MB 8+9	2.5 + 0.8	↔	3.3
372	*Cla* I	MB 8+9	2.5	↔	1.8 + 0.7
373	*Cla* I	MB 11	8.5 + 1.3	↔	9.8
374	*Cla* I	MB 11	5.1	↔	4.7 + 0.4
375	*Cla* I	MB 11	5.1 + 3.8	↔	8.9

376	*Cla* I	MB 11	3.5 + 0.4	↔	3.9
377	*Cla* I	MB 1	2.3	↔	1.9 + 0.4
378	*Cla* I	MB 1	1.7	↔	1.5 + 0.2
379	*Bgl* II	MB 5	6.2	↔	2.6 + 3.6
380	*Bgl* II	MB 5	3.6	↔	3.2 + 0.4
381	*Bgl* II	MB 5	2.6	↔	2.4 + 0.2
382	*Bgl* II	MB 7	8.6	↔	6.1 + 2.5
383	*Bgl* II	MB 7	6.1	↔	4.4 + 1.7
384	*Bgl* II	MB 7	8.6	↔	4.9 + 3.7
385	*Bgl* II	MB 7	4.9	↔	3.4 + 1.5
386	*Bgl* II	MB 7	2.3 + 1.7	↔	5.0
386	*Bgl* II	MB 7	2.3	↔	2.0 + 0.3
387	*Bgl* II	MB 7	1.5 + 0.3	↔	1.8
389	*Bgl* II	MB 7	4.4	↔	3.9 + 0.5
390	*Bgl* II	MB 8+9	4.5	↔	3.9 + 0.6
391	*Bgl* II	MB 8+9	5.2 + 0.6	↔	5.8
392	*Sma* I	MB 11	18.0	↔	3.0 + 15.0
393	*Bcl* I	MB 5	5.0 + 0.7	↔	5.7
394	*Pvu* II	MB 8+9	4.5	↔	2.75 + 2.75
395	*EcoR* V	MB 2	10.8	↔	6.2 + 4.6
396	*HinD* III	MB 7	4.3	↔	1.0 + 3.3
397	*HinD* III	MB 8+9	4.7	↔	2.3 + 2.4
398	*HinD* III	MB 8+9	11.0	↔	9.0 + 2.0
399	*HinD* III	MB 8+9	11.2	↔	5.2 + 6.0
400	*HinD* III	MB 12	13.0	↔	10.0 + 3.0
401	*Bam* HI	MB 3	4.2 + 9.2	↔	13.4
402	*Bam* HI	MB 7	x + y	↔	1.9
403	*Bam* HI	MB 7	9.0	↔	7.7 + 1.3
404	*Bam* HI	MB 11	2.3 + 0.2	↔	2.5
405	*Cla* I	MB 1	5.5	↔	3.8 + 1.7
406	*Cla* I	MB 3	3.5	↔	2.3 + 1.2
407	*Bgl* II	MB 7,8+9	4.2 + 0.5	↔	4.7

In the above table; character no. refers to the character number used in the data matrix (see following pages); enzyme and probe number refer to which restriction enzyme and mung bean probe combination the site mutation was visible with. The last two columns detail the restriction site change. The weight of the fragments is in kilobases. Large fragment refers to a fragment whose size could not be determined accurately because of its large molecular weight, x and /or y refers to fragments whose size could not be determined due either to their small size, or their presence on an adjacent probe which was not used, e.g. MB4.

Input data matrix

Taxa	Character No. 1111111111222222222233333333334444444444555555555 1234567890123456789012345678901234567890123456789012345678
Faidherbia	100001000111000010010001?0000101010001001000010000000100000
niloticakraussia	000000010000010000000000000000000000000000000000010000010010100
niloticatomentos	000000010000010000000000000000000000000000000000010000010010100
tortilisheteraca	000000010000010000000000000000000000000000000000010000000010100
tortilisspirocar	000000010000010000000000000000000000000000000000010000000010100
abyssinica	000000010000010000000000000000000000000000000000010000000010100
arenaria	001000010000010000000000000000000000000000000000010000000010100
erioloba	000000010000010000000000000000000000000000000000010000000010100
galpinii	000000000001001000000000000000000000000000000000000000101000000
karroo	000000010000010000000000000000000000000000000000100010000000100
seyalseyal	000000010000010000000100010000000000000000000000010000000010100
sieberana	000000010000010000000000000000000000000000000000010000000010100
xanthophloea	000000010000010000000100010000000000000000000000010000000010100
tortilisraddiana	000000010000010000000000000000000000000000000000010000000010100
polycanthacampyl	000000000001001000000000000000000000000000000000000000101000000
caffra	000000000001001000000000000000000000000000000000000000101000000
nigrescens	000000000001001000000000000000000000000000000000000000101000000
burkei	000000000001001000000000000000000000000000000000000000101000000
senegalleiorhach	000000000001001000000000000000000000000000000000000100101000000
senegalsenegal	000000000001000000000000000000000000000000000000000100101000000
melliferadetinen	000000000001000000000000000000000000000000000100000000101000000
chariessa	000000000001001000000000000000000000000000000000000000101000000
angustissima	100100100011011010000000001000001001000000001000000000000001
niloticasubalata	000000010000010000000000000000000000000000000000010000010010100
seiberanawoodii	000000010060010000000000000000000000000000000000010000000010100
Albiziaversicolo	100000100111000010010001?001001001001100000000000000000000000000
Albiziaharveyi	100000001111100001001000?001100001000100000000000100000000
Albiziaschimpera	100010001111000010010101?001000001000100010000000010000000
Albiziatomentosa	100000000111010010010001?001000011100100000000000010000010
Albiziasaman	100000001111010010010101?001000001001100011100010000001000
Pithecellobium	110000000111000010010?11?100000000000100000000000000000000
Enterolobioum	100000001111000010010001?001000001000100000000000000000000
Piptadenia	000000010000000101000000000000000000000000000010001000010000
farnesiana	000000010000010000000000000000000000000000010001000000010100
schaffneri	000000010000010000000000000000000000000010001000000000010100
pringlei	000000010000010000100000000000000000000010001000000000010100
amentacea	000000010000010000100000000000000000000010001000000000010100
pennatula	000000010000110000000000000000000000000010001000000000010100
hebecladachobien	000000010000010000000000000000000000000000000010000000010100
exuvialis	000000010000010000000000000000000000000000000010000000010100
rhemmiana	000000010000010000000000000000000000000000000010000000010100
leuderitziileud	000000010000010000000000000000000000000000000010000000010100
drepanolobium	000000010000010000000000000000000000000000000010000000010100
seyalfistula	000000010000010000001000100000000000000000000010000000010100
melanoyxlon	100000000111010010010001000000000100010000000000000000000000
alata	100000000111011000010001000000000100010000000000000010000000
koa	100000000111010010010001000000000100010000000000000000000000
pyncantha	1?000000011101000010?010000000001000100000000000100000000
paradoxa	100000000111010010010101000000000100010000000000000000000000
mearnsii	1?000000011101001001?010000000001000100000000000000000000
caven	000000010000010000000000000000000000000000010001000000010100
dolichocephala	000000010000010000000000000000000000000000010000000010100
greggii	000000000010010100000000000000000000000000000000000001000000
glomerosa	000000000010010000000000000000000000000001000000000000001000000
mammifera	100000000111010010010001000100000110010000000000000010000000
sericea	000000000011010010000000000000000000000000000000000000000?00
rosei	100100100011011010000000001000001001000000001000000000000001
chamelensis	?000001000100110100000000?00000?000000000001000000000000001
montigena	000000000010010100000000000000000000000001000000000000010?????
ataxacantha	000000000010??000000010000000000000000000000000000000000000000
breviscapa	000000000010010100000000000000000000000001000000000000001000000
Prosopisjuliflor	?000000000000010000000000000000000000000001000100000000010000
deamon	000000010000010000000000000000010000000001000100??000101000
cucuyo	000000010000010000000000000000000000000000000001000??000101000
roigii	000000010000010000000000000000010000000001000100??000101000
choriophylla	000000010000010000000000000000000000000001?001000??000101000
tequilana	100100100011011010000000001000001001000000001000??000000001
riparia	000000000010010100000000000000000000000000000000??0?000000
gaumeri	000000000010010100000000000000000000000010000000000??0?000000
macracantha	000000010000010000000000000000000000000001000100010000??000101000
persicflora	000000000010010000000000000000000000000010000000000??01000000

cpDNA data

Input data matrix (continued)

```
                                                                  1111111111111111
                 5666666666677777777778888888889999999999000000000001111111
                 9012345678901234567890123456789012345678901234567890123456
-----------------------------------------------------------------------------
Faidherbia       0000100010001000000100100000000001000010101000011001010001
niloticakraussia 0000000000000000000000000000001000000000000001000000000100000
niloticatomentos 0000000000000000000000000000001000000000000001000000000100000
tortilisheteraca 0000000000000000000000000000001000000000000000000000000000000
tortilisspirocar 0000000000000000000000000000001000000000000000000000000000000
abyssinica       0000000000000000000000000000001000000000000000011000000000000
arenaria         0000000000000000000000000000001000000000000000000000000000000
erioloba         0000000000000000000000000000001000000000000000000000000000000
galpinii         0000000000000001000000000001000000001000000000101001000101
karroo           0000000000000000000000000000001000000000000000000000000000000
seyalseyal       0000000000000000000000000000001000000000000000000000000000000
sieberana        0000000000000000000000000000001000000000000000000000000000000
xanthophloea     0000000000000000000000000000001000000000000000000000000000000
tortilisraddiana 0000000000000000000000000000001000000000000000000000000000000
polycanthacampyl 0000000000000001000000000001000000001100000000101001000101
caffra           0000000000000001000000000001000000001000000000101001000101
nigrescens       0000000000000001000000000001000000001000000000101001000101
burkei           0000000000000001000000000001000000001000000000101001000101
senegalleiorhach 0000000000000001000000000001000000001000000000101001000101
senegalsenegal   0000000000000001000000000001000000001000000000101001000101
melliferadetinen 0000000000000001000000000001000000001000000000101001000101
chariessa        0000000000000001000000000001000000001000000000101001000101
angustissima     0000?1101011?1000000000010000000000000101000011100100010l
niloticasubalata 0000000000000000000000000001000000000000000001000000000100000
seiberanawoodii  0000000000000000000000000001000000000000000000000000000000000
Albiziaversicolo 1010100001001000010000000010100000000100000000110010007?11
Albiziaharveyi   1000?0000100100000000100000001000010001000000000110010107?1
Albiziaschimpera 1010100001007010000010100110000000000010000000011101000?11
Albiziatomentosa 1010100000001000000001000000001000010001000000001100100l101
Albiziasaman     1010100000001000000011100110000000100010000000011101000?1
Pithecellobium   107110000000100001100111000000000000010001000011011100101
Enterolobioum    1010100001001000000001000100000000000100000001110010007?1
Piptadenia       0100000100000000000000100001000010000010000000011001000101
farnesiana       0000000000000000100000000000000100000000000000010001000000
schaffneri       0000000000000000100000000000000100000000000000000000000000
pringlei         0000000000000000000000000000000000000000000000000000000000
amentacea        0000000000000000000000000000000000000000000000000000000000
pennatula        0000000000000000000000000000000000000000000000000000000000
hebecladachobien 0000000000000000000000000000001000000000000000000000000000000
exuvialis        0000000000000000000000000000001000000000000000000000000000000
rhemmiana        0000000000000000000000000000001000000000000000000000000000000
leuderitziileud  0000000000000000000000000000001000000000000000100000000000000
drepanolobium    0000000000000000000000000000001000000000000000000000000000000
seyalfistula     0000000000000000000000000000001000000000000000000000000000000
melanoyxlon      0010100000001000000011007?????????000000000000?????0??????
alata            0010100000001000000011007?????????010000000000?????0??????
koa              0010100000001000000011007?????????000000000000?????0??????
pyncantha        0010100000001000000011107?????????000000000000?????0??????
paradoxa         0010100000001000000011107?????????000000000000?????0??????
mearnsii         0070100000001000000011107?????????000000000000?????0??????
caven            0000000000000000000000007?????????000001000000?????0??????
dolichocephala   0000000000000007?????000?????????????????000000?????0??????
greggii          0000000000001007?????000?????????????????000000?????0??????
glomerosa        0000000000001007?????000?????????????????000000?????0??????
mammifera        1010000000001000000000007?????????????????010000?????0??????
sericea          0000000000001000000000007?????????????007000000000?????0??????
rosei            0000011071011001000000007?????????????????010000?????0??????
chamelensis      0000???0?011001?????????????????????????????010000?????0??????
montigena        0000000000001000000000007?????????????????000000?????1??????
ataxacantha      0000000000000007?????000?????????????????000000?????0??????
breviscapa       0000000000001000000000007?????????????????000000?????1??????
Prosopisjuliflor 0000???10000000?????7000?????????????????000000?????0??????
deamon           0000000010100000000000000000000000000007?????????????????
cucuyo           0000000010000001000000000000000000000007?????????????????
roigii           0000000010100001000000000000000000000007?????????????????
choriophylla     0000000000000001000000000000000000000007?????????????????
tequilana        0000011010110100000000001000000000010107?????????????????
riparia          0000100070000001000000000000000000007?0000?????????????????
gaumeri          0000100000000000000000000000000000001000?????????????????
macracantha      0000000000000001000000000000000000000007?????????????????
persicflora      0000000000000001000000000010000000100007?????????????????
```

Input data matrix (continued)

```
                  1111111111111111111111111111111111111111111111111111111111111111
                  1112222222222333333333344444444455555555566666666667777777
                  7890123456789012345678901234567912345678901234567890123456
-----------------------------------------------------------------------------------
Faidherbia        01011??000000000001000001000111000001010110000001100001000
niloticakraussia  00000000000000010000000000000000000000000000000000000000000
niloticatomentos  00000000000000010000000000000000000000000000000000000000000
tortilisheteraca  00000000000000010000000000000000000000000000000000000000000
tortilisspirocar  00000000000000010001000000000000000000000000000000000000000
abyssinica        00000000000000010000000000000000000000000001000000000000000
arenaria          00000000000000010000000000000000000000000000000000000000000
erioloba          00100000000000010000000000000000000001000000000000000000000
galpinii          11000100000000107010000000001000000100001000000007000000
karroo            00000000000000010000000000000000000000000000000000000000000
seyalseyal        00000000000000010000000000000000000000000000000000000000000
sieberana         00000000000000010000000000000000000000000000000000000000000
xanthophloea      00000000000000010000000000000000000000000000000000000000000
tortilisraddiana  00000000000000010010000000000000000000000000000000000000000
polycanthacampyl  11000100000000101010000000001000000100001000000001000000001
caffra            11000100000000100010000000001000000100001000000001000000000
nigrescens        11000100000000100010000000001000000100001000000001000000000
burkei            11000100000000100010000000001000000100001000000001000000001
senegalleiorhach  11000100000000100010000000001000000100001000010001000000100
senegalsenegal    11000100000000100010000000001000000100001000010001000000100
melliferadetinen  11000100000000100010000000001000000100001000000001000001000
chariessa         11000100000000100010000000007000000100001000000001000001000
angustissima      11000?10001010000000000100011000000100000000010000000000000
niloticasubalata  00000000000000010000000000000000000000000000000000000000000
seiberanawoodii   00000000000000010000000000000000000000000000000000000000000
Albiziaversicolo  01010??000000000000000001001001000011001010000001100101000
Albiziaharveyi    01010??000000000000000001001001000001001010001001100101000
Albiziaschimpera  01010??100000000010100001001001001001001010000001100110000
Albiziatomentosa  01010??000000100000100101001001000001001010001001100100000
Albiziasaman      01010??100000000000000000100001100000110101000100110010000
Pithecellobium    01011???100000000001000?101001110000100000001001100000010
Enterolobioum     01010??000001000000000100101?001000??????71000000110110000
Piptadenia        01000100000000000000100010000011001010000101010001000000000
farnesiana        00000000000000000000000000000000000010000000000100000000000
schaffneri        00000000000000000000000000000000000010000000000100000000000
pringlei          00000100010000000000000000000000000010010010000100000000000
amentacea         00000100010000000000000000000000000010010010000100000000000
pennatula         00000000000000000000000000000000000000000000000000000000000
hebecladachobien  00000000010000010000000000000000000000000000000000000000000
exuvialis         00000000000000001000000000000000000000000000000000000000000
rhemmiana         00000000000000001000000000000000000000000000000000000000000
leuderitziileud   00000000000000001000000000000000000000000000000000000000000
drepanolobium     00000000000000001000000000000000000000000000001000000000000
seyalfistula      00000000000000001000000000000000000000000000000000000000000
melanoyxlon       ???100000000000?????????001001001000001000010000001010100000
alata             ???100000000000?????????001000001000001001010000001000100000
koa               ???100000000000?????????001001001000001000010000001000100000
pyncantha         ???100000000000?????????00100117100000100001000001000100000
paradoxa          ???100001000000?????????001001001000001000010000001010100000
mearnsii          ???100000000000?????????001001?010000010000100000001010100000
caven             ???000000100000?????????000000000000001000000000001000000000
dolichocephala    ???000000100000?????????000000000000001000000010000000000000
greggii           ???001000000000?1??????000000010000001000010000000000000100
glomerosa         ???001000000000?????????000000010000001000010000000000000107
mammifera         ???100000000001070???????101000001000001001010001001000100100
sericea           ???001000000000?????????000000000000001000010000000000000100
rosei             ???000100010100?????????010000100000001000000000010000000000
chamelensis       ??????1??????0?????????????00??000000100000000???0000000000
montigena         ???001000000000?????????000000010000001000010000000000000000
ataxacantha       ???001000000000?????????????????000000100001000000000000000
breviscapa        ???001000000000?????????000000010000001000010000000000000100
Prosopisjuliflor  ???000000000000?????????000000010000001000000000100100000000
deamon            ???000000000000000000000000000000000001000010000010000000000
cucuyo            ???000000000000000000000000000000000001000000000010000000000
roigii            ???000000000000000000000000000000000001000010000010000000000
choriophylla      ???000000000000000000000000000000000001001000000010000000000
tequilana         ???000100010100000000001000010000000100100000010000000000000
riparia           ???001000000001000001000000010000000100101000001001000010000
gaumeri           ???000100000001000001001000001000000100001000000010000100000
macracantha       ???000000000000?????????000000000000001000000000000000000000
persicflora       ???001000000000000000000000001000000100001000000011000000000
```

Input data matrix (continued)

```
                  1111111111111111111111122222222222222222222222222222222222
                  7778888888888999999999900000000011222222222233333333334444
                  7890123456789012345678901234567890123456789012345678901234
-------------------------------------------------------------------------------
Faidherbia        1000101010000010100000001010100010000100000000000010010111
niloticakraussia  0000000000000000000000100000010000000000010000000000000000
niloticatomentos  0000000000000000000001000000100000000000100000000000000000
tortilisheteraca  0000000000000000000000000001001001000001000000000000000000
tortilisspirocar  0000000000000000000000000001001001000001000000100000000000
abyssinica        0000000000000000000000000000100000000000100000000000000000
arenaria          0000000000000000000000000001000000000001000000000000000000
erioloba          0000000000000000000000001001000000000001000000000000000000
galpinii          0000000001000111000000100010000000000100000100000000000001
karroo            0000000000000000000000000001000000000001000000000000000000
seyalseyal        0000000000000000000000000001000000000001000000000000000000
sieberana         0000000000000000000000000011000000000001000000000000000000
xanthophloea      0000000000000000000000000001000000000001000000000000000000
tortilisraddiana  0000000000000000000000000001000000000001000000100000000000
polycanthacampyl  1000100000100011100000001000100000000001000001001000000001
caffra            1100100000100011100000001000100000000??????0001000000000001
nigrescens        1100100000100011100000001000100000000????0001000000?????
burkei            1100100000100011100000101000100000000001000001000000000001
senegalleiorhach  1100100000100011100000101000100000000001000001000000000001
senegalsenegal    1100100000100011100000001000100000000001000001000000000001
melliferadetinen  1100100000100011100000001000100000000001000001000000000001
chariessa         1100100000100011100000001000100000000??????0001000000000001
angustissima      1110100001000010101010001000100000000000010100001000000000
niloticasubalata  0000000000000000000001000000100000000000100000000000000000
seiberanawoodii   0000000000000000000000000011000000000001000000000000000000
Albiziaversicolo  1100000000000010100000101010100010000?001010001001000?????
Albiziaharveyi    1100100010000010000000000101000100007001000000000100000011
Albiziaschimpera  1100110010000010?1000010101010001000010010000000100000011
Albiziatomentosa  1101001010000010100100111011100010011100100000000000101011
Albiziasaman      1100110010000010100000101010100010000110001000000000100011
Pithecellobium    1100100110010010000000001011010011000100100000000100000011
Enterolobioum     1100000010001010100000101010001010000100100000000100000011
Piptadenia        1100100000000000100010001000010000000100100011000010000001
farnesiana        0000000000100100000000000001100000000000110000010000000000
schaffneri        0000000000100100000000000001100000000000110000010000000000
pringlei          0000000000000010000000000000010000000000011000000000000000
amentacea         0000000000000010000000000000010000000000????000000000000000
pennatula         0000000000100100000000000000010000000000010000000000000000
hebecladachobien  0000000000000000000000000000100100100000100000000000000000
exuvialis         0000000000000000000000000000010000000000010000000001000000
rhemmiana         0000000000000000000000000000010000000000010000000000000000
leuderitziileud   0000000000000000000000000000010000000000010000000000000000
drepanolobium     0000000000000000000000000000010000000000010000000000000000
seyalfistula      0000000000000000000000000000010000000000010000000000000000
melanoyxlon       1000100010010010000000101010100?????10010000?000000010011
alata             1100100010010010101000?010101000?????00010010?000100001011
koa               1000100010010010100000101010100?????10010000?0000??010011
pyncantha         1100100010010010100000101010100?????10010000?000?00000011
paradoxa          1000100010010011100000101010100?????00010000?0000??010011
mearnsii          1100100010010010100000101010100?????10010000?000000000011
caven             0000000000000000000000000001100000000000110007010001000000
dolichocephala    0000000000000010000000000000010000000000???000?0000??000000
greggii           1100100000010011100000101000100000000001000007000000000001
glomerosa         1??000000010011100000101???100??????00???000?000000000001
mammifera         1100101010010010100100?0101110000000000000?0?000000000000
sericea           1100100000000010100000101000100000000000???000?000001000001
rosei             1110000001010010101010101000100000000000???010?000010000000
chamelensis       1110000001010010101010101000100000000000???010?000010000000
montigena         1100100010110011100000101000100000000001000007000000000001
ataxacantha       ??00?000?00?00111000001010??100??????00???000?000000000001
breviscapa        1100100010010011100000101000100000000001000007000000000001
Prosopisjuliflor  1100000010010100100000101000110100100001000007000000000001
deamon            0000000000000000010000101000100000000000017?000000???01000
cucuyo            0000000000000000000000000001000000000001?7000000???01000
roigii            0000000000000000010000001001100000000000017?000000???01000
choriophylla      0000000000000000000000000001000000000007????000000???01000
tequilana         1110000001010110101010001??100000000007???01000000???10000
riparia           1100100000010111100000001000100000000007????000000???00001
gaumeri           1000100010010111000000001000100000000000100??000000???00001
macracantha       0000100000000000100000000000100000000000017?000000???00000
persicflora       1100100000100111100000001000100000000000100??010000???00001
```

Input data matrix (continued)

```
                2222222222222222222222222222222222222222222222222222223333333
                4444455555566666666667777777777888888888899999999990000000
                4578905678901234567890123456789012345678901234567890123456
-------------------------------------------------------------------------------
Faidherbia      0100010000000100000101010001100007000000000000000000000000
niloticakraussia 0000000000000000000000000000000007000000000000000000000000
niloticatomentos 0000000000000000000000000000000007000000000000000000000000
tortilisheteraca 0000000000000000000000000000000007000000000000000000000000
tortilisspirocar 0000000000000000000000000000000007000000000000000000000000
abyssinica      0000000000000000000000000000000007000000000000000000000000
arenaria        0000000000000000100000000000000007000000000000000000000000
erioloba        0000000000000000000000000000000007000000000000000000000000
galpinii        0000101100001100000000011001000007000000000000000000000000
karroo          0000000000000000000000000000000007000000000000000000000000
seyalseyal      0000000000000000000000000000000007000000000000000000000000
sieberana       0000000000100000000000000000000007000000000000000000000000
xanthophloea    0000000000000000000000000000000007000000000000000000000000
tortilisraddiana 0000000000000000000000000000000007000000000000000000000000
polycanthacampyl 0000101100001100000000011001000007000000000000000000000000
caffra          0000007100001100000000011001000007000000000000000000000000
nigrescens      7000001100001100000000011001000007000000000000000000000000
burkei          0000001100001100000000011001000007000000000000000000000000
senegalleiorhach 0000001100000100000000011001000007000000000000000000000000
senegalsenegal  0000001100000100000000011001000007000000000000000000000000
melliferadetinen 0000001100001100000000011001000007000000000000000000000000
chariessa       0000007100001100000000011001000007000000000000000000000000
angustissima    1110000000000110000000000110000007000000000000000000000000
niloticasubalata 0000000000000000000000000000000007000000000000000000000000
seiberanawoodii 0000000000?0000000000000000000007000000000000000000000000
Albiziaversicolo 7100011010001100000001100010000?000000000000000000000000
Albiziaharveyi  0100011010000100000001100010000?000000000000000000000000
Albiziaschimpera 0100011010000101001000110001000007000000000000000000000000
Albiziatomentosa 0100010010000100000010????7100000?000000000000000000000000
Albiziasaman    0100011010010100000001100010000?000000000000000000000000
Pithecellobium  0101011010000101000000010001000007000000000000000000000000
Enterolobioum   0100010010000110000000110001000007000000000000000000000000
Piptadenia      0000000001001100000000110001000007000000000000000000000000
farnesiana      0000000000000010000000000000000007000000000000000000000000
schaffneri      0000000000000101000000000000000007000000000000000000000000
pringlei        0000000000000100000000000000000007000000000000000000000000
amentacea       0000000000000000000000000000000007000000000000000000000000
pennatula       0000000000000001000000000000000007000000000000000000000000
hebecladachobien 0000000000000000000000000000000007000000000000000000000000
exuvialis       0000000000000000000000000000000007000000000000000000000000
rhemmiana       0000000000000000000000000000000007000000000000000000000000
leuderitziileud 0000000000000000000000000000000007000000000000000000000000
drepanolobium   0000000000000000010000000000000007000000000000000000000000
seyalfistula    0000000000000000000000000000000007000000000000000000000000
melanoyxlon     0100017????700100000001?1000101000010000100000100000000000
alata           0100017????700100000007?1000100100001100111010110010010000
koa             0100017?????0?1000000017?1000101000010000100000100000000010
pyncantha       0100017????70?1?0000001?1000100010000001111111010001000000
paradoxa        0100017????701100000001?0000100000000000100000100000000000
mearnsii        0100017?????001?0000001?0000100010000010101100100001100010 1
caven           0?0000?????0101000000?0000000000000000000000000000000000000
dolichocephala  0?0000?????0000000000?0000000000000000000000000000000000000
greggii         0?0000?????0110000000?0100100000000000000000000000000000000
glomerosa       0?0000?????0?000?00000?0100101000000000000000000000000000000
mammifera       0100017????701100000107000010000000000000010100001000000000
sericea         0100007????701100000007000010000000000000000000001000101000
rosei           1110007????701110000000?????100000000000000000000000000000
chamelensis     1?1000?????0111000000?????????000000000000000000000000000000
montigena       0?0000?????0110000000?0100100000100000000000000000000??7000
ataxacantha     0?0000?????0110070000?0100000000000000007?0000000000000000
breviscapa      0?0000?????0110000000?0100100000100000000000000000000000000
Prosopisjuliflor 0100017????701100000007?1000100001000000001000010000000000
deamon          0000000000??0100000000000000000007000000000000000000000000
cucuyo          0000000000??0100000000000000000007000000000000000000000000
roigii          0000000000??0100000000000000000007000000000000000000000000
choriophylla    0000000000??0100000000000000000007000000000000000000000000
tequilana       1110000000??1100000000001100000007000000000000000000000000
riparia         0010001101071000000001100000000?000000000000000000000000
gaumeri         0000011010??1000000001100000000?000000000000000000000000
macracantha     0000000000??0100000000000000000007000000000000000000000000
persicflora     0010001100??1000000001100000000?000000000000000000000000
```

Input data matrix (continued)

```
                 3333333333333333333333333333333333333333333333333333333333333333
                 0001111111111222222222233333333334444444445555555555666666
                 7890123456789012345678901234567890123456789012345678901234
-----------------------------------------------------------------------------------
Faidherbia       0000000000000000000000000001000000000000000000000100010000
niloticakraussia 0000000000000000000000000000000000000000000000000000000000
niloticatomentos 0000000000000000000000000000000000000000000000000000000000
tortilisheteraca 0000000000000000000000000000000000000000000000000000000000
tortilisspirocar 0000000000000000000000000000000000000000000000000000000000
abyssinica       0000000000000000000000000000000000000000000000000000000000
arenaria         0000000000000000000000000000000000000000000000000000000000
erioloba         0000000000000000000000000000000000000000000000000000000000
galpinii         0000000000000000000000000000000000000000000000000000000000
karroo           0000000000000000000000000000000000000000000000000000000000
seyalseyal       0000000000000000000000000000000000000000000000000000000000
sieberana        0000000000000000000000000000000000000000000000000000000000
xanthophloea     0000000000000000000000000000000000000000000000000000000000
tortilisraddiana 0000000000000000000000000000000000000000000000000000000000
polycanthacampyl 0000000000000000000000000000000000000000000000000000000000
caffra           0000000000000000000000000000000000000000000000000000000000
nigrescens       0000000000000000000000000000000000000000000000000000000000
burkei           0000000000000000000000000000000000000000000000000000000000
senegalleiorhach 0000000000000000000000000000000000000000000000000000000000
senegalsenegal   0000000000000000000000000000000000000000000000000000000000
melliferadetinen 0000000000000000000000000000000000000000000000000000000000
chariessa        0000000000000000000000000000000000000000000000000000000000
angustissima     0000000000000001000000000000000001000000100000000000000000
niloticasubalata 0000000000000000000000000000000000000000000000000000000000
seiberanawoodii  0000000000000000000000000000000000000000000000000000000000
Albiziaversicolo 0000000000000000000000000?1000000000000000000000100010000
Albiziaharveyi   0000000000000000000000000?1000000000000000000000100010000
Albiziaschimpera 0000000000000000000000000?1000000000000000000000100010000
Albiziatomentosa 0000000000000000000000000000000000000000000000000100010000
Albiziasaman     0000000000000000000000000?1000000000000000000000100010000
Pithecellobium   0000000000000000000000000000000000000000000000000100010000
Enterolobioum    0000000000000000000000000?1000000000000000000000100010000
Piptadenia       0000000000000000000000000000000000000000000000000000000000
farnesiana       0000000000000000000000000000000000000000000000000000000000
schaffneri       0000000000000000000000000000000000000000000000000000000000
pringlei         0010010000000000000000000000000000000?0000000000000000000
amentacea        0010000000000000000000000000000000000000000000000000000000
pennatula        0000000000000000000000000000000000000000000000000000000000
hebecladachobien 0000000000000000000000000000000000000000000000000000000000
exuvialis        0000000000000000000000000000000000000000000000000000000000
rhemmiana        0000000000000000000000000000000000000000000000000000000000
leuderitziileud  0000000000000000000000000000000000000000000000000000000000
drepanolobium    0000000000000000000000000000000000000000000000000000000000
seyalfistula     0000000000000000000000000000000000000000000000000000000000
melanoyxlon      1100000001000000000000000010001001000?1100000000000110010101
alata            1000100101100010010000001100010010010011000000000111010111
koa              1000000001000000001000000100010010001?1101010000001100101101
pyncantha        1000000011100000000000001101O?101000101??001000001110101101
paradoxa         0000001101101000000000001100000100000101000000001100111101
mearnsii         1000000011000000000000011000??11000001??0001100011101010l
caven            0000000000000000100000000000000?1000000000000000000000000
dolichocephala   0000000000000???000000000000000000000000000000000000000000
greggii          0000000000000???000001000000000000000000000001110000000100
glomerosa        0000000000000???000000000000000000000000000???0000000001
mammifera        0000000000101010000000000000000000000000000000100110101
sericea          0000000000000000001000000000000000000000000000000000000000
rosei            0000000002????010000000000000??1000000100000000000000101
chamelensis      0000000000000???000000000000000??1000000100000000000000101
montigena        0000000000000000000011100000000000000000000011000000101
ataxacantha      00010?000000007??000000100000000000007?000000000000000?01
breviscapa       0000000000000000000001110000000000000000000011000000101
Prosopisjuliflor 0000000000100???000100000000000??1000000000000110010100
deamon           0000000000000001000??000010000000000000000000000000000000
cucuyo           0000000000000000000??000010000000000000000000000000000000
roigii           0000000000000000000??000010000000000000000000000000000000
choriophylla     0000000000000000000??000010000000100000000000000000000000
tequilana        0000000000000001000??000010000000000000001000000100000100
riparia          0000000000000000000??011010000000000000000000000000010100
gaumeri          0000000000000001000??001010000001000000000000000000010100
macracantha      0000000000000000000??000010000001000000000000000000000000
persicflora      0000000000000000000??000010000000100000000000000000010000
```

Input data matrix (continued)

```
                    333333333333333333333333333333333333344444444
                    666667777777777888888888899999999990000000000
                    567890123456789012345678901234567890123456789
-------------------------------------------------------------------
Faidherbia          00000000000000000000000000000000000000000000
niloticakraussia    00000000000000000011000000000000000000000000
niloticatomentos    00000000000000000011000000000000000000000000
tortilisheteraca    00000000000000000011000000000000000000000000
tortilisspirocar    00000000000000000011000000000000000000000000
abyssinica          00000000000000000011000000000000000000000000
arenaria            00000000000000000011000000000000000000000000
erioloba            00000000000000000011000000000000000000000000
galpinii            00000000000000000011001000000000000000000000
karroo              00000000000000000011000000000000000000000000
seyalseyal          00000000000000000011000000000000000000000000
sieberana           00000000000000000011000000000000000000000000
xanthophloea        00000000000000000011000000000000000000000000
tortilisraddiana    00000000000000000011000000000000000000000000
polycanthacampyl    00000000000000000011001000000000000000000000
caffra              00000000000000000011001000000000000000000000
nigrescens          00000000000000000011001000000000000000000000
burkei              00000000000000000011001000000000000000000000
senegalleiorhach    00000000000000000011001000000000000000000000
senegalsenegal      00000000000000000011001000000000000000000000
melliferadetinen    00000000000000000011001000000000000000000000
chariessa           00000000000000000011001000000000000000000000
angustissima        00000000000000000011000000000010000000100000
niloticasubalata    00000000000000000011000000000000000000000000
seiberanawoodii     00000000000000000011001000000000000000000000
Albiziaversicolo    00000000000000000000000000000000000000000000
Albiziaharveyi      00000000000000010100000000000000000000000000
Albiziaschimpera    00000000000000010110000000000000000000000000
Albiziatomentosa    00000000000000010100000000000000000000000000
Albiziasaman        00000000000000010100000000000000000000000000
Pithecellobium      00000000000000010100000000000000000000000000
Enterolobioum       00000000000000010100000000000000000000000000
Piptadenia          00000000000000000011000001000000000000000000
farnesiana          00000000000000000011000000000000000000000000
schaffneri          00000000000000000011000000000000000000000000
pringlei            00000000000000000011000000000000000000000000
amentacea           00000000000000000011000000000000000000000000
pennatula           00000000000000000011000000000000000000000000
hebecladachobien    00000000000000000011000000000000000000000000
exuvialis           00000000000000000011000000000000000000000000
rhemmiana           00000000000000000011000000000000000000000000
leuderitziileud     00000000000000000011000000000000000000000000
drepanolobium       00000000000000000011000000000000000000000000
seyalfistula        00000000000000000011000000000000000000000000
melanoyxlon         00000010000001101001000000?000000000000000000
alata               00000000000011101001000000?000000000000000000
koa                 01110??000001101001000000?000000000000000000
pyncantha           01100??000101111001101100?000000000000000000
paradoxa            01111000000011010010000000?000000000000000000
mearnsii            01100000001011010010011007000000000000000000
caven               00000000000000000011000000000000000000000000
dolichocephala      00000000000000000011000000100000000000000000
greggii             01000010000000000110000000000000000100000000
glomerosa           01000000000000?????????????000000000000000000
mammifera           11000000000010101000000000000000000000000000
sericea             01000000000000000000000000000000000000000000
rosei               01000000000000000010000000000010000000100000
chamelensis         01000000000000001000000000010000000100000000
montigena           01000010000000100000000011000000001000000000
ataxacantha         01000000001000?????????????000000000000000000
breviscapa          01000010000000100000000011000000001000000000
Prosopisjuliflor    01000101110000000110000000100000000000000000
deamon              000000??000000000011000000000000000000000100
cucuyo              000000??000000000011000000000000000001001000
roigii              000000??000000000011000000000000000000000101
choriophylla        000000??000000000011000000000000000000000000
tequilana           010000??000000?????????????001000001110000O
riparia             010000??000000000011000001000001010100010000
gaumeri             010000??000000000011001001001001011000000010
macracantha         000000??000000000011000000001000000001000000
persicflora         010000??000000000011001000000001000000000000
```

cpDNA data

Appendix D:
RAPD data.

Character No.	Primer	Size (base pairs)
1	OPH-03	1387
2	OPH-03	1270
3	OPH-03	1243
4	OPH-03	1163
5	OPH-03	1065
6	OPH-03	1018
7	OPH-03	925
8	OPH-03	882
9	OPH-03	800
10	OPH-03	761
11	OPH-03	724
12	OPH-03	688
13	OPH-03	620
14	OPH-03	557
15	OPH-03	469
16	OPH-12	2245
17	OPH-12	2103
18	OPH-12	1911
19	OPH-12	1795
20	OPH-12	1585
21	OPH-12	1535
22	OPH-12	1441
23	OPH-12	1311
24	OPH-12	1193
25	OPH-12	1120
26	OPH-12	1051
27	OPH-12	1018
28	OPH-12	986
29	OPH-12	956
30	OPH-12	926
31	OPH-12	897
32	OPH-12	842
33	OPH-12	816
34	OPH-12	790
35	OPH-12	765
36	OPH-12	717
37	OPH-12	650
38	OPH-12	629
39	OPH-12	549
40	OPH-12	512
41	OPH-12	462
42	OPH-12	429
43	OPH-12	378
44	OPH-12	344
45	OPH-01	1824
46	OPH-01	1714
47	OPH-01	1517
48	OPH-01	1347
49	OPH-01	1233
50	OPH-01	1129
51	OPH-01	1003
52	OPH-01	888
53	OPH-01	810
54	OPH-01	761
55	OPH-01	670
56	OPH-01	607
57	OPH-01	549
58	OPH-01	512
59	OPH-01	460
60	OPH-01	396
61	OPH-16	1585
62	OPH-16	1489
63	OPH-16	1313
64	OPH-16	1233
65	OPH-16	1157
66	OPH-16	1051
67	OPH-16	985
68	OPH-16	862
69	OPH-16	806
70	OPH-16	729
71	OPH-16	681
72	OPH-16	635
73	OPH-16	551
74	OPH-16	531
75	OPH-16	495
76	OPH-16	461
77	OPH-16	396
78	OPH-14	1033
79	OPH-14	972
80	OPH-14	885
81	OPH-14	805
82	OPH-14	731
83	OPH-14	685
84	OPH-14	641
85	OPH-14	579
86	OPH-14	540
87	OPH-14	468
88	OPH-14	435
89	OPH-14	358
90	OPH-14	268
91	OPH-07	1554
92	OPH-07	1467
93	OPH-07	1334
94	OPH-07	1281

95	OPH-07	1211
96	OPH-07	1149
97	OPH-07	1130
98	OPH-07	1076
99	OPH-07	1026
100	OPH-07	956
101	OPH-07	885
102	OPH-07	840
103	OPH-07	732
104	OPH-07	671
105	OPH-07	632
106	OPH-07	557
107	OPH-07	488
108	OPH-07	441
109	OPH-13	1195
110	OPH-13	1127
111	OPH-17	1660
112	OPH-17	1463
113	OPH-17	1372
114	OPH-17	1245
115	OPH-17	1165
116	OPH-17	1127
117	OPH-17	1053
118	OPH-17	955
119	OPH-17	867
120	OPH-17	812
121	OPH-17	735
122	OPH-17	686
123	OPH-17	640
124	OPH-17	574
125	OPH-17	512
126	OPH-17	492
127	OPH-17	452
128	OPH-05	1810

129	OPH-05	1720
130	OPH-05	1636
131	OPH-05	1554
132	OPH-05	1477
133	OPH-05	1368
134	OPH-05	1334
135	OPH-05	1267
136	OPH-05	1204
137	OPH-05	1115
138	OPH-05	1031
139	OPH-05	978
140	OPH-05	904
141	OPH-05	856
142	OPH-05	790
143	OPH-05	747
144	OPH-05	688
145	OPH-05	650
146	OPH-05	595
147	OPH-05	561
148	OPH-05	481
149	OPH-02	1560
150	OPH-02	1329
151	OPH-02	1245
152	OPH-02	1165
153	OPH-02	1090
154	OPH-02	981
155	OPH-02	847
156	OPH-02	730
157	OPH-02	676
158	OPH-02	626
159	OPH-02	602
160	OPH-02	494
161	OPH-02	423

Input data matrix

```
          Character No.          1111111111222222222233333333334444444444555555555566
Accession                123456789012345678901234567890123456789012345678901234567890123456789012
-------------------------------------------------------------------------------------------------
het30/90                 00010001010100000110010010010001001001000000000001001000000100
het69/92                 00010000010100000000010010000100000011000010000000000000011000
het27/92                 00100000100001000110010000000001010001000000????????????????01
het29/92                 00100100000101100110010010000010001001000000000111000000000001
radeilat                 00010000000101100000001110100101000000001000000001001000100000
radlamu                  00010000001000000000010010101010000001000000000001001000100000
rad86/5555               00011000000000001011011010100100000000000000000110101010100000
rad1402/84               00010010000100000000000000000000000000010000100010101010110000
rad1240/84               00010000010100101110011010000100000000000000000101000000100000
radwadi                  00010000000101110110011010000100000000100000????????????????00
tortnegev                00010101001000001100110101010010000000000000000001001110110000
tort1065/82              00010010010000000000000001010000000110101000000001001110110??
tortA/3                  10010000001000000001010000100010000000000000000001000110100011
spiro70/92               10010000001000000000011100000000100100100000000001001001100000
spiro110/87              00010001000100000000011011010000110010100000000001001000100011
spiro1340/84             00010101010000100000010100010100010110000010000001001100110000
spiro15/92               00010000100101101010011100000000001010000000110000110100010000
spiro131/91              00010100010101101010011000000001001011010000000100100100010000
planifrons               01000001001010001001010000010000000001100100000100000101000000
```

Input data matrix (continued)

```
                                                    11111111111111111111111111
          66666667777777777888888888899999999900000000001111111111222222
          34567890123456789012345678901234567890123456789012345678901234
-----------------------------------------------------------------------------
het30/90  01110110010000000110011000000110100101100000000100000100000000
het69/92  00010110011010000011101010100000100101100000010100000100000000
het27/92  00111110011010010010001010100000000001111000000100000100000100
het29/92  00010111111010001010001000101110101001100000000111001100000000
radeilat  00010010011001100010101001010001010001100000101100011100001000
radlamu   00010001010100000010011010102????????????????????0100000001000000
rad86/5555 10010010011001100010111000000001010001100010001100000100010100
rad1402/84 00010010011001100011111001110001010001000010111100000100010010?
rad1240/84 00010010011001100010111001000001010001100000101100001001000000
radwadi   00010010011001100010101001010001000001100010101101000100000000
tortnegev 00110110001001101011011000100000000011000000110100000100111000
tort1065/82 ????????????????00011011000100000000001000000002?000001001100000
tortA/3   00110100000001101010011000000000000100000001001000111001100000
spiro70/92 01010000100001101011101000100000111001000001100100100101010100
spiro110/87 00011010100000101010011100000000001001000001100100100101010010
spiro1340/84 00010100100001101010011000000000000001000000110100011100100000
spiro15/92 01010000011000000010000010002?????????????????1100000101010000
spiro131/91 00010000011000000010001000000010100100100100011100000100000001
planifrons 00010011000000000101001000000001000000000000010001001001100010000
```

Input data matrix (continued)

```
                111111111111111111111111111111111111
                222223333333333444444444455555555566
                567890123456789012345678901234567890l
-------------------------------------------------------
het30/90        00010000100100000000000000001000001000
het69/92        00000010000000100000001010010010010000
het27/92        00000000010000001000101000010001110000
het29/92        01000000010000001000110000010100110000
radeilat        00001000000010000010101100010100010000
radlamu         00000001000010000010111100010000000000
rad86/5555      10001001100010100010101001010000001000
radl402/84      ??????????????????????10110101010011001
radl240/84      10001100000010000010101001010111010000
radwadi         10001001000110010010101011010111010000
tortnegev       01100010001011000010101011010000010101
tort1065/82     00000000000001000000000000111000010101
tortA/3         01000110011011000000101010011100101010
spiro70/92      00110100010000000010101000101100010000
spiro110/87     00101000001100100010111001010100100l0
spiro1340/84    01010100011000000110101001000100101010
spiro15/92      00000001110000001000011011010000011010
spiro131/91     00100000100000100010101000110000011000
planifrons      00010000000100000000110000001000001000
```