o l at T

A molecular systematic study of the genus *Encephalartos*.

Maria Elizabeth Coetzer

Dissertation presented in order to qualify for the degree *Magister Scientiae* in the Faculty of Natural and Agricultural Sciences (Department of Botany and Genetics: Division Genetics) at the University of the Orange Free State.

June 2000

Supervisor: Prof. J.J. Spies
Co-Supervisors: Prof. A-M. Oberholster and Dr P.J. Vorster

I am a cycad. You did not know me in the Triassic period of the Mesozoic era, hundred and ninety million years ago. I evolved more than hundred and eighty million years before your ancestors. I saw the appearance of the dinosaur and their death. I saw the appearance of the mammals and your kind, the primates. I was hundred and sixty years old when the Alps, Andes and Himalayas were infants, striving to become giants on my earth. My numbers were many before the Pleistocene Age of Ice I, one million years ago. Your mind cannot comprehend my antiquity. Yet your mind, that which involvement set you above animals, give you an appreciation and even reverence for me and others of my kingdom. My strength is my antiquity and tenacity ... and you.

C. A. William

Florida, USA

Table of contents

Abbreviations

AFLP	Amplified Fragment Length Polymorphism
AP-PCR	Arbitrarily primed PCR
bp	Base pair
BgI I	*Bacillus grobigii I*
CI	Consistency index
CITES	Convention in International Trade in Endangered Species
cpDNA	Chloroplast DNA
CTAB	Hexadecyltrimethyl ammonium bromide
D	Genetic distance
DAF	DNA amplification fingerprinting
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dH_2O	Distilled water
EDTA	Ethylene diaminetetra acetic acid
ethanol	Ethyl alcohol
F	Coefficient of similarity
G	Gravitational force
g	Gram
$g.l^{-1}$	Gram per litre
HCl	Hydrochloric acid
HI	Homoplasy index
Hind III	*Haemophilus influenzae Rd III*
Hinf I	*Haemophilus influenzae RF 1*
ITS	Internal transcribed spacer
IUCN	International Union for the Conservation of Nature and Natural Resources
kb	Kilobase
M	Molar
MAAP	Multiple arbitrary amplicon profiling
$MgCl_2$	Magnesium chloride
ml	Millilitre
mm	Millimetre
mM	Millimolar
mg	Milligram
m/v	Mass per volume
NaCl	Sodium chloride
ng/ul	Nanogram per microlitre
nt	Nucleotide
OTU	Operational taxonomic units
OPA	Operon primer kit A
OPB	Operon primer kit B
OPC	Operon primer kit C

P	Product yield
PAUP	Phylogenetic analysis using parsimony
PCR	Polymerase chain reaction
pmol	Picomoles
$pmole.ul^{-1}$	Picomole per microlitre
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RI	Retention index
S.A.	South African
SNL	Signal-to-noise
TAE	Tris-acetic acid EDTA
Taq	*Thermus aquaticus*
TBR	Tree bisection and reconnection
TE	Tris-EDTA
tecMAA P	Template endonuclease cleavage MAAP
TPU	Threatened Plant Unit
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UOFS	University of the Orange Free State
UV	Ultraviolet
V	Volt
v/v	Volume per volume
u	Units
$\mu g.l^{-1}$	Microgram per litre
µl	Microlitre
°C	Degrees Celsius
%	Percentage

Summary

The extant cycads (Cycadophyta) are divided into three families, with 11 genera (Johnson 1959, Stevenson 1990), and at least 210 species (Stevenson *et al.* 1995). In this study a molecular systematic study was done on the genus *Stangeria* T. Moore from the family Stangeriaceae, and the genus *Encephalartos* Lehm. from the family Zamiaceae.

Fifty eight specimens representing 35 species of the genus *Encephalartos*, and one specimen from the genus *Stangeria*, *Stangeria eriopus* (Kunze) Nash were used. The effect and behaviour of hybrids in cladistic analysis were also investigated, such that they might be detectable during cladistic analysis.

A pilot study of two molecular techniques, RAPD and DAF was done to determine the applicability, time effectiveness and most suitable primers for a molecular systematic study using these techniques. All the RAPD primers tested gave a high coefficient of similarity and adequate number of characters per specimen. The speed and simplicity of the RAPD technique, combined with the findings of the pilot study, made this an ideal method to generate a large amount of fingerprinting data for a phylogenetic assessment of the genus *Encephalartos*.

Three variations of the DAF technique were tested in the pilot study i.e. simple DAFs, tecMAAP and multiplex DAFs. The coefficient of similarity was high in all the primers, with an increasing number of characters per specimen being generated. The multiplex DAF technique was the most successful, and most cost effective DAF variation. tecMAAP is also a very useful method in this study, but prior cleavage with restriction endonuclease, makes this a time consuming method. Multiplex DAFs combined with simple DAFs were, therefore, applied to more specimens in this study.

Phylogenetic relationships were determined using the computer software PAUP, and the genetic distances between species were also determined. The RAPD and DAF cladograms show some correspondence, but also some discrepancies as expected because of the low resolution in the cladograms. Combining the data sets increased the resolution of the cladogram. The molecular results concur with morphological and biochemical studies that have been done on the species in this study, as observed in the close relationship between *Encephalartos cupidus* and *E. eugene-maraisii*, and *E. umbeluziensis* and *E. villosus*. The RAPD and DAF techniques were also successfully applied in hybrid analysis. An increase in hybrids to the cladistic analysis, do influence the tree topology, but not the cladogram resolution, with an increased tree length. Hybrids, therefore, are no more problematic cladistically than the increased inclusion of specimens. In this study the molecular techniques RAPD and DAF have proven to be a fast and cost effective method in successfully generating fingerprints for phylogenetic analysis.

Keywords: DNA amplification fingerprinting (DAF), *Encephalartos*, genetic distances, phylogenetic relationships, random amplified polymorphic DNA (RAPD), *Stangeria eriopus*.

Opsomming

Broodboom spesies (Cycadophyta) van die wêreld is geklassifiseer in 3 families, met 11 genera (Johnson 1959, Stevenson 1990), en omtrent 210 spesies (Stevenson *et al.* 1995). 'n Molekulêre sistematiese studie is op die genus *Stangeria* van die familie Stangeriaceae, en die genus *Encephalartos* van die familie Zamiaceae gedoen.

Vyf-en-dertig *Encephalartos* spesies verteenwoordig deur 58 eksemplare, asook een *Stangeria eriopus* eksemplaar is in die studie gebruik om die filogenetiese verwantskappe van die spesies te bepaal. Die invloed van basters tydens kladistiese analises is ook bestudeer, met die hoop dat basters op die wyse geïdentifiseer kan word tydens analise. Die toepasbaarheid, tyd en mees informante inleiers van twee tegnieke, RAPD en DAF, is tydens 'n loodstudie bepaal.

RAPD tegniek is vinnig, goedkoop en 'n hoë koëffisient van ooreenstemming dui aan dat die tegniek gebruik kan word vir 'n filogenetiese studie van die genus *Encephalartos*.

Drie verskillend variasies van die DAF tegniek is getoets, nl. eenvoudige DAF, tecMAAP en multipleks DAF. Die inleiers wat getoets is, het 'n hoë koëffisient van ooreenstemming gehad en 'n toenemende hoeveelheid fragmente is waargeneem met die tegnieke. Alhoewel tecMAAP suksesvol in die loodstudie gebruik was, maak vooraf snyding met 'n beperkingsensiem die tegniek meer tydsaam. Multipleks DAF was die mees suksesvolle en koste effektiewe tegniek, en is saam met eenvoudige DAFs verder toegepas op meer eksemplare.

Filogenetiese verwantskappe van die eksemplare is bepaal met behulp van die rekenaar sagteware PAUP. Die genetiese afstand tussen spesies is ook bepaal. RAPD en DAF resultate stem grotendeels ooreen, met verhoogde resolusie van die kladogramme wanneer die RAPD en DAF data gekombineer is. Die resultate van hierdie studie stem ooreen met resultate verkry vanaf morfologiese en biochemiese studies soos byvoorbeeld die verwantskap tussen *Encephalartos cupidus* en *E. eugene-maraisii*, asook *E. umbeluziensis* en *E. villosus*.

Die invloed van basters tydens kladistiese analise het bewys dat basters nie die resolusie van kladogramme beïnvloed nie, maar wel die kladogram topologie. Kladogram lengte neem toe wanneer meer basters in die studie gebruik is. 'n Soort gelyke toename word egter ook gevind indien meer "tipiese" eksemplare gebruik word.

In die studie is die molekulêre tegnieke RAPD en DAF suksesvol toegepas om die genetiese afstand en filogenetiese verwantskappe tussen spesies van die genus *Encephalartos* te bepaal.

Sleutelwoorde: DNA amplifiserings vingerafdrukke (DAFs), *Encephalartos*, filogenetiese verwantskappe, genetiese afstande, lukraak geamplifiseerde polimorfiese DNA (RAPDs), *Stangeria eriopus*.

Acknowledgements

A special word of thanks to Prof. Spies for his guidance and patience throughout this study. I would also like to thank my co-supervisors Prof. Oberholster, for her useful suggestions and financial assistance, Dr. Vorster for his generous provision of plant material and guidance.

My gratitude to the University of the Orange Free State for their facilities and the Foundation for Research Development for financial support.

Thank you to Christ, for giving me the strength; to my family and friends, for your love and support; especially my parents who made all this possible.

In loving memory of my grandmother, Maria.

CHAPTER ONE

INTRODUCTION

1.1 History of cycads

Cycads are the most primitive living, seedbearing plants known (Loconte & Stevenson 1990). They have survived major climatic changes during the last 200 million years, and although small numbers are found, these 'living fossils' ability to adapt to changing environments has allowed them to survive to modern times (Whitelock 1995).

The first cycad recorded in Africa was *Encephalartos longifolius* (Jacq.) Lehm. in 1772 by Carl Peter Thunberg, accompanied by Francis Mason from Kew Gardens (Giddy 1989). By the end of the nineteenth century more than half of the species of South Africa known today had been discovered and described (Giddy 1989). The only real danger to cycads is Man, and by 1971, cycads had become a "status" plant desired by plant lovers, which forced Nature Conservation authorities to declare all cycads Specially Protected Plants[1]. For future conservation of cycads it is important to know as much as possible about the plants and this study is a contribution towards this goal.

1.2 General characteristics of cycads

Cycads are found in the tropical and subtropical regions of both hemispheres in Central America, Africa, Asia and Australia. The wide distribution of the cycad families, and especially the Zamiaceae, was probably caused by continental drift (Goode 1989).

Cycads are dioecious, which means that the male and female cones are

[1]http://www.ucmp.berkeley.edu/seedplants

on separate plants. The female cone bears the seed and consists of megasporophylls. The seeds of all the species, except *Cycas* subsection *Rumphiae*, sink in water, excluding water dispersal (Dehgan & Yuen 1983). The seeds of all the genera are covered with a fleshy outer coat (sarcotesta), that attracts animals and this serves as a dispersal method of the seeds (Giddy 1989). The inside kernels of most tested species {*Encephalartos cycadifolius* (Jacq.) Lehm., *E. friderici-guilielmi* Lehm., *E. eugene-maraisii* Verdoorn, *E. ferox* Bertol. f., *E. horridus* (Jacq.) Lehm., *E. lehmannii* Lehm., *E. longifolius* and *E. villosus* Lem.}, have proved to be toxic to at least some mammal species (Giddy 1989). The male cones produce the pollen. It was previously accepted that pollen was wind dispersed (Chamberlain 1935), but recent studies indicate that cycads are mostly insect pollinated (Donaldson 1995).

1.3 Origin and classification of cycads

The ancestor of the present-day cycads is extinct (Arnold 1953). All modern cycads are now separated by these 'missing links', and thus render phylogenetic studies difficult. According to Chamberlain (1920) cycads have not given rise to other plants, and will probably be extinct by the next geological period (Arnold 1953). Although cycads have developed characteristics of their own, they probably originated from Palaeozoic pteridosperms. Some of the features suggesting such an origin is: the motile male gametes, the structure of the ovules, the thick persistent cortex and the frond shaped leaves (Arnold 1953). Fossils recovered from the lower Permian of China (Zhifeng & Thomas 1989), classified as Cycadophyta, also indicate that the cycad group has risen from within the medullosan pteridosperms of the late Palaeozoic (Crane 1988). According to Nixon *et al.* (1994) the perception that cycads were the dominant plants of the Mesozoic era, is somewhat misplaced, because cycads were confused with the now extinct order Bennettitales.

The extant cycads (Cycadophyta) are divided into three families, with 11 genera (Johnson 1959, Stevenson 1990), and at least 210 species (Stevenson *et*

al. 1995). The families are Cycadaceae, Stangeriaceae and Zamiaceae. Johnson (1959) also recognised a fourth family, Boweniaceae. This study will focus on the families Stangeriaceae and Zamiaceae.

1.4 The family Stangeriaceae

Stangeriaceae is the smallest family of the cycads and has two genera, i.e. *Stangeria* T. Moore and *Bowenia* Hooker ex Hooker f. (Stevenson 1990). *Stangeria* has only one representative in Africa, *Stangeria eriopus*, (Kunze) Nash. This species is confined to the coastal areas of the eastern Cape and KwaZulu-Natal. Controversy regarding the recognition of one or more species of *Stangeria,* because of the morphological differences between plants found in the open grassveld and those found in forest areas (Dyer 1965b). These morphological differences could be attributed to adaptations to the different habitat conditions. Whether these adaptations are genotypic or phenotypic is unknown.

No fossil record of *S. eriopus* has been found and its relationship with other cycads is unknown (Giddy 1989). *Stangeria* shows by far the greatest number of other primitive and fern-like characters, yet in its habit and mode of growth it is unlike any known pteridosperm (Sporne 1967). If one were to look at the vegetative instead of reproductive characters for criteria of primitiveness among cycads, *Stangeria* would probably go to the base of the series (Arnold 1953). *Stangeria eriopus* was included in this study as an outgroup, to determine tree rooting points and character transformation. Outgroups are brought into an analysis to provide a broader phylogenetic context, and to aid in determining the root of the ingroup or ancestral states (Farris 1982, Maddison *et al.* 1984, Nixon & Carpenter 1993).

1.5 The family Zamiaceae

Zamiaceae is more widely distributed and consists of eight genera: *Chigua* Stevenson (South America), *Ceratozamia* Brongniart (Mexico,

Guatemala & Belize), *Dioon* Lindley (Mexico, Honduras & Nicaragua), *Lepidozamia* Regel (Australia), *Macrozamia* Miquel (Australia), *Microcycas* (Miquel) A.D.C. (Cuba), *Zamia* L. (South, Central & North America), and *Encephalartos* Lehm. (Africa) (Stevenson *et al.* 1995).

1.5.1 The genus *Encephalartos*

The genus *Encephalartos* has about 54 recognised species (Stevenson *et al.* 1995). The name *Encephalartos* is derived from the Greek word '*en*' meaning 'within', '*kephali*' meaning 'head' and '*artos*' meaning 'bread' (Giddy 1989). Appropriately, cycads are known as Bread trees (Afrikaans = 'broodbome'), referring to the use of the pith from stems of *Encephalartos* species for making crude bread by the local tribes in southern Africa (Dyer 1965b). *Encephalartos* (Figure 1.1) species have adapted to their environment and occupy a wide variety of habitats, from dry areas with harsh winters, to high rainfall and frost-free areas. In southern Africa the genus *Encephalartos* is found in the eastern half of the country, on the cool mountain slopes of the Drakensberg, the humid coastal forests of KwaZulu-Natal, and the rocky areas of the Karoo and Mpumalanga (Goode 1989).

1.5.1.1 Previous studies of the genus *Encephalartos*

Thorough morphological (Dyer 1965b, Whitelock 1986, Osborne 1988), taxonomic (Dyer & Verdoorn 1966), cytological (Norstog 1980), anatomical (Koelemann *et al.* 1981), tissue culture (Koeleman & Small 1982), propagation and biochemical studies (Tustin 1974, Osborne *et al.* 1988, Van der Bank *et al.* 1998) have been done on cycads, but very few molecular studies. Cytogenetic studies of somatic cells from *Encephalartos* show a chromosome number of 2n = 18 (Marchant 1968, Mogford 1979).

Stangeria has a chromosome number of 2n = 16 (Sporne 1967). No sex chromosomes have been identified in cycads (Mehra 1986). Cytogenetic studies could help shed light on the peculiar gender changes that have been reported in several specimens (Osborne 1985), including an *E. umbeluziensis* R.A. Dyer specimen in a Pretoria garden (Van Wyk & Claassen 1981).

Figure 1.1 Illustrations of the genus *Encephalartos* in their natural environment. A – *E. paucidentatus* Stapf & Burtt Davy with a stem of 6 meters in height. B – *E. natalensis* Dyer & Verdoorn. C – female cones of *E. natalensis*. D – *E. middelburgensis* Vorster.

Norstog (1980) speculated that adaptations of the ancestral cycad stock to harsher environments, represent advances. According to this hypothesis *E. altensteinii* Lehm., *E. lebomboensis* Verdoorn, *E. natalensis* and *E. transvenosus* Stapf & Burtt Davy are classified as primitive species. Species that are considered more advanced are those with reduced leaflets (*E. cycadifolius*), those with subterranean caudices (*E. villosus*) and with highly armed leaves in arid regions (*E. horridus*) (Osborne *et al.* 1988). This observation is, however, greatly influenced by the characteristics observed and cannot be used as an overall ranking of species.

Koeleman *et al.* (1981) attempted to establish an identification method within *Encephalartos* by studying the anatomy of the pinnae of some *Encephalartos* species. They found a correlation between anatomical characteristics and the distribution of the different species of *Encephalartos*. They also suggested two developing lines, which could have evolved from a hypothetical mutual ancestor. The first 'line' ends in the *E. humilis* Verdoorn group, which have narrow pinnae suited for a colder and humid climate, distributed in high-laying sour grassvelds. The second 'line' varies from the mesophyte group with true dorsiventral pinnae distributed in warmer, humid areas, to the smaller group of amphistomatic isobilateral pinnae such as *E. eugene-maraisii* and *E. cupidus* R.A. Dyer.

Vorster (1986) grouped the *Encephalartos* species according to the external morphology of the fronds and cones into the following groups:

- *E. horridus, E. lehmannii, E. trispinosus* (Hook) R.A. Dyer, perhaps *E. arenarius* R.A. Dyer and *E. latifrons* Lehm.,
- *E. altensteinii, E. lebomboensis, E. natalensis* and *E. woodii* Sander,
- *E. villosus, E. umbeluziensis,* perhaps *E. hildebrandtii* A. Braun & Bouché,
- *E. eugene-maraisii* and *E. cupidus,*

- *E. paucidentatus* and *E. transvenosus,*
- *E. chimanimaniensis* R.A. Dyer & Verdoorn, *E. concinnus* R.A. Dyer & Verdoorn, *E. gratus* Prain, *E. manikensis* Gilliland (Gilliland) and *E. pterogonus* R.A. Dyer & Verdoorn,
- *E. cycadifolius* and *E. ghellinckii* Lem.,
- *E. humilis, E. laevifolius* Stapf & Burtt Davy and *E. lanatus* Stapf & Burtt Davy.

Species with uncertain affinities are *E. caffer* (Thunb.) Lehm., *E. ferox* Bertol. f., *E. heenanii* R.A. Dyer, *E. inopinus* R.A. Dyer, *E. munchii* R.A. Dyer & Verdoorn, *E. ngoyanus* Verdoorn and *E. princeps* R.A. Dyer.

Phylogenetic relationships, based on allozyme data, indicate that the genetic distance between *E. altensteinii* and *E. natalensis* is very small (Van der Bank *et al.* 1998). *Encephalartos villosus* is a sister to a group that includes *E. lehmannii*, *E. altensteinii* and *E. natalensis* (Van der Bank *et al.* 1998).

1.5.1.2 Hybridisation in the genus *Encephalartos*

Due to intraspecific variability, it is sometimes difficult to identify isolated cycad specimens. It is important to distinguish whether these differences were caused by adaptation to climatic or geographical factors, or if the cycad could be a hybrid between two closely related species.

Chamberlain (1919) believed, in contrast to Henderson (1945), that hybridisation in the genus *Encephalartos* does take place in nature (quoted by Dyer 1965b). According to Vorster (1986) and Norstog (1990), evidence suggests a weak fertility barrier between cycad species and, therefore, natural hybridisation could take place where different species come in contact. An example of the cone of a natural hybrid is shown is in Figure 1.2. Artificial hybrids have been produced (Osborne *et al.* 1988), and a number of natural *Encephalartos* hybrids have been recorded (Vorster 1986), but geographical isolation and different coning times make natural hybrid specimens rare.

Natural hybridisation occurs between sympatric species or subspecies of a plant where ecological conditions are permissive (Grant 1981). Reproduction of a natural hybrid (in a cross-breeding system), follows the pathway of back-crossing or introgressive hybridisation (Grant 1981). Introgression is the incorporation of genes of one species into the gene pool of another species by hybridisation and backcrossing (Anderson & Hubricht 1938). In the Buffalo River Valley, Bathurst and Bushmans River areas, specimens have been found which suggest that hybridisation took place between *E. altensteinii* and *E. villosus* (Dyer 1965b). Whether these hybrids are sterile or if introgressive hybridisation took place, has not been recorded. Interspecific hybridisation in plants has been the focus of many studies (Anderson & Hubricht 1938). An unknown origin, due to interspecific hybridisation, makes classification arduous, and this type of hybridisation is a potential threat for some parental species as distinct breeding groups (Dyer 1965b). Interspecific hybridisation could have the greatest impact on small or rare populations, such as *E. woodii*.

There are only a few *E. woodii* trees left in the world (Figure 1.3). These are offsets from the last remaining male tree that was discovered by Wood in 1895, near Ngoye in Zulu- land. The reason for the rarity of this plant is unknown, but several theories exist. The species could have been depleted through the use in Zulu ceremonies, or some suggest that the plant could be a natural hybrid, representing the only one of its kind (Osborne 1986).

1.5.1.3 The necessity for a molecular systematic study of the genus *Encephalartos*

The Threatened Plant Unit (TPU) of the International Union for the Conservation of Nature and Natural Resources (IUCN) lists three *Encephalartos* species as endangered: *E. cupidus, E. inopinus* and *E. latifrons*; five as rare: *E. altensteinii, E. friderici-guilielmi, E. lehmannii, E. natalensis* and *E. transvenosus*. *Encephalartos woodii* is classified as extinct (Gilbert 1984). Cycads' habitat is being threatend by agriculture and human

Figure 1.2 The cone of a putative hybrid female specimen, thought to be a cross between *E. trispinosus* and *E. arenarius.*

Figure 1.3 A specimen of *E. woodii* in the Durban Botanical Gardens.

development. Unscrupulous collectors and syndicates remove mature specimens to private gardens, where they are unlikely to reproduce (Osborne *et al.* 1988). Cycads are protected locally by legislation that is enforced by the Endangered Species Protection Unit of the S.A. Police Services. Internationally, cycads are protected by the CITES convention, which has banned all trade of endangered species (CITES Secretariat 1982).

A microchip tracking device has recently been used to assist the police in tracing any illegal removal and retrieval of cycads[2]. A scanner picks up the microchip embedded in the cycad's stem, which gives the cycad's identity number. The identity number is used to retrieve the data about the cycad, such as the exact location, sex and plant's diameter. Despite many conservation strategies, natural populations of cycads are becoming extinct due to drastic human exploitation (Johnson & Wilson 1990).

A reliable identification system could also assist the police and nature conservationists in cases where cycads are illegally removed from nature. A combination of morphological and molecular research could help in the identification of specimens. In this study, the application of molecular systematics were tested to aid in further understanding of the genus, which could help in these issues.

1.6 Molecular systematics

Systematic research recognises and describes species and establishes the evolutionary relationship between them[3]. According to Hennig (1966), phylogenetic systematics investigates the phylogenetic relationship between existing species and presents the results in a clear form, such as a phylogenetic cladogram.

Traditionally, morphological data supported by anatomical data, ultrastructure, similarities with respect to secondary metabolites, isozymes and

[2]http://www.wildnetafrica.com
[3]http://cycad.org

other protein systems, were used to determine the relationship between plant groups (Clegg & Durbin 1990). In the last few years, molecular biology is increasingly being employed in systematic research to study genetic relationships and phylogeny (Clegg & Durbin 1990). New and refined molecular genetic techniques provide exciting discoveries on a regular basis that help improve systematic analysis (Sokal 1994).

Since there is a variety of choices to be made when initiating a molecular systematic study, it is important to determine the most appropriate genome or gene, as well as the molecular method that will be used. Each approach has its own advantages, disadvantages and technical difficulty (Clegg & Durbin 1990).

1.6.1 Techniques that could be used in a molecular systematic study

Amongst the many techniques that could be used to determine the phylogenetic relationships of *Encephalartos* species are:

1 Heun *et al.* (1994) used isozyme data for analysing the genetic relationships among *Avena sterilis* L. accessions. They concluded that isozyme analyses could be used to distinguish between *A. sterilis* accessions, but that DNA-based markers will be more accurate in determining the relationship between species that are too closely related to be accurately differentiated by isozymes.

2 Allozymes analysis can be used to study genetic variation in unambiguously identified species and hybrids (Rajora 1990). This is a relatively fast and inexpensive method, but proven ineffective in some plant groups due to an insufficient number of markers for genome analysis (Liu & Furnier 1993). The method has been applied successfully to *Encephalartos transvenosus*, *E. villosus*, *E. friderici-guilielmi*, *E. natalensis*, *E. altensteinii* and *E. lehmannii* (Van der Bank *et al.* 1998).

3 Restriction fragment length polymorphisms (RFLPs) show differences at DNA level in coding and non-coding regions. This method is based on differences in fragment lengths obtained by restriction enzyme digestion. The resulting fragments are separated by gel electrophoresis, blotted onto a membrane and then probed (Karp *et al.* 1996). The differences (polymorphisms) obtained, are due to the presence or absence of restriction sites in the genomes being compared (Yu *et al.* 1993). This method is useful for providing large numbers of genetic markers, and can be used to differentiate between species and also between individuals within species (Liu & Furnier 1993). Detection of RFLPs is often laborious, time-consuming and expensive (Yu *et al.* 1993). The method has been used in the phylogenetic relationship analysis of the American Zamiaceae (Cycadales) (Caputo *et al.* 1991), as well as a phylogenetic analysis of *Dioon* (Zamiaceae) (Moretti *et al.* 1993). RFLP analysis of chloroplast DNA has also been applied to the molecular systematics of cycads (De Luca *et al.* 1995).

4 Amplified fragment polymorphism (AFLP) is based on the selective amplification of restriction fragments from, enzyme digested, genomic DNA (Lin & Kuo 1995). Although this technique is fast and generates high-density genetic maps, it is much more expensive than RAPDs (Yu *et al.* 1993).

5 In recent years the most relevant method for molecular phylogenetic analysis has been sequencing (Sang *et al.* 1994, Suh *et al.* 1993, Wolciechowski *et al.* 1993). A popular region to sequencing is the ITS (internal transcribed spacer) region containing two internal transcribed spacers (ITS-1 and ITS-2) and 5.8S rRNA (Baldwin 1992, Baldwin 1993, Baldwin *et al.* 1995). The relative high rate of nucleotide substitutions in the transcribed spacers, allow for systematic comparison of recently diverged taxa (Baldwin *et al.* 1995). Liston *et al.* (1996) studied the approximated ITS region size in non-flowering seed plants, which

included *Stangeria eriopus* (1450 base pairs). Sequencing of the ITS region were unsuccessfully applied in this study.

6 Multiple Arbitrary Amplicon Profiling, MAAP (Caetano-Anollés *et al.* 1992a, 1992c, 1993, 1994) involves the use of a short arbitrary chosen oligonucleotide primer that anneals to DNA. This technique is comprehensive, so that a primer could be used repeatedly for other species, even if the evolutionary distances between the species are large (Gresshoff 1993). When the reaction conditions are optimised for a cultivar, amplification conditions can be used to compare several cultivars of the same species or genus. Multiple regions of the genome are then amplified, and the fragments are called amplicons (Mullis 1991). MAAP procedures were developed independently in three laboratories:

- Welsh & McClelland (1990), developed AP-PCR, which uses primers (18 to 32 nucleotides) of arbitrary sequence to amplify target DNA under low stringency annealing conditions for two amplification cycles, and further amplification at normal temperature.

- Williams *et al.* (1990) invented the random amplified polymorphic DNA analysis (RAPD) procedure in which arbitrary primers of either 9 or 10 nucleotides produce amplification products after temperature cycling.

- Caetano-Anollés *et al.* (1991) developed DNA amplification fingerprinting (DAF), and is based on the use of arbitrary oligodeoxyribonucleotide primers to amplify anonymous genomes, and generate diagnostic fingerprints. DAF utilises the shortest primers of which the optimal length was found to be 8 nucleotides (Caetano-Anollés *et al.* 1993).

The techniques that were used in this study were RAPDs and variations of the DAF technique. These techniques have been successfully applied to

many studies to determine phylogenetic relationships between and within species (Bassam *et al.* 1992, Caetano-Anollés *et al.* 1993, Chalmers *et al.* 1992, Demeke *et al.* 1992, Jayarao *et al.* 1992, Oxelman 1996).

Many methods have also been used to investigate hybridisation in plants, including artificial hybridisation (Byrne & Morley 1976, Roelefs 1979), chromosomal variations (Heiser 1949, Pipkin 1972), and isozymes (Dickson & Weeden 1991, Werth 1991). More recently DNA has been used as a basis for investigations of hybridisation in plants (Crawford *et al.* 1993, Cruzan & Arnold 1993). In this study true hybrids were used to analysis the influence of hybrids in phylogenetic studies, as discussed in 1.5.1.2. No publications have yet been found where these techniques (RAPDs and DAFs) have been applied to phylogenetic analysis of the genus *Encephalartos*.

1.6.2 Random amplified polymorphic DNA analysis

Random amplified polymorphic DNA analysis is the detection of DNA polymorphism through the amplification of random DNA segments by single primers of arbitrary nucleotide sequence (Williams *et al.* 1990). RAPDs are based on the polymerase chain reaction, in which single-stranded DNA is used as a template for synthesis of a complementary new strand (Williams *et al.* 1990). A RAPD-PCR reaction consists of template DNA, *Taq* Polymerase, dNTPs, magnesium chloride, primer and reaction buffer. To ensure optimal performance of the PCR reaction, the magnesium chloride, dNTPs, primer and template DNA, must be used in optimum concentration combinations.

The reaction mixture is placed in a thermal cycler where it undergoes 35 – 40 cycles of three basic steps: (1) denaturation, (2) annealing, and (3) amplification (Watson *et al.* 1992). During the denaturation step the double-stranded DNA molecules are separated and form single-strand templates. The temperature is then lowered for the annealing step. At the annealing temperature the primer molecules anneal to their complementary sequence on the single-strand DNA. In the next step the temperature is raised for the amplification step, during which the complimentary DNA fragment is

synthesised. A final temperature raise follows, during which the now double stranded DNA is again separated to form the template for the next cycle.

The amplification product is separated on a gel medium and the fragments scored to create a data matrix. The data matrix is analysed to determine genetic variation and polymorphisms. The polymorphisms observed may have originated from point mutations, insertions, deletions and inversions and some of these polymorphisms can be used as markers (Williams *et al.* 1990). Although most of these markers are inherited in a simple Mendelian fashion (Demeke *et al.* 1992), they are always dominant (Liu & Furnier 1993, Ragot & Hoisington 1993). According to Liu & Furnier (1993), a problem with RAPDs is the assumption that each fragment represents one RAPD locus with only two alleles, which corresponds to the presence or absence of a fragment observed on the gel. This is not always true, since deletion and insertion could occur in the region between the primers and this will result in an amplified fragment migration to a different position. The fragment will then be scored as a separate locus rather than an additional allele at the first locus (Liu & Furnier 1993). Another problem that may arise is that different RAPD fragments may have similar molecular weights and therefore co-migrate on gels, and they may not necessarily be homologous (Backeljau *et al.* 1995).

There are some concerns over the validity of this technique in certain applications, such as systematics, where characters are generally assumed to be independent and homologous (Swofford & Olsen 1990). Studies comparing RAPD based phylogenies with other techniques, such as RFLP analysis of cpDNA, sometimes show conflicting results (Gilles & Abbott 1998). Harris (1995) used RAPD data to represent the phylogeny of the genus *Leucaena* and concluded that the technique did not provide enough resolution. Thormann *et al.* (1994) found that RFLPs are more reliable than RAPD data in the phylogenetic analysis of *Brassica*. But Chalmers *et al.* (1992) and Demeke *et al.* (1992) found that phylogenies derived from RAPD data are consistent with those produced by other means. The RAPD technique has also successfully been applied to phylogenetic analysis by Heun *et al.* (1994), Oxelman (1996), Ramser *et al.*

(1997) and Gilles & Abbott (1998). Gilles & Abbott (1998) concluded that, although it is unwise to use RAPDs for phylogenetic analysis between distantly related taxa, the technique can be used in assessing relationships at lower taxonomic levels. By using this technique in conjunction with techniques such as Southern analysis, sequencing, and endonuclease digestion of amplification products (Backeljau *et al.* 1995), problems concerning the reliability of RAPD data can be overcome.

Keeping these problems in mind, the RAPD procedure has been used successfully for genome mapping, gene tagging, and related studies (Goodwin & Annis 1991). RAPD analysis can also be used for reconstruction of phylogenies (Wilkie *et al.* 1993, Borowsky *et al.* 1995, Hoey *et al.* 1996, Såstad *et al.* 1999), as well as analyses of genetic variation in more than one species (Chalmers *et al.* 1992, Demeke *et al.* 1992). The technique is not widely accepted in the construction of phylogenies (Landry & Lapointe 1996). Borowsky *et al.* (1995) suggested that, although RAPD markers contain phylogenetic information, the phylogenies are not easily derived from the markers and the results obtained are often difficult to interpret. Demeke *et al.* (1992), as well as Landry & Lapointe (1996) found that a stable phylogeny could be obtained with an increase of the number of primers. The RAPD method is used in phylogenetic analysis because of the possibility to obtain relatively quickly and inexpensively, large samples of polymorphisms from sites all over the genome (Oxelman 1996). The major advantages of the RAPD analysis approach are that prior DNA sequence information is not required and amplification procedures are simple to perform (Williams *et al.* 1990).

The reproducibility in RAPDs was investigated by Penner *et al.* (1993), using the same DNA and primers in different laboratories. They found that most RAPD markers were reproducible, with differences between PCR machines accounting for most of the variations observed (Penner *et al.* 1993). In studies where the RAPD technique was compared to other analytic techniques it was concluded that RAPDs generated data faster, have a protocol that require less DNA and no radioactivity, and the problem of reliability can

be eliminated by optimising the experimental conditions and by following precisely a chosen experimental protocol (Williams *et al.* 1990, Heun & Helentjaris 1993).

1.6.3 DNA amplification fingerprinting

DNA amplification fingerprinting (DAF) was originally described by Caetano-Anollés *et al.* (1991). DAF generates characteristic signatures from DNA. These signatures are the result from the amplification of multiple anonymous sites from template DNA or RNA molecules, using oligodeoxynucleotides. The amplification produces arbitrary, but entirely characteristic signatures or fingerprint patterns. DNA can now be scanned for polymorphisms without prior knowledge of the template (Caetano-Anollés *et al.* 1991).

The DAF protocol consists of template amplification, followed by the separation and visualisation of the amplification product (Caetano-Anollés 1997). The amplification of the template is determined by complex kinetic and thermodynamic processes, and influenced by the interaction between the primer, template annealing sites, and enzyme used in the amplification reaction (Caetano-Anollés *et al.* 1992b). During optimisation a 'reproducibility window' is necessary in which the amplification parameters exhibit little or no variation (Bassam & Bentley 1994). Although optimum conditions are seldom identified, this laborious task can be simplified by determining the most important reaction components, using a modified Taguchi method (Cobb & Clarkson 1994). DNA amplification is influenced by the concentration of the primer, magnesium chloride, deoxynucleotide triphosphate and template (Caetano-Anollés 1997).

Polymorphisms result from changes in the DNA sequence that were targeted by the primer, as well as from deletions, insertions or inversion of the priming site or segments between the priming sites. Polymorphisms could also arise from conformational changes in the DNA molecule that would alter the efficiency of amplification or annealing of primer to the specific genomic

site (Caetano-Anollés 1997). These molecular polymorphisms generaerated, are influenced by the variation in primer sites on the target DNA, length variation between the primer sites, and possible changes in the secondary structure of the target DNA between or flanking the primer recognition sites (Gresshoff 1993).

DAFs can be tailored to increase the amount of polymorphisms obtained. Tailoring of DAF reactions can increase the number of sites being probed by changing primer design, endonuclease restriction or post-amplification manipulation, as well as changing the kinetics of the reaction by changing the stringency of amplification and primer-template interaction (Caetano-Anollés *et al.* 1996). Primer design can be improved by designing the sequence to anneal to a specific site on the genome, or primers with secondary structure called mini-hairpin primers (Caetano-Anollés & Gresshoff 1994a). Cleavage of template with restriction endonuclease prior to amplification, also known as tecMAAP (template endonuclease cleavage MAAP), has been used for the identification of near isogenic soybean lines (Caetano-Anollés *et al.* 1993). In tecMAAP the template DNA length is reduced and it eliminates possible priming sites (Caetano-Anollés *et al.* 1993). This strategy increases the generaeration of polymorphic DNA two to four-fold (Caetano-Anollés 1994). Digestion of the template could result in the differential destruction of amplicons and selective amplification of those products that lack internal restriction sites (Caetano-Anollés 1994).

Another method of increasing polymorphisms is the use of arbitrary primers in pairwise combinations, called multiplex DAFs (Caetano-Anollés *et al.* 1991, Callahan *et al.* 1993, Micheli *et al.* 1993). With this method the amplification product is not the combination of one fingerprint pattern with another, but a new fingerprint pattern is formed. Certain fragments disappear and new ones are generaerated with few being shared (Callahan *et al.* 1993). New fragments could arise due to the overlapping of the extension products started by each primer, and fragments could disappear because one primer annealing site is located between two other primer annealing sites. Competition for annealing sites during amplification can cause the generation

of new fingerprint patterns (Callahan *et al.* 1993).

The detected polymorphic DNA can be used as molecular markers in genetic mapping; or character loci in population biology, systematics, phylogenetic, or pedigree analysis (Caetano-Anollés 1994, Schierwater 1995). DAF has been used successfully for legal proof of genetic individualism between turfgrass species and cultivars (Callahan *et al.* 1993). The method has also been applied in phytoforensics, in identifying bermudagrass plant material based on unique reference profiles generateraed with selected primers (Gresshoff 1996). DAF has successfully been applied to breeding, identification, and phylogenetic analysis of plants (Caetano-Anollés *et al.* 1993, Baum *et al.* 1994, Caetano-Anollés & Gresshoff 1994a, Weaver *et al.* 1995). DNA fingerprinting using DAF was proven accurate, repeatable and not random (Callahan *et al.* 1993, Gresshoff & MacKenzie 1994).

Artifacts are a potential problem in the analysis of genetic variation, especially when the mode of inheritance of DNA polymorphisms is unknown. It is, therefore, important that a reliable molecular technique is used to generate data for relationship analysis. According to Caetano-Anollés (1994), DAF profiles are produced with minimal experimental variability and appear free of artifactual fragments. The DAF technique have many applications and have successfully been applied to determining phylogenetic relationships (Jayarao *et al.* 1992, Kaemmer *et al.* 1992, Baum *et al.* 1994, Caetano-Anollés *et al.* 1995). The consistency of the DAF technique for amplification of DNA as shown by Caetano-Anollés (1994), Caetano-Anollés & Gresshoff (1994b) and Gresshoff & McKenzie (1994) makes this a reliable technique for analysis of phylogenetic relationships. DAF consistency of fragments have been proven for DNA amplification of prokaryotic (Bassam *et al.* 1992, Jayarao *et al.* 1992) and eukyotic organisms (Caetano-Anollés 1994, Caetano-Anollés & Gresshoff 1994b).

Compared to RFLPs, DAFs are fast, easy to perform, need very small amounts of DNA, and use no radioactivity (Callahan *et al.* 1993). DAF generates relatively complex amplification profiles (Caetano-Anollés *et al.* 1991) in contrast to the simple profiles produced by RAPDs (Williams *et al.*

1990). Generally DAFs are separated from other scanning techniques by the high primer-to-template ratios, as well as the excellent reproducibility and high multiplex ratios (Caetano-Anollés 1997).

1.6.4 Phylogenetic analysis

Based on incomplete information, co-existing species are used to determine the 'best estimate' of an evolutionary history (Swofford & Olsen 1990). To prevent misinterpretation, the results are presented as a phylogenetic cladogram. The main aim in phylogenetic reconstruction is to locate sister species or sister groups by using the unique, derived features inherited by members from the immediate ancestor of the group, called synapomorphies (Bremer & Wanntorp 1978). Similarity is an important criterion for determining the relationship of species, and recently shared homologies (synapomorphies) are evidence that two organisms are closely related (Lipscomb 1998).

There are many methods of determining phylogeny, and to conclude phylogenetic relationships from molecular data requires the selection of an appropriate analytical method (Swofford & Olsen 1990). Evolutionary systematists infer a phylogeny from a classification system or from the databases of the system, whereas cladists' first produces a cladogram with the help of computer software and from the cladogram, systematic conclusions are derived (Grant 1998).

A data matrix is created, consisting of specific characters representing the molecular data. Characters are classified as qualitative (in which the possible states are two or more discrete values) or quantitative (in which characters vary continuously and are measured on an interval scale) (Swofford & Olsen 1990). Qualitative characters can be further subdivided into binary (two possible states) or multistate (three or more states). In this study binary characters were used, which typically represent the presence or absence of a fragment. In most character-based analyses, independence of characters is assumed. If covariance among characters were taken into account, the

computational methods would become vastly more complicated (Swofford & Olsen 1990). They made a second assumption that characters must be homologous. This means that the states observed in all the taxa for a particular character must be derived, perhaps with modifications, from a corresponding state observed in the common ancestor of the taxa.

A cladogram is a branching diagram, in which the sequence of branching points is based on the pattern of distribution of synapomorphous characters in the group being investigated so that a nested series of sister groups is established (Brothers 1978). There are two ways of constructing a cladogram, the Hennig argumentation, as described by Hennig (1966); and the Wagner method that was used in this study (Lipscomb 1998). Wagner trees are constructed by adding one taxon at a time, and the joining of the taxon to the tree must be in such a way that there will be a minimum number of character state changes (Lipscomb 1998). Cladograms do not hypothesise ancestor-descendant relationships, but are relative statements of relationships between specimens used in the study. Certain terms are used to describe how much homoplasy was required to construct the cladogram.

Homoplasies are 'extra steps' required to explain the same character state in two or more taxa due to inheritance from the ancestor and are caused by reversals, parallelisms and convergences (Swofford 1993). If no homoplasy occurred in the data, the taxa would be grouped in a phylogeny according to the shared derived character states (synapomorphies) (Swofford 1998). The tree length or steps are the number of character state changes necessary to support the relationship of a specimen in a cladogram (Lipscomb 1998). This means that a shorter tree length represents less homoplasy and fewer character state changes. The consistency index (CI) measures the homoplasy as a fraction of the character changes on a cladogram (Farris 1989). It is also inversely proportional to the length of a cladogram, which means that the CI decreases with increasing number of taxa (Farris 1989). The retention index (RI) reflects the degree to which similarities in the data can be retained as homologies on a cladogram and it is not influenced by the inclusion of autapomorphies in data (Farris 1989).

Constructing a cladogram so that the number of changes from one character state to the next is minimised, as applied in this study, is called parsimony (Lipscomb 1998). Finding the most parsimonious cladogram becomes more arduous as datasets and character conflicts increase. There are different methods of searching for the most parsimonious cladogram.

- An exhaustive search is the preferred method because it analyses every possible cladogram, but this method is only suitable for small data sets.

- A Branch-and-bound search, seeks cladograms that are likely to be the shortest, but this is a time consuming method.

- In heuristic searches the branches of a cladogram are rearranged to search for a shorter topology (Lipscomb 1998). However, some data sets yield equally parsimonious cladograms, which means that some parts of the phylogenetic analysis cannot be resolved.

A consensus cladogram can be constructed representing the information which all the parsimonious cladograms have in common. Consensus techniques have been designed to handle problems with different data sets, but not the problem of multiple cladograms for a single data set (Adams 1972, Carpenter 1988). This means that the consensus technique is appropriate for different data sets but not for cladograms from a single data set, because consensus cladograms are highly unresolved (Barrett *et al.* 1991) and will sometimes require more character state changes than any of the separated cladograms (Miyamoto 1985). The consensus cladogram may, therefore, be a misleading guide to patterns of character evolution (Miyamoto 1985), and can contradict the most parsimonious cladogram obtained from the pooled data (Barrett *et al.* 1991). It is, therefore, important to establish whether the consensus cladogram is consistent with the best cladogram based on the pooled data (Barrett *et al.* 1991). There are a number of consensus techniques such as: Strict (Sokal & Rohlf 1981), Adams (Adams 1972), majority rule (Margush & McMorris 1981), and combinable component or semistrict consensus (Bremer 1990). A Strict consensus cladogram is constructed by

combining only those components that appear in all the parsimonious cladograms (Lipscomb 1998). According to Barrett *et al.* (1991) Strict is most commonly used as it is the most conservative. Although Adams cladograms can contain components not present in the most parsimonious cladograms, Funk (1985) and Hillis (1987) found that these cladograms could be use for pinpointing taxa responsible for incongruence, because they place conflicting taxa at the internal node (Adams 1972). An internal node is the branching point in a tree (Lipscomb 1998). Semistrict or combinable component consensus cladograms are similar to a Strict consensus but will include clades that are not contradicted by all the cladograms (Bremer 1990). In majority rule consensus a taxon is placed were it is most frequently found in the most parsimonious cladograms (Swofford 1991).

Another technique that can be used when choosing a cladogram from a number of equally parsimonious cladograms, is successive character weighting. Successive character weighting is the weighting of characters differentially according to their degree of correlation with cladistic relationships (Farris 1969). This is based on cladistic reliability, which means the degree of fit between a character and the phylogeny (Farris 1969). The unit character consistencies provide measures of cladistic reliability, and this can then be used to weight characters. The reweighted characters are used to construct a new estimated cladogram, the process is repeated until two successive cladograms have the same form (Farris 1969). A problem with successive weighting is that different initial weights may lead to different final solutions (Lipscomb 1998).

Some of the most generally used computer software programs to determine cladograms are:

- Hennig86 (Farris 1988), a fast parsimony program that uses branch-and-bound search for the most parsimonious cladograms;
- Random cladistics (version 2.2.1, Siddall 1994), a program that can carry out bootstrapping, jacknifing (Lanyon 1985) and search for "island" trees, using Hennig86 to analyse the data;

- Phylogeny inference package (PHYLIP version 3.5) (Felsenstein 1991), consists of programs for molecular sequence data, distance matrix data (NEIGHBOR), gene frequencies and continuous characters (GENDIST), 0-1 discrete state data and programs for plotting trees and consensus trees (CONSENSE, RETREE); and

- Phylogenetic analysis using parsimony (PAUP* version 4.0 beta) (Swofford 1998), a computer program for inferring phylogenies from discrete character data under the principle of maximum parsimony (Swofford 1998), as used in this study.

Another method of analysing data generated from molecular techniques is calculating the similarity or distance between two specimens (Swofford & Olsen 1990). In this study the index of genetic similarities (F) proposed by Nei & Li (1979), was used to calculate pairwise genetic distances (D) for all the species. Genetic distance is a quantitative measure of genetic relationship between two individuals expressed as a single number (Smith 1977). An F-value of one illustrates full similarity, and a lower values illustrate less similarity between two individuals (Weir 1996).

1.7 Aim of study

The aim of this investigation is to use molecular techniques, i.e. random amplified polymorphic DNA analysis and DNA amplification fingerprinting, to study phylogenetic relationships of species in the genus *Encephalartos*. Although more successful techniques, such as sequencing, has been applied to phylogenetic analysis of many different species, RAPDs and DAFs were used in this study, as a fast and cost effective method to create a phylogenetic hypothesis of the genus *Encephalartos*, based on molecular data. A few hybrids were included in the study as a preliminary analysis of the impact of hybrids on cladistic analysis, because hybrids can cause character conflict in cladogram construction and, therefore, affect the reconstruction of phylogenies.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Materials

Fifty-eight specimens, representing 35 *Encephalartos* species and one *Stangeria eriopus* specimen (included as an outgroup), were analysed in the study (Table 2.1). Several more species were excluded because DNA could not successfully be extracted from them. The hybrids included in this study is true hybrids.

Super-Therm Taq-Polymerase from Southern Cross Biotechnologies and OPERON primers were used. DAF primers, DNA molecular weight marker VI (pBR328 DNA cleaved with *Bgl* I and *Hinf* I) and restriction endonuclease *Hind* III were from Boehringer Mannheim. All the chemicals used in the study were of analytical grade.

Table 2.1 A list of species with the specimens used during this study.

Species	Voucher number
Encephalartos aemulans Vorster	*Vorster 229*[1]
E. altensteinii Lehm.	*FS 01, PRE 06*
E. aplanatus Vorster	*Vorster 422, Vorster 312*
E. arenarius R.A. Dyer	*PRE 08, Vorster 998*[2]
E. bubalinus Melville	*Vorster 978*
E. caffer (Thunb.) Lehm.	*Vorster 289*
E. cf.*chimanimaniensis* R.A. Dyer & Verdoorn	*Vorster 160*
E. cupidus R.A. Dyer	*SBG 01*

[1] *Vorster 229* (*Encephalartos aemulans*) has been donated to Kirstenbosch.
[2] *Vorster 998* represents the so-called 'blue *E. arenarius*', a very glaucous variant which may turn out to have closer affinities to *E. horridus* than *E. arenarius*.

E. eugene-maraisii Verdoorn	*PRE 21, Vorster 268*
E. ferox Bertol. f.	*Vorster s.n., Vorster s.n.*[3]
E. friderici-guilielmi Lehm.	*SBG 02*
E. ghellinckii Lem.	*Vorster 322*[4]
E. gratus Prain	*Vorster 834*
E. heenanii R.A. Dyer	*Vorster 622*
E. hildebrandtii A. Braun & Bouché	*Vorster 507*
E. horridus (Jacq.) Lehm.	*PRE 12, Vorster 362*
E. humilis Verdoorn	*PRE 15*
E. inopinus R.A. Dyer	*PRE 05*
E. laevifolius Stapf & Burtt Davy	*PRE 19, Vorster 260, Vorster 258*[5]
E. lanatus Stapf & Burtt Davy	*PRE 17, Vorster s.n.*
E. laurentianus De Wildeman	*Vorster1000*
E. lebomboensis Verdoorn	*PRE 01, PRE 14, Vorster 313*
E. lehmannii Lehm.	*PRE 11, Vorster s.n.*
E. macrostrobilus	*Vorster 1073*
E. msinganus Vorster	*PRE 20*
E. munchii R.A. Dyer & Verdoorn	*Vorster s.n.*
E. natalensis R.A. Dyer & Verdoorn	*PRE 18*
E. senticosus Vorster	*Vorster 273*
E. trispinosus (Hook.) R.A. Dyer	*Vorster s.n.*
E. turneri Lavranos & Goode	*Vorster 1110*
E. umbeluziensis R.A. Dyer	*PRE 23, Vorster 267*
E. villosus Lem.	*PRE 02, Vorster 265, SBG 03*
E. cf. *woodii*[6] Sander	*Vorster 682*
E. sp. (Didinga)[7]	*Vorster 1101*

[3] *Vorster s.n.* (*E. ferox*) comes from Vila Joao Bello in Mozambique. These plants are morphologically distinct from those in northern KwaZulu-Natal.

[4] *Vorster 322* (*E. ghellinckii*) represents the lowland form from near Port Shepstone, which is morphologically distinct from the montane form from near Cathedral Peak.

[5] *Vorster 258* (*E. laevifolius*) represents the glaucous form from near Kaapse Hoop; while *Vorster 260* represents the green form from Mariepskop.

[6] *Encephalartos* cf. *E. woodii* (*Vorster 682*) is an enigmatic taxon, like *E. woodii* known from a single clone, of unknown origin. It strongly resembles *E. woodii*, but the juvenile foliage is morphologically different.

[7] *Encephalartos* sp. (Didinga) (*Vorster 1101*) is an as yet, undescribed species from the Didinga Hills in south-eastern Sudan. Apparently it is allied to *E. septentrionalis*.

E. sp. (Moyo)[8]	*Vorster 1094*
E. senticosus x E. trispinosus	*Vorster 502*
E. umbeluziensis x E. lehmannii	*Vorster 438*
E. umbeluziensis x E. villosus	*Vorster 416a*
E. altensteinii x E. trispinosus	*Vorster 701*
E. trispinosus x E. altensteinii	*Vorster 506*
E. trispinosus x E. ferox	*Vorster 504*
E. transvenosus x E. woodii	*Vorster 787*
Stangeria eriopus (Kunze) Nash	*Vorster 337*

PRE: GAUTENG.—2528 (Pretoria): National Botanical Garden (-CA).
FS: FREE STATE.—2926 (Bloemfontein): University of the Orange Free State Botanical garden (-AA).
Vorster: WESTERN CAPE.—3318 (Stellenbosch): P.J.Vorster's garden (-DD). *Vorster* numbers are garden accession numbers, referring to Vorster's private living collection in Stellenbosch.
SBG: WESTERN CAPE.—3318 (Stellenbosch): Stellenbosch Botanical Garden.

2.2 Methods

2.2.1 DNA extraction method

Fresh leaves from the different specimens were collected and stored in a saturated mixture of hexadecyl trimethyl ammonium bromide (CTAB) and sodium chloride and stored at 4°C (Rogstad 1992).

The DNA extraction method of Rogstad (1992) was slightly modified. Approximately 1g of leaf material was ground in liquid nitrogen to a fine powder and incubated in extraction buffer at 65°C for 60 minutes. The extraction buffer consists of 0.1M 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) (pH 8), 1.4M sodium chloride, 0.025M ethylene diamine tetracetic acid (EDTA) and 1% sodium N-lauroyl sacosine sodium, to which 1% (v/v) 2-mercapto-ethanol had been added immediately prior to use. The organic and liquid phases were separated by adding chloroform:iso-amylalcohol (24:1 v/v) to the mixture, vortexed and centrifuged at 5000G for five minutes. The

[8] *Encephalartos* sp. (Moyo) is yet another undescribed species from Moyo in Uganda. Apparently it is also allied to *E. septentrionalis*.

supernatant was transferred to a clean tube. The DNA in the supernatant was precipitated at -20°C for 60 minutes with two volumes ethylalcohol containing 3M sodium acetate (25:1). The mixture was centrifuged at 10 000G for ten minutes. The supernatant was removed and the pellet washed with 70% (v/v) ethylalcohol containing 0.01M ammonium acetate, air dried and dissolved in sterile water to a final concentration of ±2.5ng.μl^{-1} and stored at -20°C.

The genomic DNA samples were mixed with 6X loading buffer {0.25% (m/v) bromophenol blue, 0.25% (m/v) xylene cyanol FF and 30% (v/v) glycerol in water} and loaded on to 0.8% (m/v) agarose gel, using a TAE running buffer of 0.04M Tris (pH 8.0), 0.001M EDTA and 35% (v/v) glacial acetic acid (Sambrook *et al.* 1989). Ethidium bromide was added to a final concentration of 0.5ng.μl^{-1}. The genomic fragment was separated at 100V for fifteen minutes and examined under ultraviolet (UV) light, to visualise the success of DNA extraction.

The DNA samples were purified using a modified phenol-chloroform extraction method (Sambrook *et al.* 1989). An equal volume of phenol (pH 7.8-8.0) and chloroform were added to the DNA sample. The solution was mixed and centrifuged for 3 minutes at 5000G to separate the organic and aqueous phases. The aqueous phase was transferred to a clean tube and the DNA was recovered by precipitation with two volumes 100% ethylalcohol at -20°C for 30 minutes. The mixture was centrifuged for 5 minutes at 10 000G, after which the supernatant was removed. The pellet was washed with 70% ethylalcohol, centrifuged for 5 minutes at 10 000G. The samples were air dried then dissolved in distilled water.

2.2.2 PCR optimisation

A modified Taguchi method (Cobb & Clarkson 1994) was used to determine the optimal conditions for the RAPD and DAF PCR reactions. The components of the PCR reaction that have a major effect on amplification, were identified and called the control factors (Taguchi & Wu 1980; Taguchi 1986). These control factors were used in a combination of nine reactions,

with three different concentrations for each reaction component tested. The variable components tested were dNTPs, primer, magnesium chloride and DNA concentrations.

The product yield of each Taguchi reaction was used to estimate the effect of individual components on amplification. This is called the signal-to-noise ratio and was done using the quadratic loss functions (Taguchi 1986). The gel was scored qualitatively according to the number and distribution of the products of each reaction and the data were used to determine the optimal reaction by calculation of:

$$SNL = - 10 \log[1/n \sum 1/y^2],$$

where SNL is the signal to noise ratio, n is the number of levels and y is the yield. The largest SNL gives an indication of the most optimal condition for each component. The SNL values for each control factor were projected on to a graph (y-axis), with the different concentrations of each reagent tested, on the x-axis. Using the second order function (k=2), the plotted values are linked into a curve. The highest SNL value is used to determine the optimum concentration of each reagent.

RAPD profiles were scored for specific characteristics such as a larger fragment size and a greater fragment size range and, therefore, the yield was determined differently using the equation:

$$P = (r \times s) + 1,$$

where P is the product yield, r is the number of products and s is the size range. When the fragment size was 1kb or smaller a value of s = 1 was given to the profile. If the fragment size was up to 2kb, s = 2 were given to the profile. Reactions that did not amplify or produce a smear of amplification products were given a score of P = 1. The product yield (P) is used in the first equation to determine the SNL value of RAPD profiles.

2.2.3 Random amplified polymorphic DNA analysis (RAPD)

Molecular data are subject to experimental errors which differ depending on the technique (Karp *et al.* 1996). Reproducibility is mostly used as a way of assessing the quality of the data, but it is impossible to take all experimental variables into account and, therefore, controls such as negative controls and duplicated reactions were included in the study (Karp *et al.* 1996). All RAPD reactions were performed on a Perkin-Elmer GeneAmp PCR system 9600, using the following cycling conditions. A denaturation step of 60 seconds at 94°C, followed by 40 amplification cycles of 10 seconds at 94°C, 15 seconds at 37°C and 75 seconds at 72°C. Eight primers were tested for RAPD analysis of the specimens (Table 2.2).

Table 2.2 The eight, ten nucleotide primers, used for RAPD analysis of the specimens.

Primer	Sequence of primer
OPA7	5'-GAAACGGGTG-3'
OPA11	5'-CAATCGCCCT-3'
OPA16	5'-AGCCAGCGAA-3'
OPA20	5'-GTTGCGATCC-3'
OPB3	5'-CATCCCCCTG-3'
OPB6	5'-TGCTCTGCCC-3'
OPC4	5'-CCGCATCTAC-3'
OPC16	5'-CACACTCCAG-3'

2.2.4 DNA amplification fingerprinting analysis (DAF)

Duplicated DAF reactions were amplified in a Perkin-Elmer GeneAmp 9600, programmed for an initial denaturation step of 60 seconds at 94°C, followed by 35 amplification cycles of 10 seconds at 94°C, 15 seconds at 30°C, and 75 seconds at 72°C. Twelve primers were screened in the study (Table 2.3).

Table 2.3 The twelve, eight nucleotide primers screened for use in DAF analysis.

Primer	Sequence of primer
DAF1	5'-AACGGGTG-3'
DAF2	5'-GTAAGGCC-3'
DAF3	5'-GAGGGTGG-3'
DAF4	5'-CCTCGTGG-3'
DAF5	5'-GAAAGGCC-3'
DAF6	5'-GTTACGCC-3'
DAF7	5'-CTGGACTA-3'
DAF8	5'-GTAACGCC-3'
DAF9	5'-GTACTGCC-3'
DAF10	5'-GTAAGGCC-3'
DAF11	5'-CCTGCTGG-3'
DAF12	5'-CAGCTCGG-3'

2.2.5 Gel electrophoresis and DNA detection

In this study, two different separation methods (agarose and polyacrylamide) were used as a supporting medium in separating RAPD and DAF profiles:

- Agarose support medium

  The RAPD-PCR amplification products were separated on a 1.5% (m/v) agarose gel with 1X TAE running buffer [0.04 M Tris-acetate, 0.001M EDTA, 0.35% glacial acetic acid (v/v), pH 8] for 60 minutes at 80V and stained with 0.5ng.μl^{-1} ethidium bromide (Sharp *et al.* 1973), and visualised with an UV-illuminator at a wavelength of 302nm.

- Polyacrylamide support medium

DAF-PCR amplification products were separated on a 10M-8% polyacrylamide (ratio of 20:1) gel with 0.5X TAE running buffer for 90 minutes at 150V.

A modified silver staining method of Caetano-Anollés *et al.* (1994) was used to detect DNA in the polyacrylamide gels. To remove unwanted chemicals and prevent the diffusion of separated fragments in the gel matrix, the polyacrylamide gel was fixed with 7.5% acetic acid for a minimum of five minutes. The acetic acid was then removed with three 2-minute washes in deionized water. The gel was impregnated with $1.0g.l^{-1}$ silver nitrate and 0.056% formaldehyde for 30 minutes. Any silver remnants were removed with a 20 second wash in deionized water. The gel image was developed in a mixture of $30g.l^{-1}$ sodium carbonate, 0.056% formaldehyde and $400\mu g.l^{-1}$ sodium thiosulfate. The developing process was stopped with cold (4°C) 7.5% acetic acid for two minutes.

2.2.6 Gel documentation

All gel electrophoresis results were photographed (using an orange filter and a 13mm extension tube at an aperture of f/8.0 and a shutterspeed of 25–30 seconds), as well as electronically recorded using a BioRad GelDoc 1000 system.

2.3 Data analysis

2.3.1 Phylogenetic analysis using cladistics

A datamatrix for each primer was created by scoring the amplification products according to the presence (1) or absence (0) of the amplified product at a specific molecular weight. All amplification reactions were duplicated and only fragments present in both profiles were scored as being present. If a fragment was present in only one of the duplicated reactions a question mark (?) was allocated to the character. These data matrices were processed with

computer software and algorithms, to determine the most parsimonious cladogram and its accuracy.

2.3.1.1 Analysis using cladistic computer software

Phylogenetic Analysis Using Parsimony (PAUP, version 4.0, Swofford 1998) is a computer software program that performs distance matrix and parsimony analyses. The program creates phylogenies from character data and presents the results as cladograms (Swofford 1993). In this study the method of parsimony was used to search for minimum length cladograms (Swofford 1993): those that minimise the total amount of evolutionary changes needed to explain the variation in a given set of data (Swofford & Maddison 1987).

Characters were optimised as far as possible from the root using DELTRAN. Character types were assumed unordered, character weights uncertain, multistates interpreted as uncertain, and uninformative characters ignored. Some characters were temporarily removed when CI:RI ratio was less than 2, and CI value less than 0.2 (Lipscomb 1998).

Most parsimonious cladograms were determined from the molecular data matrices using heuristic searches with steepest descent, MULPAR and the global branch swapping algorithm, tree bisection and reconnection (TBR) in effect. Cladograms were described using consistency index, retention index, tree length, and number of equally parsimonious cladograms. In cases where the cladograms yielded a RI value less than 0.67, a search for multiple islands of parsimony was also conducted (Maddison 1991).

Support for the cladograms recovered was tested with the bootstrap (Felsenstein 1985), decay value (Bremer 1988), and Jackknifing (Lanyon 1985, Siddall 1995). Bootstrap is a method of resampling one's own data and thereby placing confidence intervals on phylogenies (Felsenstein 1985). In this study a bootstrap of 200 replicates, using a heuristic search, was done. The decay index was determined because a branch present in one of the most parsimonious cladograms is more strongly supported by the data if a large increase in length of additional cladograms is required before that branch is

lost in the consensus (Bremer 1994). In most cases the computer used had insufficient memory and therefore no decay values were obtained. Taxon Jackknifing was also performed whereby the taxa were sequentially deleted.

Where multiple parsimonious cladograms were produced, a Strict consensus cladogram was obtained, because it is not based on characters but on topologies, and illustrates components which are common to all the equally parsimonious cladograms of an analysis (Anderberg & Tehler 1990).

2.3.1.2 Analysis using the index of the genetic similarity and pairwise genetic distance

The index of genetic similarities (F) of Nei & Li (1979) was used to calculate pairwise genetic distances (D) for all the species:

$$F = 2N_{xy}/(N_x + N_y),$$

where N_{xy} is the number of fragments duplicated, and N_x and N_y are the number of fragments in each of the individual reactions. To calculate the genetic distance (D) between samples the following formula was used:

$$D = \ln F.$$

CHAPTER THREE

A PILOT STUDY OF TWO MOLECULAR TECHNIQUES FOR ANALYSIS OF PHYLOGENIES IN THE GENUS *ENCEPHALARTOS*

3.1 Introduction

The aim of the pilot study was to determine the applicability of two molecular techniques, as well as the most informative primers for a systematic study of the genus *Encephalartos*.

The molecular techniques tested, were random amplified polymorphic DNA analysis (RAPD) and DNA amplification fingerprinting (DAF), as well as variations of the DAF technique i.e. DAFs using DNA template digested with a restriction endonuclease (tecMAAP), and multiplex DAF reactions.

Five *Encephalartos* specimens were used in this preliminary study: *E. villosus* (*Vorster 265*), *E. altensteinii* (*FS 01*), *E. natalensis* (*PRE 18*), *E. lembomboensis* (*Vorster 313*), and *E. ferox* (*Vorster s.n.*). *Stangeria eriopus* (*Vorster 337*) was used as an outgroup. The six specimens were used to verify the applicability of the different techniques in the analysis of the genus *Encephalartos*. They were not intended to resolve the phylogenetic relationships among them.

3.2 Results and Discussions

3.2.1 Random amplified polymorphic DNA analysis

3.2.1.1 RAPD optimisation

To determine the optimum reagent concentrations for RAPD reactions, different primer, magnesium chloride, DNA template and dNTP concentrations were tested (Table 3.1). The RAPD primer OPA 20 and specimen *E. natalensis* (*PRE 18*) were used. Buffer and *Taq* polymerase concentrations were kept constant and sterile water was added to a final concentration of 25µl per reaction. The reactions were amplified (Chapter 2.2.3) and the amplified products were separated on 1.5% agarose gel (Chapter 2.2.5).

Table 3.1 The optimisation of the RAPD reaction, testing different concentrations of magnesium chloride, dNTP, primer and template.

Reaction	[$MgCl_2$] mM	[Primer] pmole.µl^{-1}	[dNTP] mM	[DNA template] ng.µl^{-1}
1	62.5	3.5	6	2
2	62.5	7.0	7	4
3	62.5	10.5	9	6
4	87.5	3.5	7	6
5	87.5	7.0	9	2
6	87.5	10.5	6	4
7	112.5	3.5	9	4
8	112.5	7.0	6	6
9	112.5	10.5	7	2

The profiles of the nine duplicated reactions (Figure 3.1) were analysed using the signal-to-noise (SNL) ratio (Chapter 2.2.2) and plotted on to graphs (Figure 3.2).

The profiles of the nine duplicated reactions (Figure 3.1) were analysed using the signal-to-noise (SNL) ratio (Chapter 2.2.2) and plotted on to graphs (Figure 3.2).

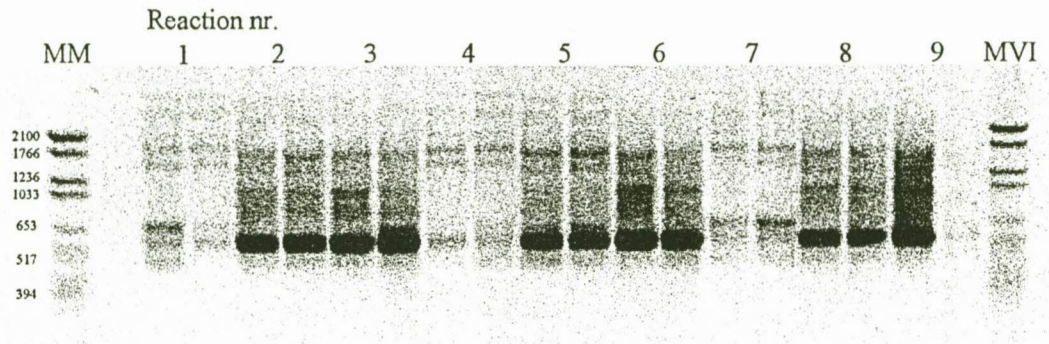


Figure 3.1 The nine duplicated RAPD reactions separated on 1.5% agarose gel to determine an optimised RAPD reaction for amplification of *Encephalartos natalensis* (*PRE 18*) specimens.

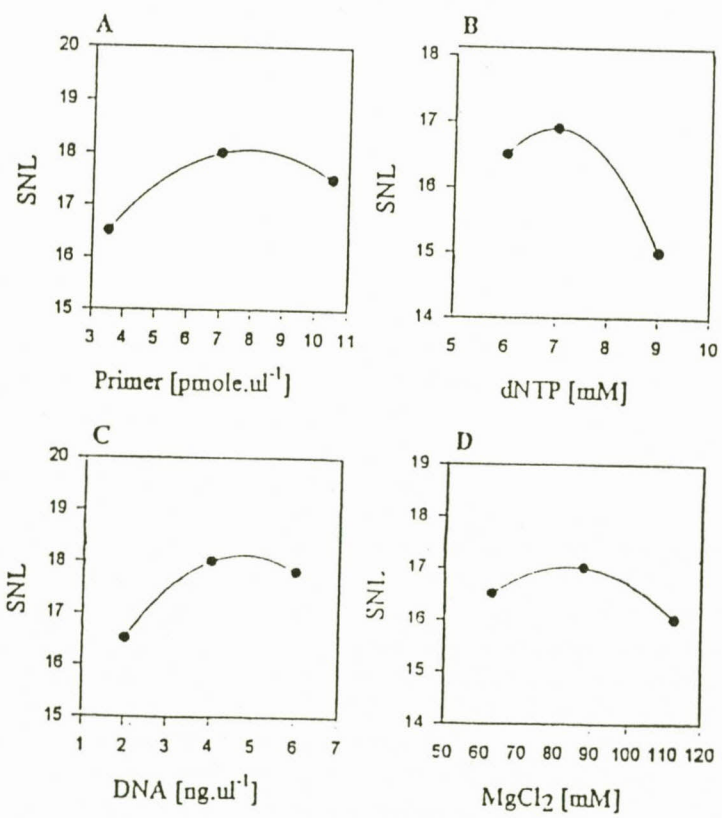


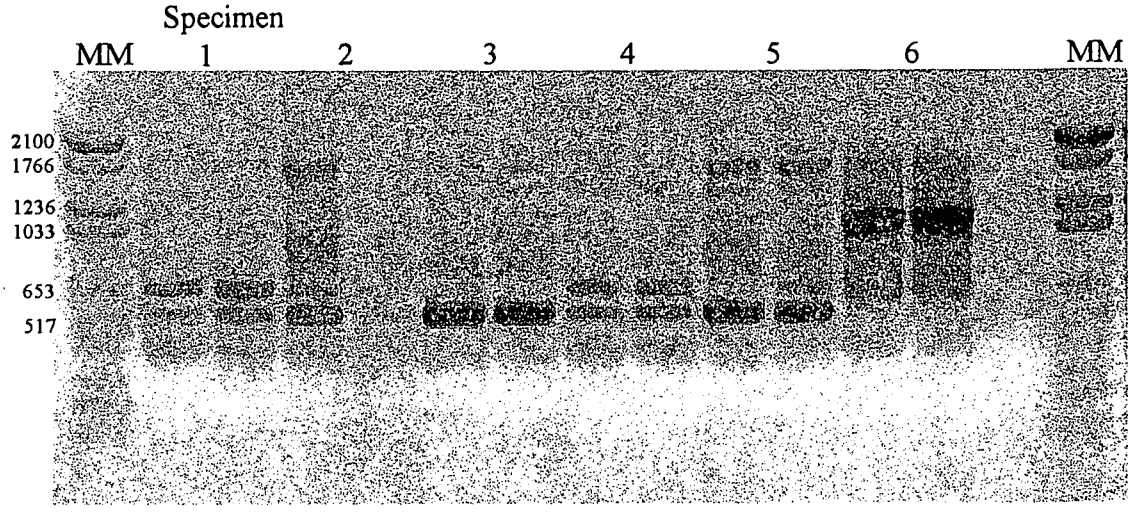
Figure 3.2 The signal-to-noise values (y-axis) plotted against: A – the primer, B – dNTP, C – DNA, and D – $MgCl_2$, concentrations (x-axis).

Sub-optimal conditions can produce PCR artifacts such as dimerization, misincorporation and incorrect extension of the primer template (Cobb & Clarkson 1994). The magnesium ion concentration in a reaction, may influence the primer annealing, product specificity, primer-dimer artifacts, and enzyme activity (Innis & Gelfand 1990). In this study the optimum concentration of magnesium chloride levels were found to be 75mM. According to the graph (Figure 3.2a), the optimum concentration for dNTP is 7mM, but a lower concentration (6mM) was found more effective in this study. Low dNTP concentration increases the specificity and accuracy of PCR, as it minimises mispairing at non-target sites (Innis *et al.* 1988). The optimum primer concentration for this study is 7 $\rho mole.\mu l^{-1}$, because high primer concentrations can result in primer-dimer formation (Innis & Gelfand 1990). A high concentration of template results in smears, without distinct amplification fragments (Williams *et al.* 1990). In this study a concentration of $6ng.\mu l^{-1}$ of template was used.

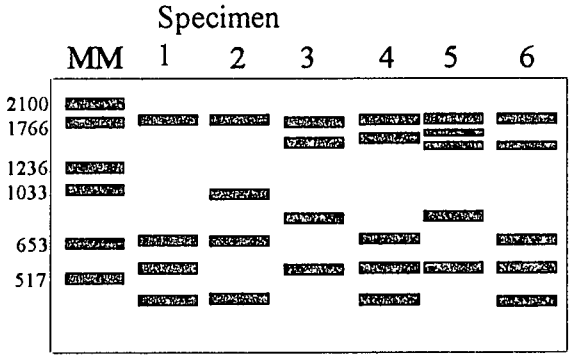
3.2.1.2 RAPD analysis

Eight RAPD primers (Table 2.2) were tested, and an example of the amplification profiles and schematic representation of primer OPA7 is given, indicating separation on agarose medium (Figure 3.3).The eight primers: OPA7, OPA11, OPA16, OPA20, OPB3, OPB6, OPC4 and OPC16 produced fifty-nine characters, present in both duplicated reactions (Appendix A). The coefficient of similarity (F) for each primer was calculated, (Table 3.2), to determine the most informative primers for this study. The number of characters generated by a primer per specimen was also compared.

A high coefficient of similarity was obtained from all the primers, ranging from $F_{OPC4} = 0.93$ to $F_{OPB6} = 1.0$. The highest number of repeated fragments (N_{xy}) was scored from primer OPA16 ($N_{xy} = 28$) and the lowest from primer OPC16 ($N_{xy} = 9$). Primers OPA11 and OPA16 generated the most characters per specimen (11 characters) and OPB6 and OPC4 generated the lowest number of characters per specimen (4). All the primers were successfully used in the production of amplification profiles.



A.



B.

Figure 3.3 A photograph (A) and a schematic representation (B) of the RAPD amplification profiles of primer OPA7 for the six specimens separated on 1.5% agarose gel. 1 – *E. villosus* (*Vorster 265*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. natalensis* (*PRE 18*). 4 – *E. ferox* (*Vorster s.n.*). 5 – *E. lebomboensis* (*Vorster 313*). 6 – *Stangeria eriopus* (*Vorster 337*).

Table 3.2 The coefficient of similarity (F), and character / primer / specimen, for each RAPD primer calculated.

		OP A7	OP A11	OP A16	OP A20	OP B3	OP B6	OP C4	OP C16
N_x	# of the fragments of reaction X for a primer.	22	25	28	27	22	14	15	9
N_y	# of the fragments of reaction Y for a primer.	22	27	30	27	21	14	13	10
N_{xy}	# of the repeated fragments found in reaction X and Y for a primer.	21	25	28	26	21	14	13	9
F	$= 2N_{xy}/(N_x + N_y)$	0.95	0.96	0.97	0.96	0.97	1.0	0.93	0.94
	Characters / primer / specimen	9	11	11	8	8	4	6	4

3.2.2 DNA amplification fingerprinting

3.2.2.1 DAF optimisation

DAF reactions were optimised using a modified Taguchi method (Cobb & Clarkson 1994). The four control factors tested were magnesium chloride, primer (DAF2), template [specimen *E. villosus* (*Vorster 265*)], and deoxynucleotide triphosphate concentrations.

The SNL values of the nine duplicated reactions (Figure 3.4) were calculated and analysed as previously done in the RAPD optimisation, to determine the optimum concentrations for each reagent.

Table 3.3 The orthogonal arrangement of the control factors tested to determine an optimised DAF reaction.

Reaction	[$MgCl_2$] mM	[Primer] $\rho mole.\mu l^{-1}$	[dNTP] mM	[DNA template] $ng.\mu l^{-1}$
1	25	25	2.0	1.0
2	25	100	6.0	3.0
3	25	200	10.0	6.0
4	75	100	2.0	6.0
5	75	200	6.0	1.0
6	75	25	10.0	3.0
7	125	200	2.0	3.0
8	125	25	6.0	6.0
9	125	100	10.0	1.0

The SNL values of the nine duplicated reactions (Figure 3.4) were calculated and analysed as previously done in the RAPD optimisation, to determine the optimum concentrations for each reagent.

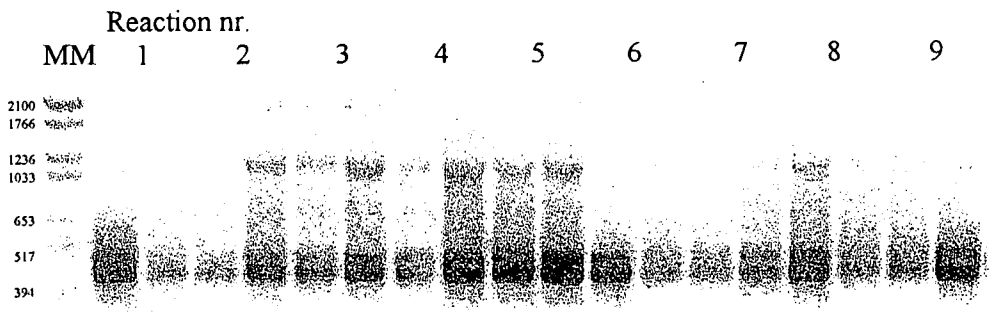


Figure 3.4 The nine duplicated DAF reactions separated on 1.5% agarose gel to determine an optimised DAF reaction for amplification of *Encephalartos natalensis* (*PRE 18*) specimens.

Magnesium ion levels are critical, depending on the complexity of the genome, with less magnesium chloride for more complex genomes such as soybean (Bassam *et al.* 1992, Jayarao *et al.* 1992). An excess of magnesium chloride levels decreases DAF amplification stringency and increases primer-template mismatching (Caetano-Anollés *et al.* 1994). The optimum magnesium chloride concentration for DAF amplification in this study is 75mM.

The primer length, primer sequence and number of primers in the DAF reaction influence the complexity of the fingerprints and detection of polymorphisms (Caetano-Anollés 1997). Decreasing the primer length (down to 5nt), minimises the number of interactions within and between the different primer and template species during amplification (Caetano-Anollés *et al.* 1992b). However, higher primer concentrations are then needed, which could lead to inhibition of the amplification reaction (Caetano-Anollés *et al.* 1992b). An increase of primer length could cause mismatching during primer annealing (Caetano-Anollés 1994). Therefore, a primer length of 8 nucleotides provides a good compromise between efficiency and specificity in DAF analysis (Caetano-Anollés *et al.* 1992b, Caetano-Anollés 1994, Caetano-Anollés & Gresshoff 1994a). The ratio of primer concentrations to template is the most important variable in the amplification reaction (Gresshoff 1996). According to Gresshoff (1996), the DAF primer concentration should be about ten times the concentration of the RAPD reaction when using an octamer primer. Although such high primer concentrations were tested in this study, a concentration of 25pmole.μl^{-1}, (only four times that of the RAPD reaction) proved to be suitable for DAF amplification of *Encephalartos* specimens.

The template used in the DAF amplification reaction is also important and a DNA concentration as low as 0.1ug.ul^{-1} can produce consistent DAF fingerprints from most plant and animal genomes (Caetano-Anollés & Gresshoff 1994a). Too low levels of template can cause stoichiometric misrepresentation and lack of reproducibility (Caetano-Anollés & Gresshoff 1994a). In this study 3ng.μl^{-1} DNA template per specimen is sufficient for amplification.

3.2.2.2 DAF reactions

Twelve DAF primers (Table 2.3) were used, but only five primers (DAF2, DAF3, DAF5, DAF6 and DAF10) gave consistent results and were further analysed in this study. The amplification products of these primers, separated on polyacrylamide medium (Figure 3.5), produced 47 characters, observed in both the duplicated reactions per specimen (Appendix B).

The most informative primers for this study were determined by calculating the F value and number of characters produced per primer for a specimen (Table 3.4).

Table 3.4 The F values for each primer calculated to determine the most informative primers for DAF analysis.

		DAF 2	DAF 3	DAF 5	DAF 6	DAF 10	DAF 12
N_x	# of the fragments of reaction X for a primer	37	20	32	14	26	10
N_y	# of the fragments of reaction Y for a primer	37	20	31	12	25	9
N_{xy}	# of the repeated fragments found in reaction X and Y for a primer	37	20	31	12	25	8
F	$= 2N_{xy}/(N_x + N_y)$	1	1	0.98	0.92	0.98	0.84
	Characters / primer / specimen	10	5	10	7	9	6

The coefficient of similarity (F) for the DAF primers ranged from F_{DAF12}=0.84 to as high as F_{DAF2}=1.0. The number of characters generated by a primer per specimen ranged from 6 characters (DAF12) to 10 characters

(DAF2 and DAF5). A higher number of fragments per specimen were expected, but an improvement of the detection method will increase this number, as no faint fragments were included in the analysis.

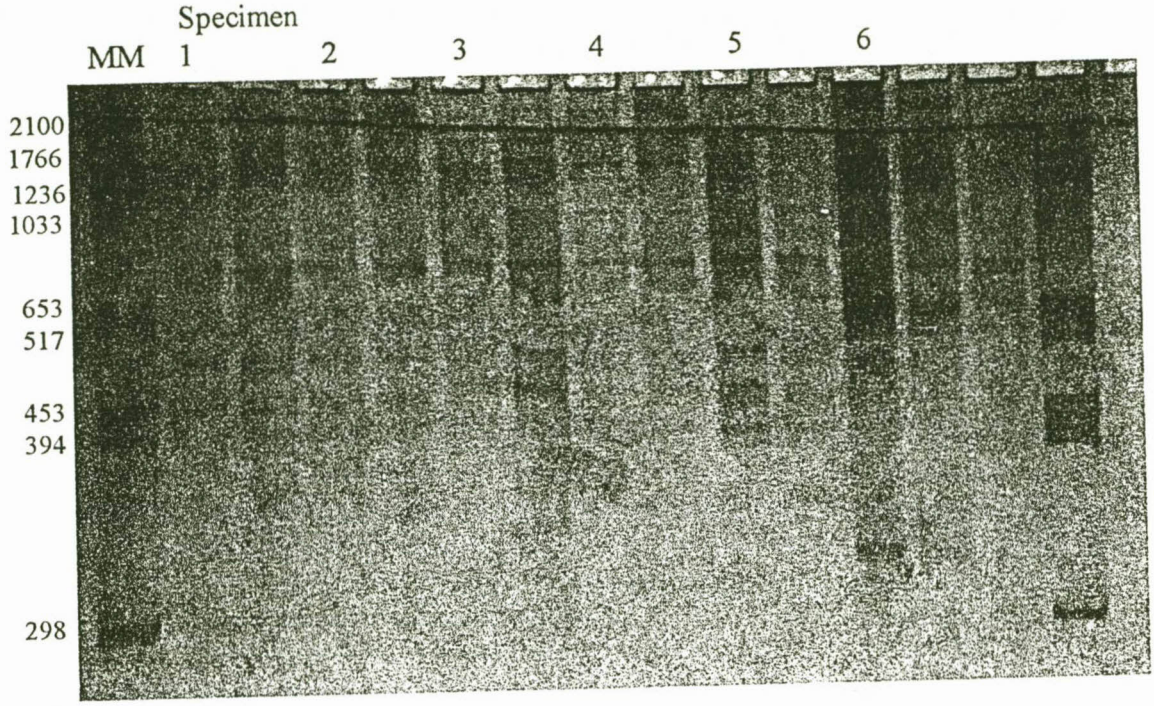
3.2.2.3 DAF using template digested with restriction enzyme (tecMAAP)

There are different DAF variations, including the digestion of template DNA with restriction endonuclease prior to amplification. In this study the restriction endonuclease *Hind III* (Boehringer Mannheim) was used and the amplification product was separated on polyacrylamide medium. A photograph and schematic representation of primer DAF2 amplification products separated on polyacrylamide medium are shown in Figure 3.6. Seventy-nine characters present in both of the duplicated reactions, were scored from 6 primers (Appendix C).

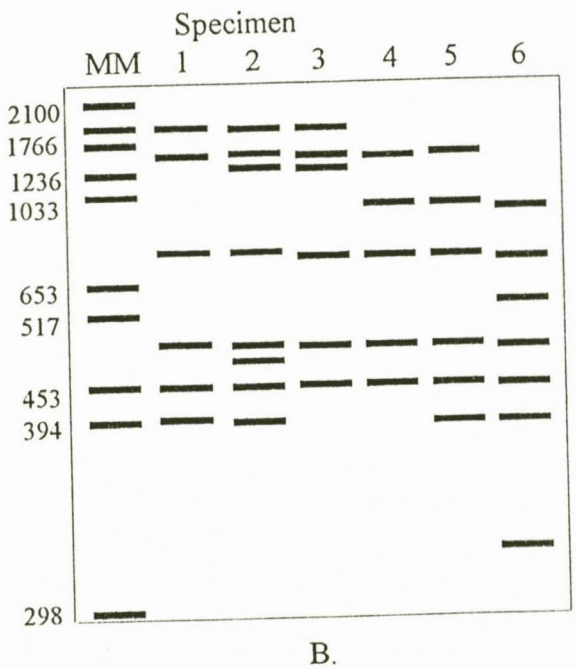
The most informative primers were determined by comparing the F value for each primer, as well as the number of characters generated by a primer for a specimen (Table 3.5).

The high F values for each primer, ranging from $F_{DAF5} = 0.95$ to $F_{DAF2} = 1.0$, indicate that the tecMAAP (digestion of template prior to amplification) is a suitable method in obtaining amplification fingerprints for phylogenetic analysis of the genus *Encephalartos*. An

increase in characters per specimen per primer was also observed, ranging from DAF2 = 11 to DAF12 = 16. It was also observed that DAF12 is more successful in tecMAAP than using the primer in normal DAF amplification reactions.



A.



B.

Figure 3.5 A photograph (A) and schematic representation (B) of the DAF amplification profile for primer DAF2, of the six specimens separated on 8% polyacrylamide gel. 1 – *E. villosus* (*Vorster 265*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. natalensis* (*PRE 18*). 4 – *E. ferox* (*Vorster s.n.*). 5 – *E. lebomboensis* (*Vorster 313*). 6 – *Stangeria eriopus* (*Vorster 337*).

Table 3.5 The coefficient of similarity (F) of primers DAF2, DAF3, DAF5, DAF6, DAF10 and DAF12, using tecMAAP technique.

		DAF 2	DAF 3	DAF 5	DAF 6	DAF 10	DAF 12
N_x	# of the fragments of reaction X, from a primer	49	54	32	51	36	51
N_y	# of the fragments of reaction Y, from a primer	48	54	29	50	36	48
N_{xy}	# of the repeated fragments found in reaction X and Y of a primer	48	54	29	50	36	48
F	$= 2N_{xy}/(N_x + N_y)$	1.0	1.0	0.95	0.99	1.0	0.97
	Characters / primer / specimen	11	15	11	15	11	16

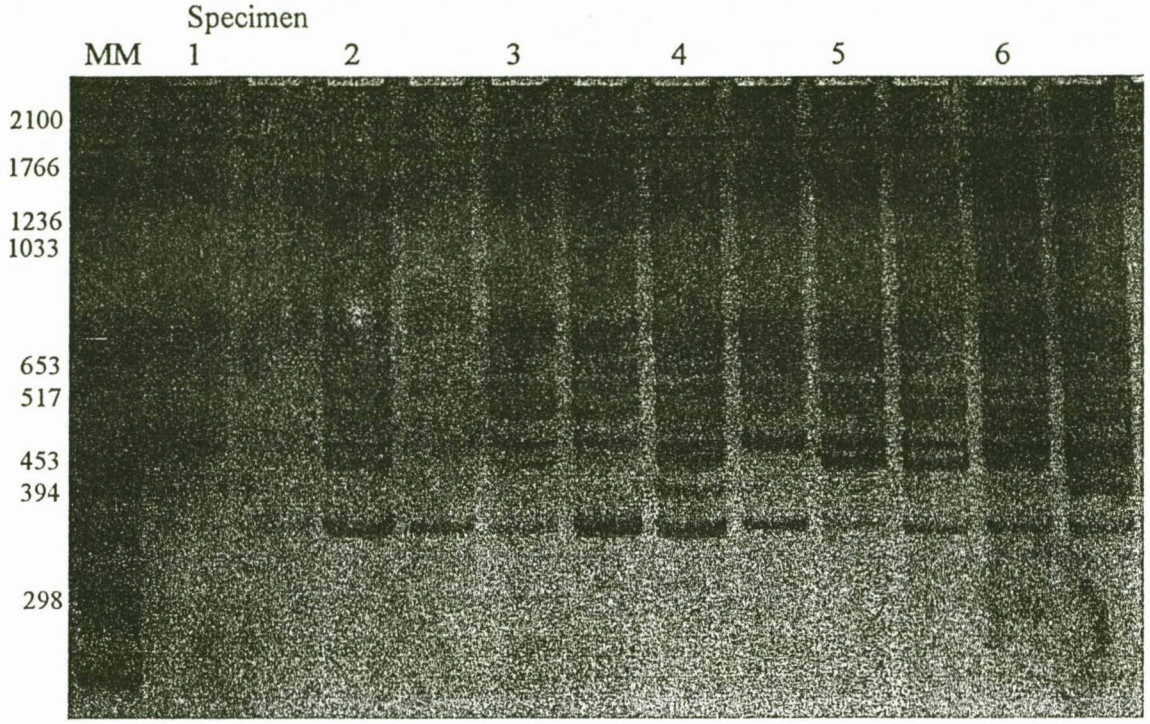
3.2.2.4 Multiplex DAF reactions

A second variation of the DAF technique is the use of multiplex DAFs. Multiplex DAF is the amplification with two or more nucleotide primers (Caetano-Anollés *et al.* 1991, Callahan *et al.* 1993). Primers DAF2, DAF3 and DAF5 were used in different combinations to form three multiplex DAF primers (Table 3.6). The amplification products were separated on polyacrylamide gel (Figure 3.7).

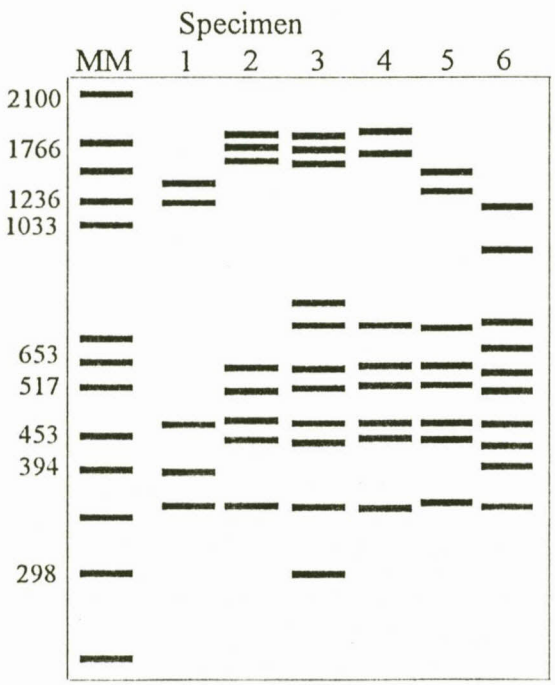
Multiplex DAF was the most successive technique with fifty characters per specimen scored from the three multiplex DAF primers (Appendix D).

3.2.2.5 Comparison of the different DAF techniques

The three different variations of the DAF technique used in this study were compared to determine the degree of variance for each technique, as well as the applicability of the techniques for phylogenetic analysis of the genus *Encephalartos* (Table 3.7).



A.



B.

Figure 3.6 A photograph (A) and schematic representation (B) of the DAF profiles, with template DNA digested with *Hind III* before amplification with primer DAF2, and separated on 8% polyacrylamide. 1 – *E. villosus* (*Vorster 265*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. natalensis* (*PRE 18*). 4 – *E. ferox* (*Vorster s.n.*). 5 – *E. lebomboensis* (*Vorster 313*). 6 – *Stangeria eriopus* (*Vorster 337*).

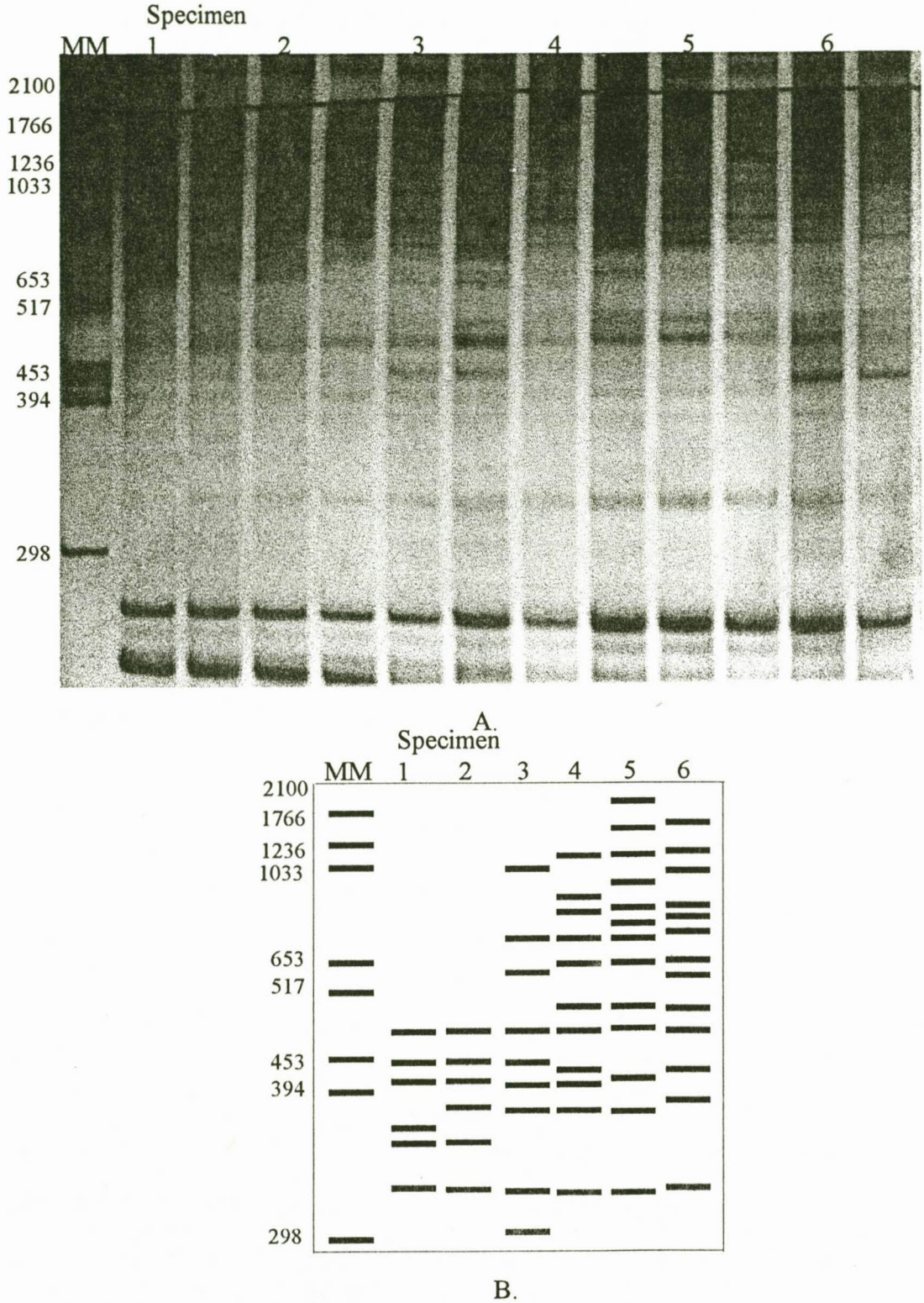


Figure 3.7 A photograph (A) and schematic representation (B) of the DAF profiles for primer K1, of the six specimens separated on 8% polyacrylamide gel. 1 – *E. villosus* (*Vorster 265*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. natalensis* (*PRE 18*). 4 – *E. ferox* (*Vorster s.n.*). 5 – *E. lebomboensis* (*Vorster 313*). 6 – *Stangeria eriopus* (*Vorster 337*).

Table 3.6 The multiplex DAF primers formed by combining three DAF primers.

DAF primer combination	New multiplex DAF primer
DAF2 and DAF3	K1
DAF2 and DAF5	K2
DAF3 and DAF5	K3

Table 3.7 The degree of variance of the three DAF techniques tested in the study, using profiles obtained from polyacrylamide separation.

		DAF reactions	tecMAAP	multiplex DAFs
N_x	# of fragments of the first repeat (X) of a reaction	139	273	178
N_y	# of fragments of the second (Y) repeat of a reaction	134	265	176
N_{xy}	Number of duplicated fragments found in both repeats of a reaction	133	265	176
F	$= 2N_{xy}/(N_x + N_y)$	0.97	0.98	0.99
	characters/specimen	47	79	50
	average number of duplicated fragments (N_{xy}) per primer	22	44	58

The three DAF techniques tested had an overall high F value, ranging from $F_{DAF} = 0.98$ to $F_{multiplex} = 0.99$. The number of characters per specimen per primer also increases from 47 fragments (DAF) to 79 fragments (tecMAAP). As previously found, multiplex DAF is the most successful technique for acquiring amplification profiles for phylogenetic analysis in this

study, with the highest average number of fragments per primer and F value.

3.3 Conclusion

PCR optimisation was done using a modified Taguchi method of Cobb & Clarkson (1994) and analysed by determining the SNL values (Taguchi 1986) for these control factors. The control factors, such as magnesium chloride, dNTPs, template, and primer concentrations, were identified because they play an important roll in the process of amplification. An optimised RAPD reaction consist of 75mM $MgCl_2$, 6mM dNTP, 7pmole primer, and 6ng.ul^{-1} template per 25µl reaction.

Optimal performance and reproducibility in DAF analysis are obtained with optimisation of the amplification parameters (Caetano-Anollés 1997). Optimisation of the DAF reaction differs from RAPD optimisation. In RAPD optimisation, larger molecular weight products are given a higher score in the signal-to-noise ratio (Gresshoff 1996). In DAF procedure the reaction combinations are sieved to remove high molecular weight products (Gresshoff 1996). An optimised DAF reaction for amplification of *Encephalartos* specimens consists of 75mM $MgCl_2$, 10mM dNTPs, 25pmole primer, and 3ng.ul^{-1} template in a 25 ul reaction volume.

The most informative primers for this study were determined by comparing the coefficient of similarity (F) with the number of characters scored per specimen for each primer. The high coefficient of similarity (F value) indicate that both RAPD and DAF can be applied in phylogenetic analysis of the genus *Encephalartos*. All the RAPD and DAF primers tested, are applicable in this study. Further RAPD analysis were done using primers OPA11, OPA16, OPA20, OPB3, and OPB6. These primers have also successfully been used in our laboratories for phylogenetic analysis of closely related specimens (Holder 1999, Kleynhans 1997, Klopper 1996, Roodt 1999, Visser 1997).

Five DAF primers (DAF2, DAF3, DAF5, DAF6 and DAF10) gave the

most consistent results. Although more fragments per primer per specimen were expected, an increase in the number of duplicated fragments was observed in the tailored DAF techniques.

Suitability of tailored DAF techniques, and the degree of variance between the different DAF techniques tested, were compared. All three techniques have a high coefficient of similarity. tecMAAP is a suitable technique for fingerprinting analysis, but it is a time consuming method in which successful digestion prior to amplification is necessary. A second tailoring technique, multiplex DAFs, successfully produced new amplification profiles. This cost effective method of producing new primers by combining three different primers was, therefore, used for further analysis in this study. Multiplex DAF and simple DAF reactions were applied to more specimens of the genus *Encephalartos*.

In conclusion, the RAPD and DAF technique is applicable in a systematic study of the genus *Encephalartos*, using the six RAPD primers (OPA7, OPA11, OPA16, OPA20, OPB3 and OPB6), separated on agarose medium, and five DAF primers (DAF2, DAF3, DAF5, DAF6 and DAF10), as well as three multiplex DAF primers (K1, K2 and K3), separated on polyacrylamide medium.

CHAPTER FOUR

Random amplified polymorphic DNA analysis

4.1 Introduction

The choice of a molecular technique to detect variation between species, particularly when examining hybridisation, is critical (Smith *et al.* 1996). The development of molecular techniques have resulted in various methods for detecting polymorphism in DNA-based procedures. Random amplified polymorphism analysis (RAPD) is one such a method.

RAPD technology is based on the polymerase chain reaction (PCR) and use single, short, arbitrary primers to amplify anonymous DNA fragments, which are then size fractioned by electrophoresis (Williams *et al.* 1990). The technique has been used in a wide variety of applications because of its speed and technical ease, with which large amounts of data can be generated without prior knowledge of any sequence information required (Gilles & Abbott 1998).

The speed and simplicity of the RAPD technique make this an ideal method to generate a large amount of fingerprinting data for a preliminary phylogenetic assessment of the genus *Encephalartos*, as very little molecular techniques have to date been applied to this plant group. The aim of this study was to applied the RAPD technique to more species of the genus *Encephalartos* for phylogenetic and a genetic distance analysis.

4.2 Results

The five RAPD primers (OPA11, OPA16, OPA20, OPB3 and OPB6) generated an average of 62 informative characters per specimen, ranging in fragment size from 400 to 2100 base pairs, and are schematically represented (Figures 4.1 – 4.5). Very faint fragments or fragments with no repetition were not included in this analysis. Some of the specimens did not successfully amplify and were, therefore, excluded from the RAPD analysis.

A binary data matrix was created from the RAPD profiles (Appendix E) and this data matrix was used to calculate the pairwise genetic distances (D) (Appendix F). The binary data matrix was also used in phylogenetic analysis. A heuristic search with the steepest descent, MULPAR, the swapping algorithm TBR in effect, yielded a consensus cladogram from 1296 equally parsimonious cladograms of 358 steps, a consistency index of 0.17 and retention index of 0.53 (Figure 4.6). The hybrid specimens were then excluded from the analysis which produced 6003 equally parsimonious cladograms of 319 steps, CI of 0.19 and RI of 0.54. A Strict consensus cladogram was calculated (Figure 4.7). Characters with a CI value of less than 0.2, CI:RI ratio of less than 2 (Libscomb 1998), and uninformative characters, as identified by PAUP, were then excluded from the analysis. This increased the resolution of the cladogram (Figure 4.8), and decreased the number of equally parsimonious cladograms to 35, the tree length to 275 steps, CI to 0.18 and gave a RI of 0.56.

4.3 Discussion

In the initial analysis all the specimens, including the hybrid specimens, were used (Figure 4.6). *Stangeria eriopus* was used as the outgroup and formed a sister group to the presumed monophyletic *Encephalartos* specimens. De Luca *et al.* (1995) also concluded from RFLP analysis that the monophyletic Stangeriaceae and the monophyletic Zamiaceae are sister groups to each other.

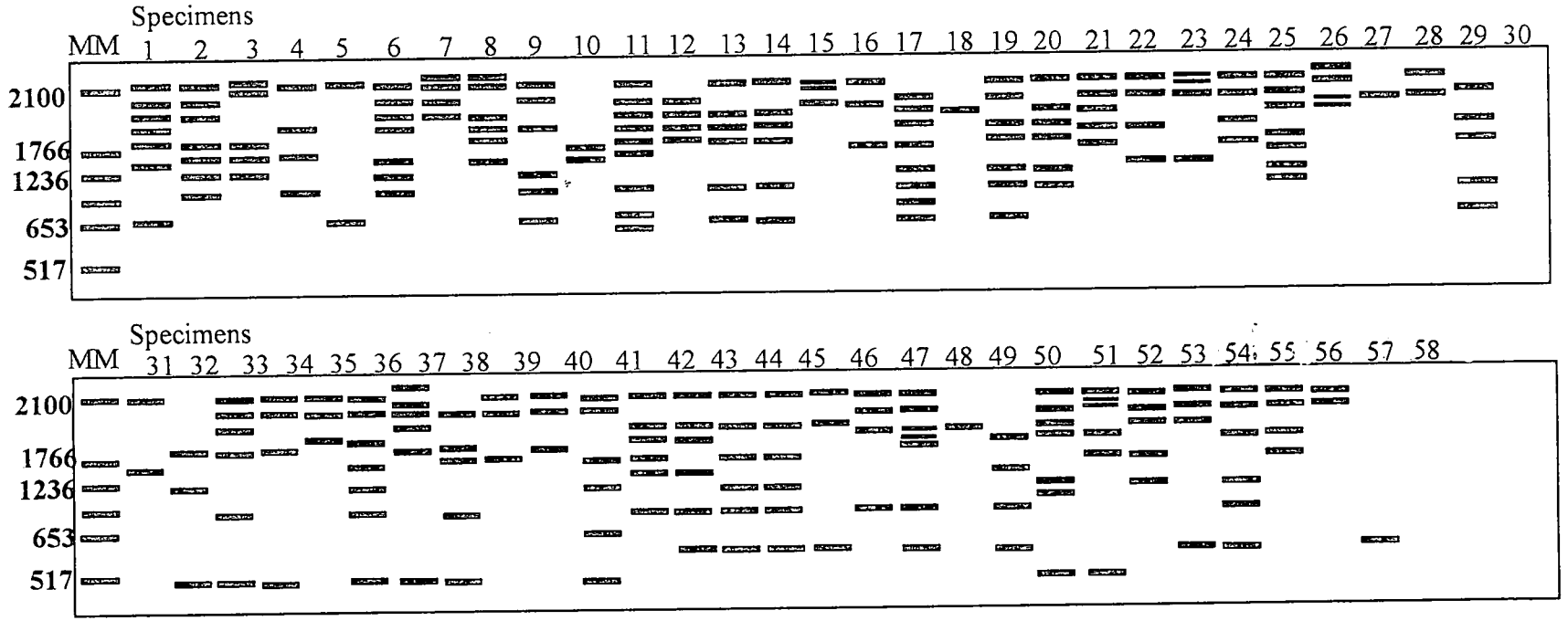


Figure 4.1 A schematic representation of the RAPD profiles generated by primer OPA 11. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

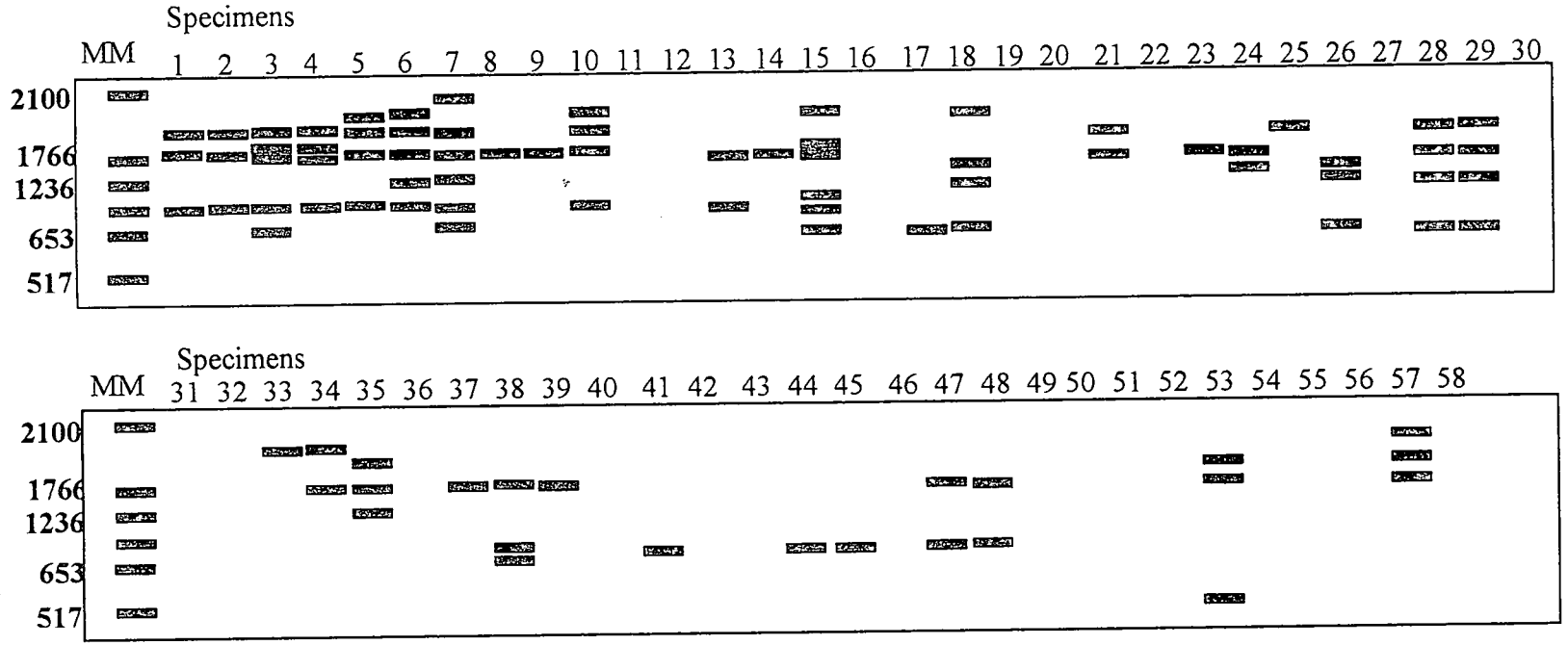


Figure 4.2 A schematic representation of the RAPD profiles generated by primer OPA 16. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

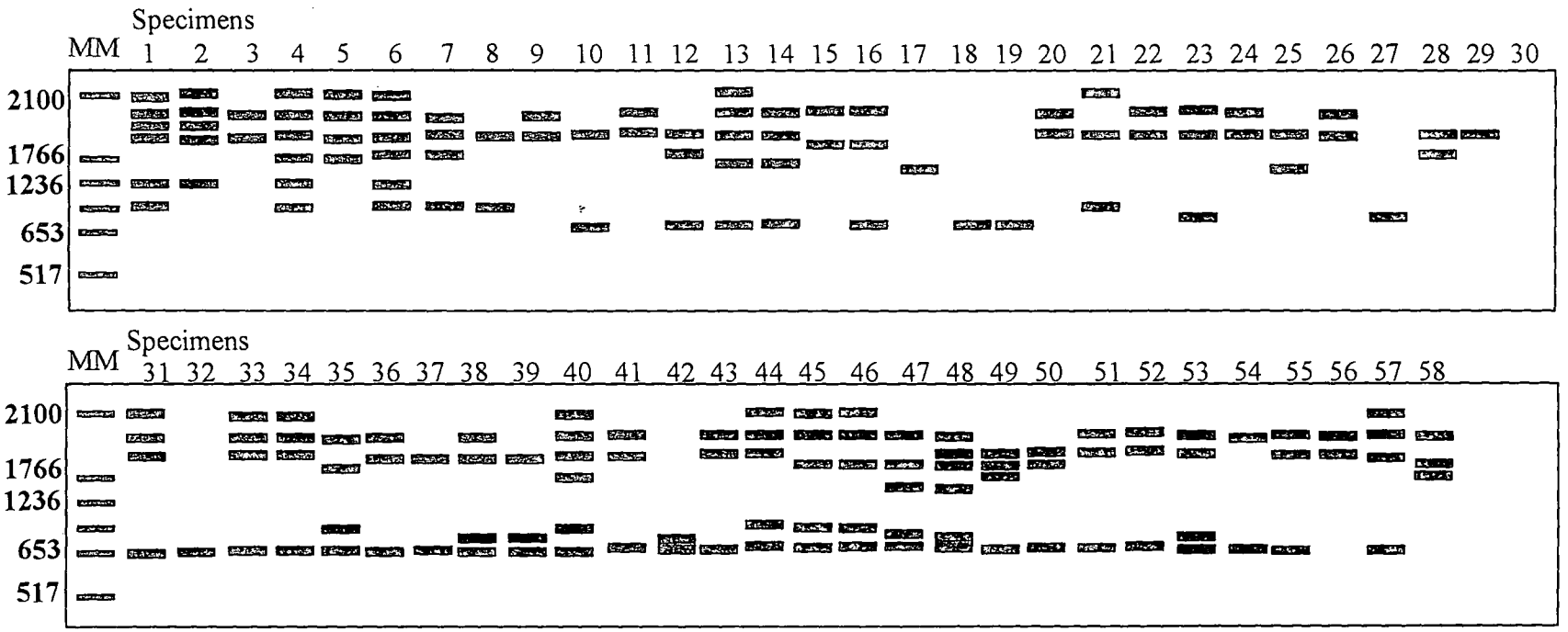


Figure 4.3 A schematic representation of the RAPD profiles generated by primer OPA 20. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

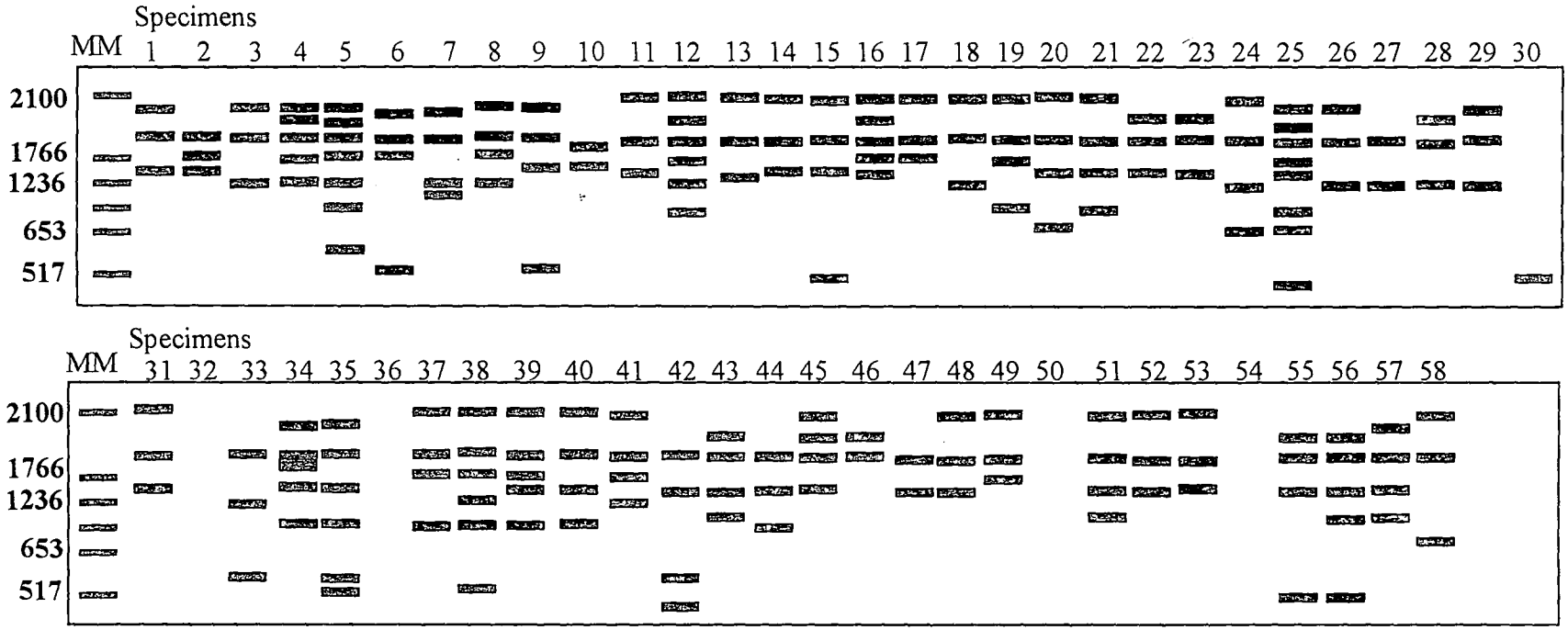


Figure 4.4 A schematic representation of the RAPD profiles generated by primer OPB 3. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

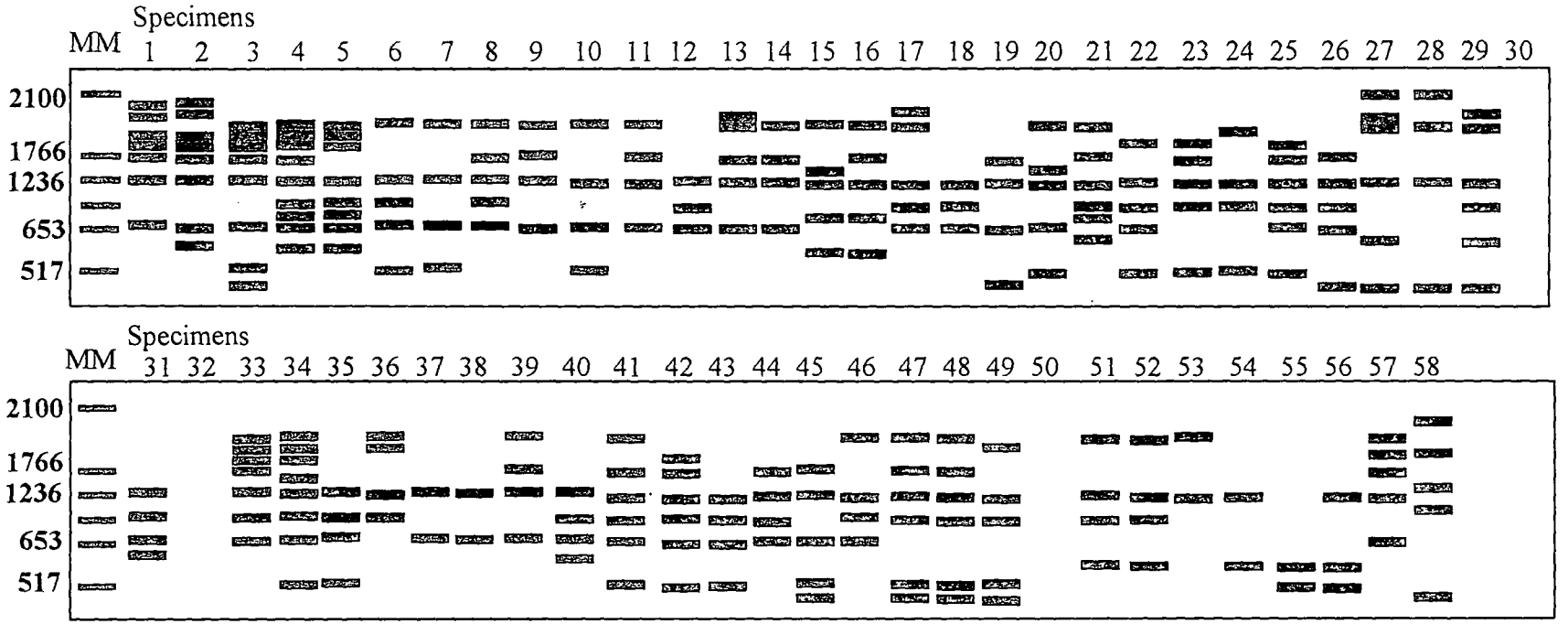


Figure 4.5 A schematic representation of the RAPD profiles generated by primer OPB 6. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

Most of the hybrid specimens included in the analysis were not grouped in the same clade as the putative parental specimens, except for the hybrid specimens *Encephalartos umbeluziensis x E. villosus* (*Vorster 416a*). Funk (1985) and McDade (1997) suggested that hybrid specimens do influence phylogenetic analysis and should be excluded from the initial analysis. The hybrid specimens were, therefore, removed from the initial analysis. The cladogram obtained without hybrid specimens had a smaller tree length (319) and higher CI (0.19), but more parsimonious cladograms were produced (Figure 4.7). Except for the excluded hybrids, the topology of the two cladograms changed very little.

CI values measure the relative homoplasy in a data set. RI indicates the number of synapomorphies expected from a data set that are retained as synapomorphies on a cladogram (Lipscomb 1998). Some of the characters were removed (where CI:RI ratio was less than 2, and CI value less than 0.2) and a heuristic search of the data set produced 35 equally parsimonious cladograms. A Strict consensus cladogram was calculated from the equally parsimonious cladograms (Figure 4.8). Compared to the previous consensus cladogram, the monophyletic ingroup was retained with an increase in resolution of the new consensus cladogram. The cladogram has the shortest length (275). According to Lipscomb (1998), the cladogram with the lowest length means that fewer homoplasies were required and fewer character state changes needed, therefore, the cladogram fits the data better and is more parsimonious.

The bootstrap and jackknife values higher than 65% are shown in the consensus cladogram indicating that, although the basal clades are not supported, most of the similar species are grouped together, indicating a close infraspecific relationships. Only the specimens of *E. lebomboensis* are grouped in more than one clade.

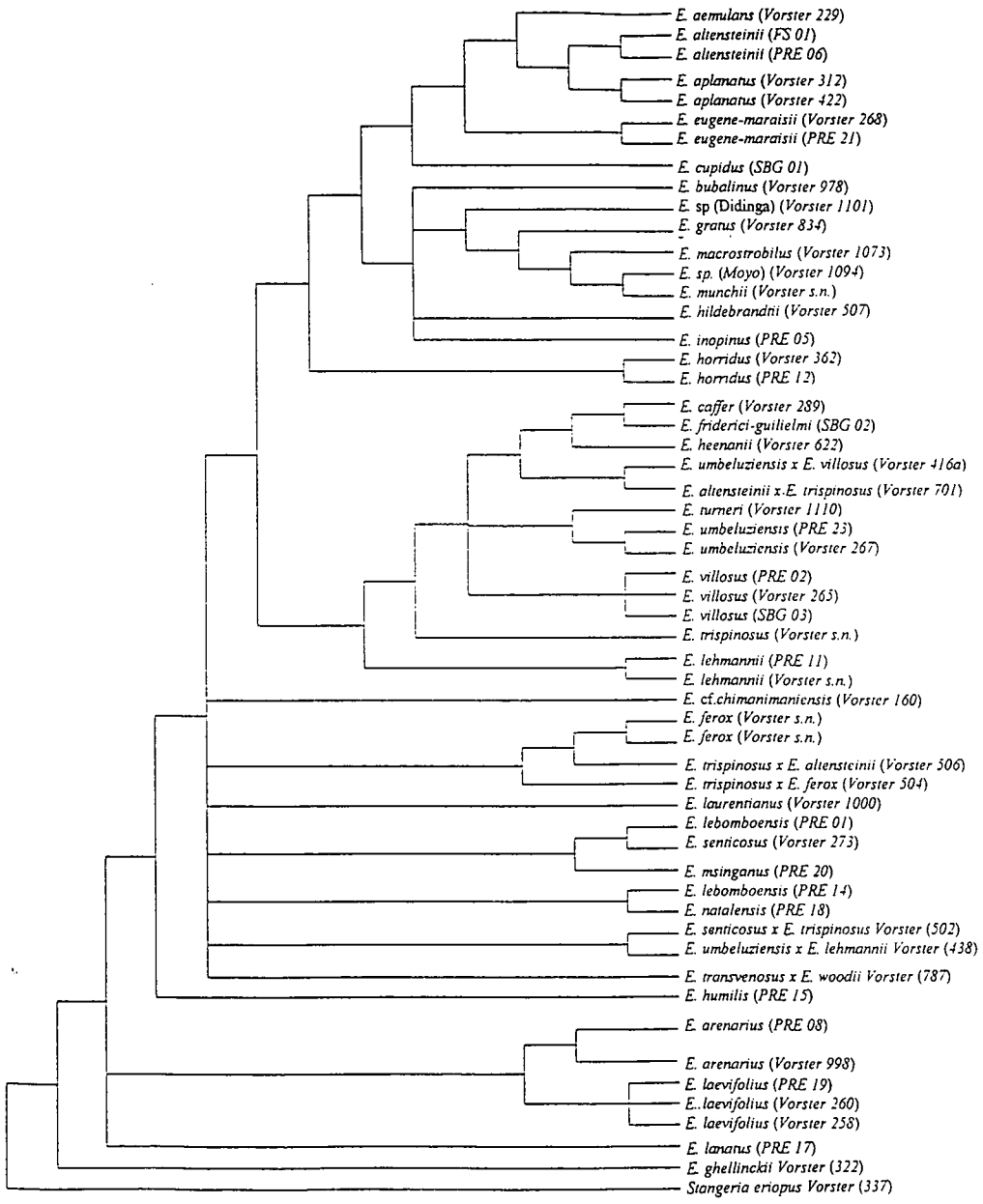


Figure 4.6 A Strict consensus cladogram from 1296 equally parsimonious cladograms, CI of 0.17, RI of 0.53, computed after a heuristic search of the RAPD data, including the hybrid specimens.

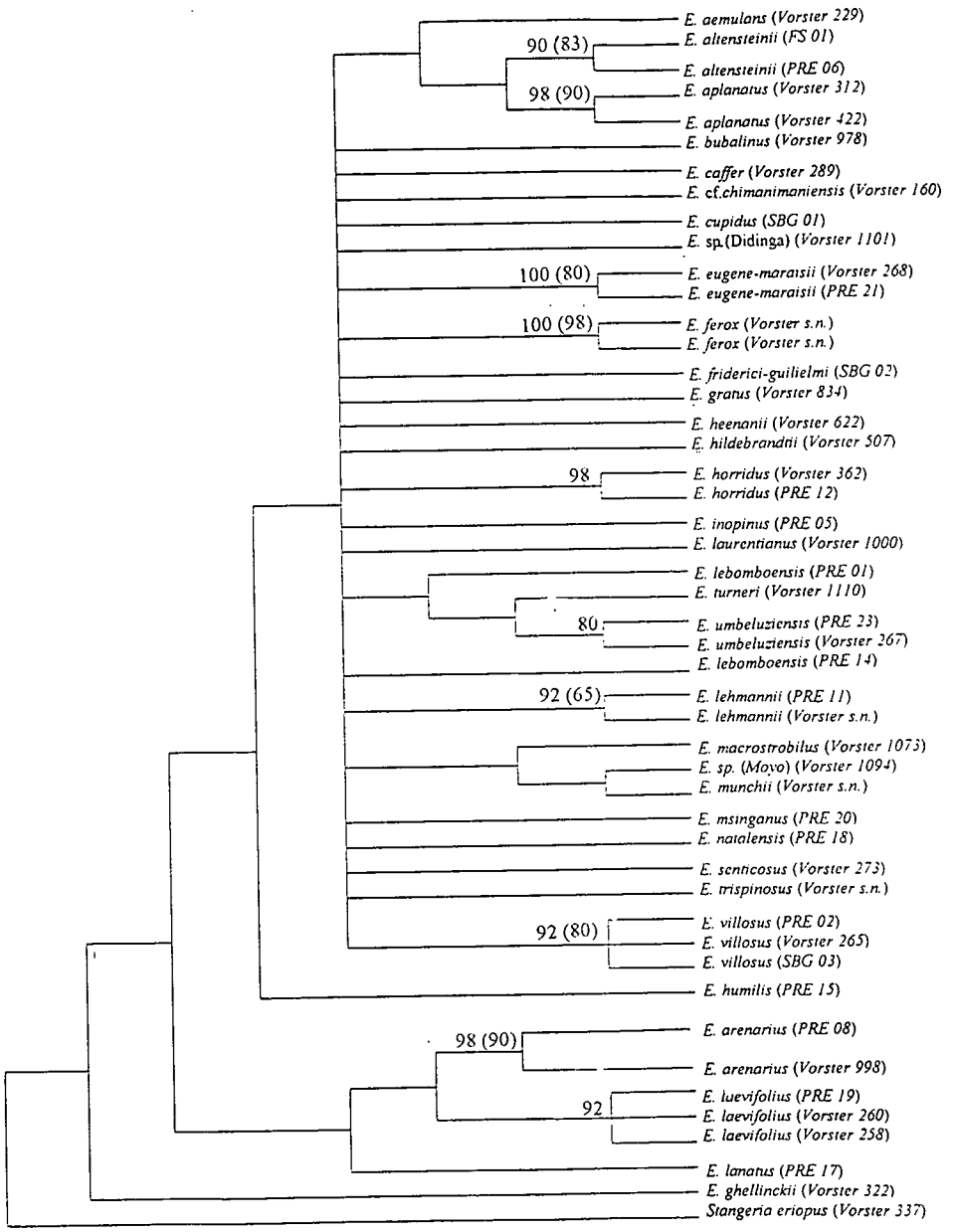


Figure 4.7 A Strict consensus cladogram from 6003 equally parsimonious cladograms, 319 steps, CI of 0.19, RI of 0.54, computed after the hybrid specimens were removed from the study. The bootstrap values and jackknife values (in brackets) are indicated on the relevant branches.

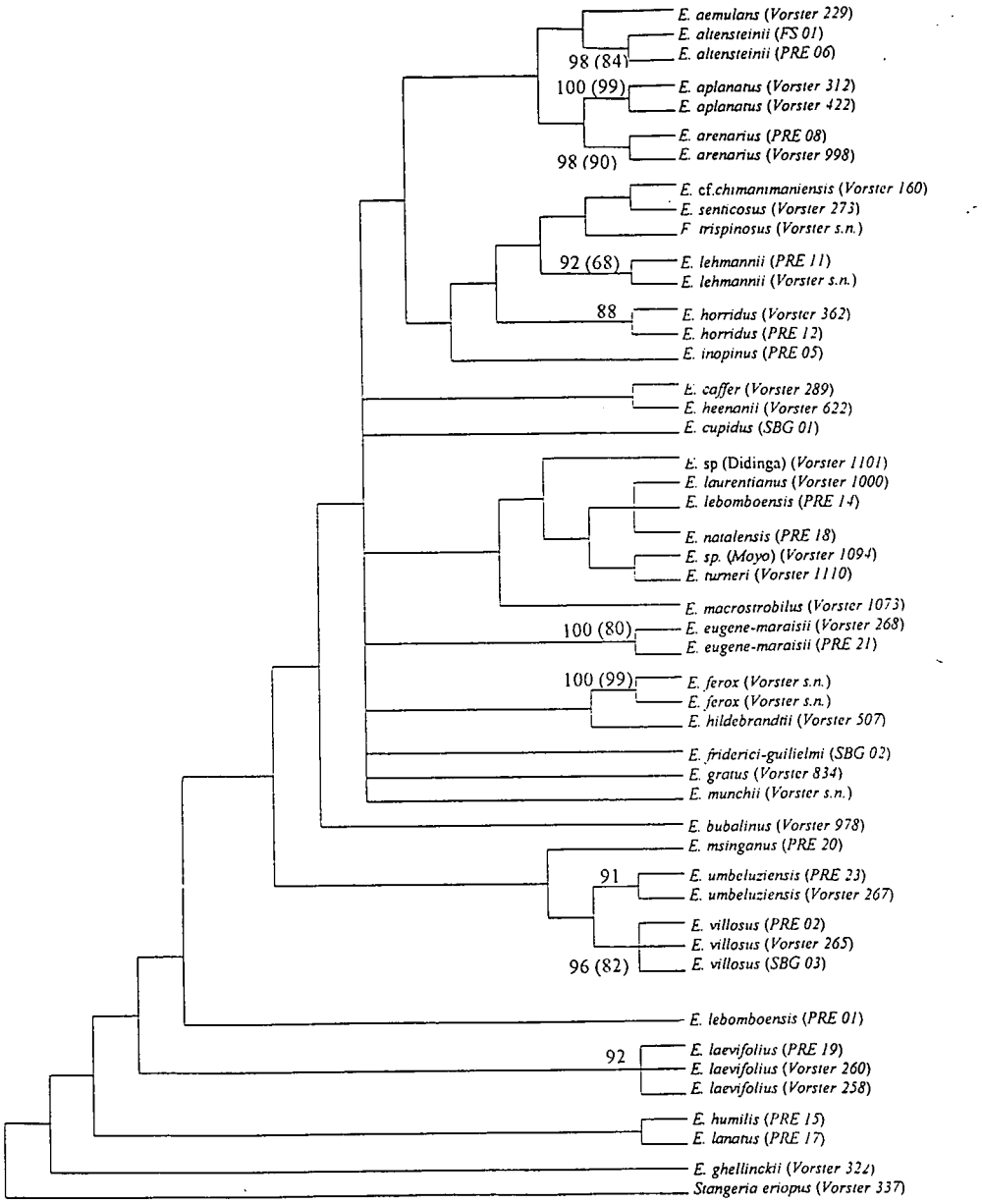


Figure 4.8 A Strict consensus cladogram computed after the removal of uninformative characters, representing 35 equally parsimonious cladograms of 275 steps, CI of 0.17 and RI of 0.56. Bootstrap values and jackknife values (in brackets) are indicated on the relevant branches.

Osborne *et al.* (1990), grouped all but four *Encephalartos* species, based on a numerical phenetic study, into five groups: the ALTENSTEINII, LEHMANNII, VILLOSUS, EUGENE-MARAISII, and CYCADIFOLIUS groups. In this study all these groupings were not observed, but some similarities were seen, such as the close relationship between *E. horridus*, *E. lehmannii* and *E. trispinosus* in the LEHMANNII group; between *E. umbeluziensis* and *E. villosus* in the VILLOSUS group; and *E. humulis* and *E. lanatus* in the CYCADIFOLIUS group.

The pairwise genetic distance (D), is based on the proportion of shared fragments between two specimens or species. The smaller the D value, the smaller the genetic distance between the two, and the closer the genetic relationship. In this study the average genetic distance between the different species were determined (Appendix E). This ranges from 0 (*E. turneri* versus *E. umbeluziensis*) to 2.35 (*E. senticosus* and *S. eriopus*).

The average genetic distance between the outgroup (*Stangeria eriopus*) and all the ingroup specimens (*Encephalartos*) is 1.28. In this study the genetic distance between different specimens of the same species (not shown) such as *E. altensteinii* (*PRE 06* and *FS 01*), *E. lehmannii* (*PRE 11* and *Vorster s.n.*), and *E. villosus* (*PRE* 02 and *Vorster* 265), indicate that the infraspecific distances of these species are much smaller than the interspecific relationships. RAPD can, therefore, be used in distinguishing between specimens of the same species. These findings are also observed on the consensus cladogram (Figure 4.8), supported by high bootstrap and jackknife values. In other species, such as *E. natalensis* and *E. lebomboensis*, the infraspecific distance is bigger than the interspecific distance and the specimens are not clustered together in the same clades (Figure 4.8). According to morphological studies (Dyer 1965a, Vorster 1990), the geographical distribution influences some morphological characters and more than one variation of a species are sometimes observed. It is, therefore, necessary that more specimens per species, from a variety of geographical areas, are studied to determine if geographical distribution influences variations within a species.

The application of RAPDs for a phylogenetic and genetic distance analysis of the genus *Encephalartos*, were not as sucsessfull as expected. Although some of the results concure wih results of other studies, the limited number of RAPD primers used in this study failed to generate sufficient variability to differentiate within species. Additional infraspecific sampling, as well as an increased number of RAPD primers, are required to demonstrate within-species variation.

CHAPTER FIVE

DNA amplification fingerprinting analysis

5.1 Introduction

The amplification of anonymous genomes with arbitrary oligodeoxyribonucleotides have proved a versatile and universal method for detecting polymorphisms for phylogenetic analysis, as well as genetic mapping and population biology (Caetano-Anollés 1994).

DNA amplification fingerprinting use arbitrary oligodeoxyribonucleotide primers to amplify anonymous genomes and generate relatively complex amplification profiles (Caetano-Anollés *et al.* 1991). The technique is fast, needs only a small amount of template DNA, and is inexpensive compared to other techniques such as RFLP-based fingerprinting (Callahan *et al.* 1993). The DAF technique can be tailored to increase the levels of polymorphic DNA by coupling the technique to endonuclease digestion of template DNA or the amplification product (Caetano-Anollés *et al.* 1993), as well as amplification with more than one primer, known as multiplex DAF (Caetano-Anollés *et al.* 1991).

The aim of this study is to generate fingerprinting data using different variations the DAF technique for phylogenetic and genetic distance analysis of the genus *Encephalartos*.

5.2 Results

Eight primers: DAF2, DAF3, DAF5, DAF6, DAF10, K1, K2 and K3, generated an average of 169 characters per specimen. The DAF profiles produced by the primers are schematically illustrated (Figure 5.1 – 5.9). Very faint fragments or fragments with no repetition were not used in this analysis.

A binary data matrix created from the DAF profiles (Appendix G), was used to calculate the pairwise genetic distances (D) (Appendix H), and determine the phylogenetic relationships of the specimens. The initial analysis yielded 1315 equally parsimonious cladograms, but these cladograms showed no resolution. Excluding the hybrids and uninformative characters yielded 315 parsimonious cladograms of 734 steps, CI of 0.18 and RI of 0.53 (Figure 5.9).

5.3 Discussion

A Strict consensus cladogram (Figure 5.9) was created from the most parsimonious cladograms, because according to Anderberg & Tehler (1990), the Strict consensus is the only way to illustrate components which are common to all the equally parsimonious cladograms of an analysis. From the consensus cladogram, groupings similar to those of Osborne *et al.* (1990) were observed such as the close relationship between *E. umbeluziensis* and *E. villosus* in the VILLOSUS group. This grouping is widely distributed, least stable statistically, and usually have leaflets reduced to prickles towards the leaf bases (Osborne *et al.* 1990). Although *E. cupidus* and *E. eugene-maraisii* are split by *E. friderici-guilielmi* in the consensus cladogram the close relationship between *E. cupidus* and *E. eugene-maraisii* is also observed in the EUGENE-MARAISII group of Osborne *et al.* (1990). Vorster (1986) also grouped *E. cupidus* and *E. eugene-maraisii*; and *E. umbeluziensis* and *E. villosus* together, according to the external morphology of the fronds and cones.

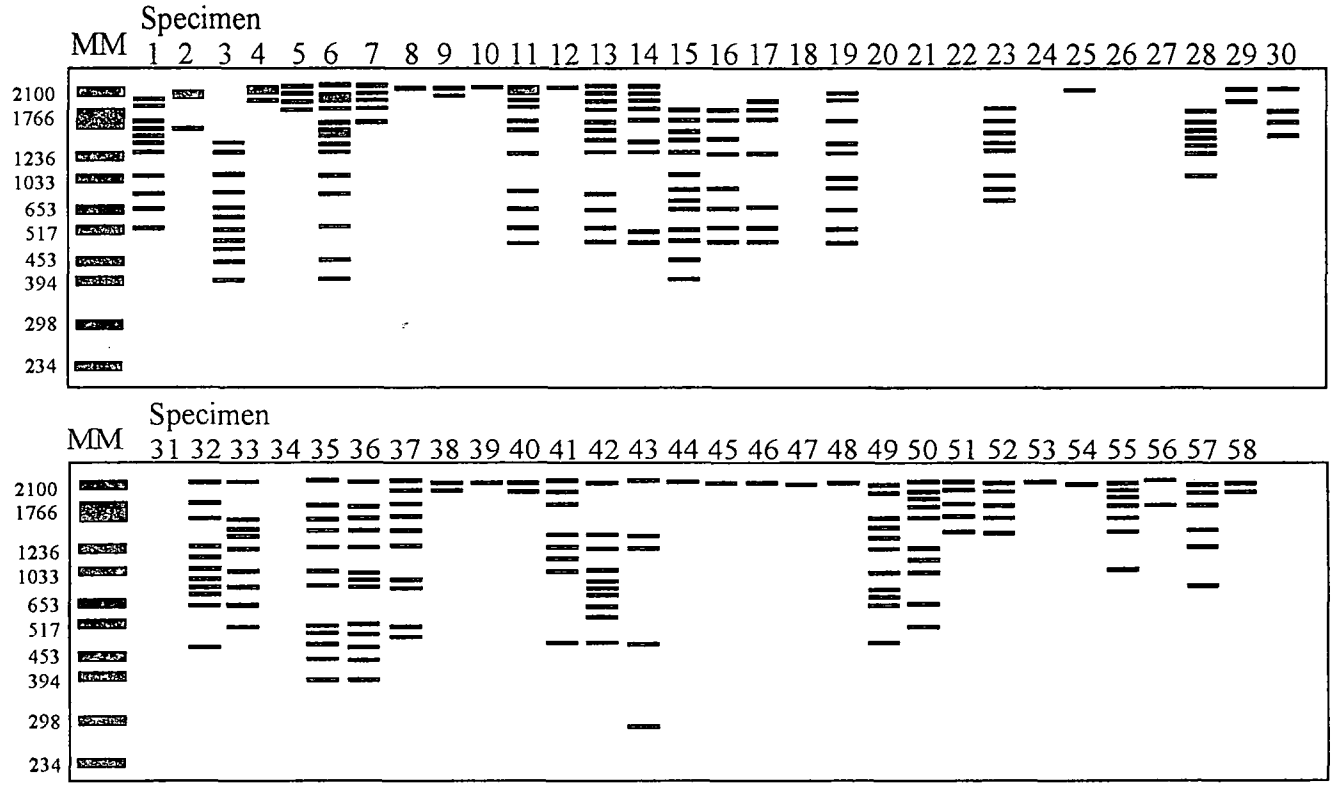


Figure 5.1 A schematic representation of the DAF profiles obtained after amplification with primer DAF 2. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

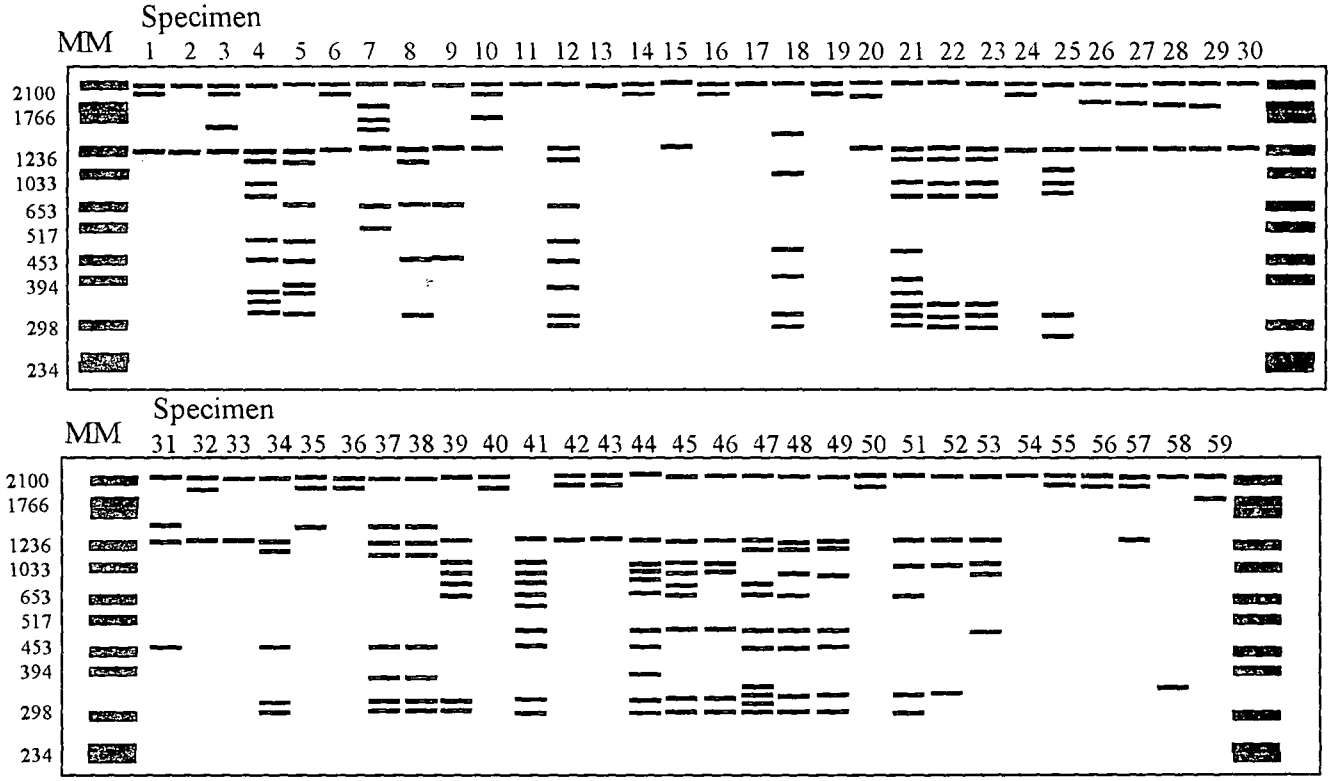


Figure 5.2 A schematic representation of the DAF profiles obtained after amplification with primer DAF 3. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

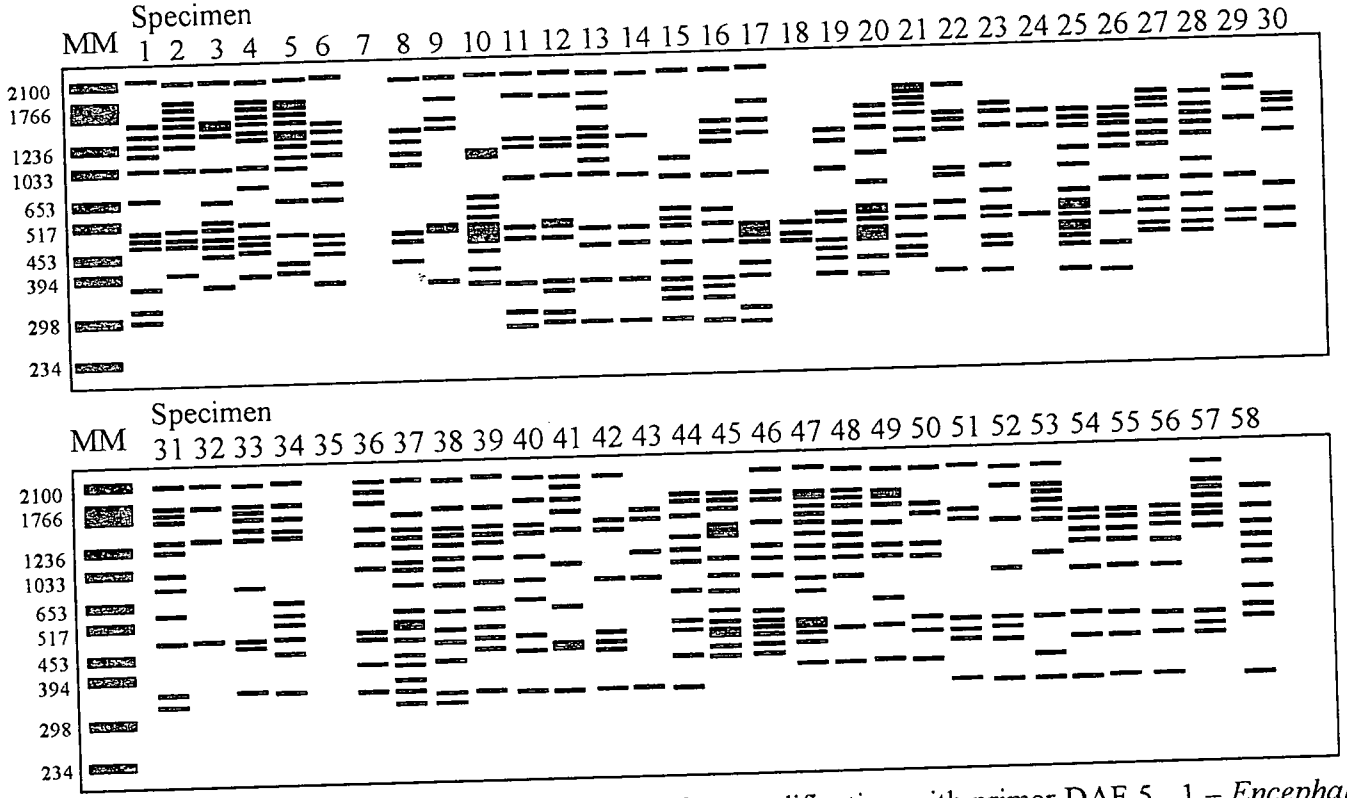


Figure 5.3 A schematic representation of the DAF profiles obtained after amplification with primer DAF 5. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

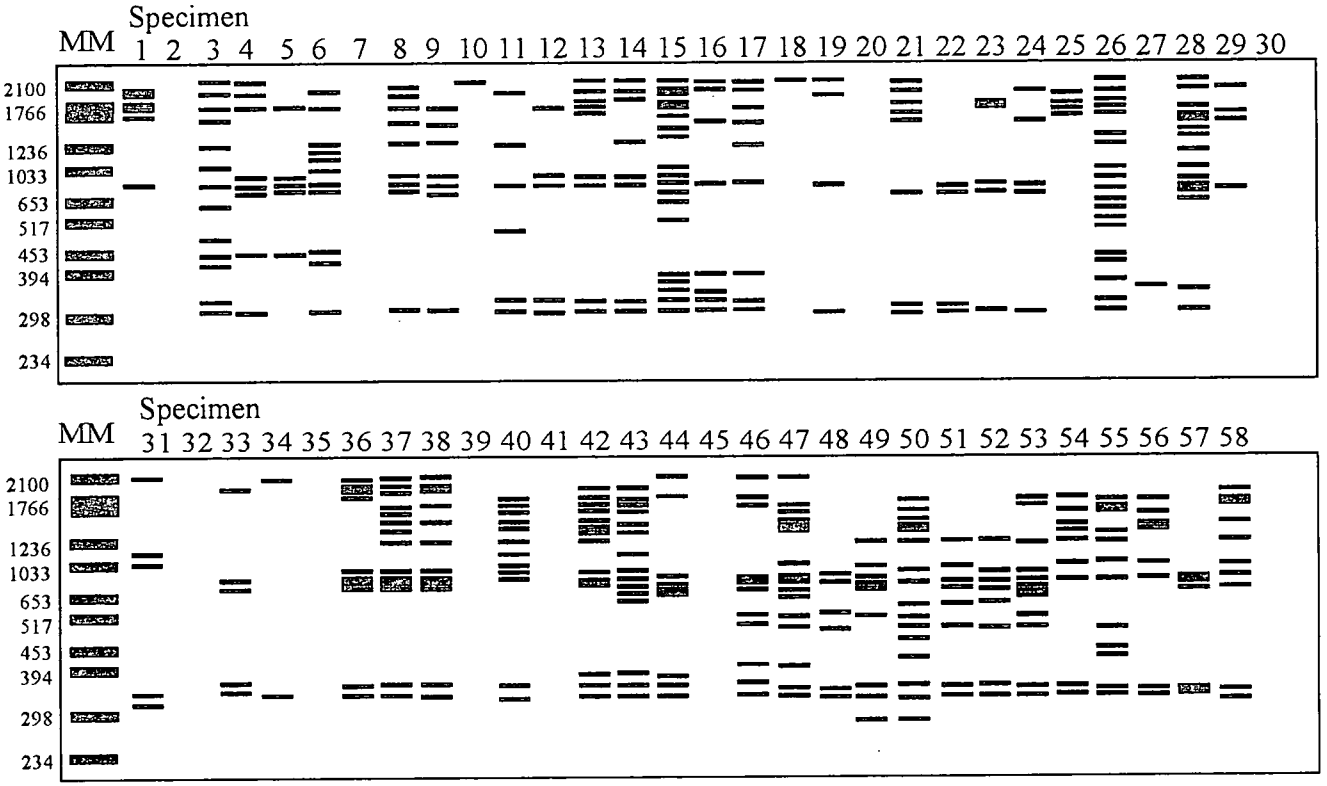


Figure 5.4 A schematic representation of the DAF profiles obtained after amplification with primer DAF 6. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

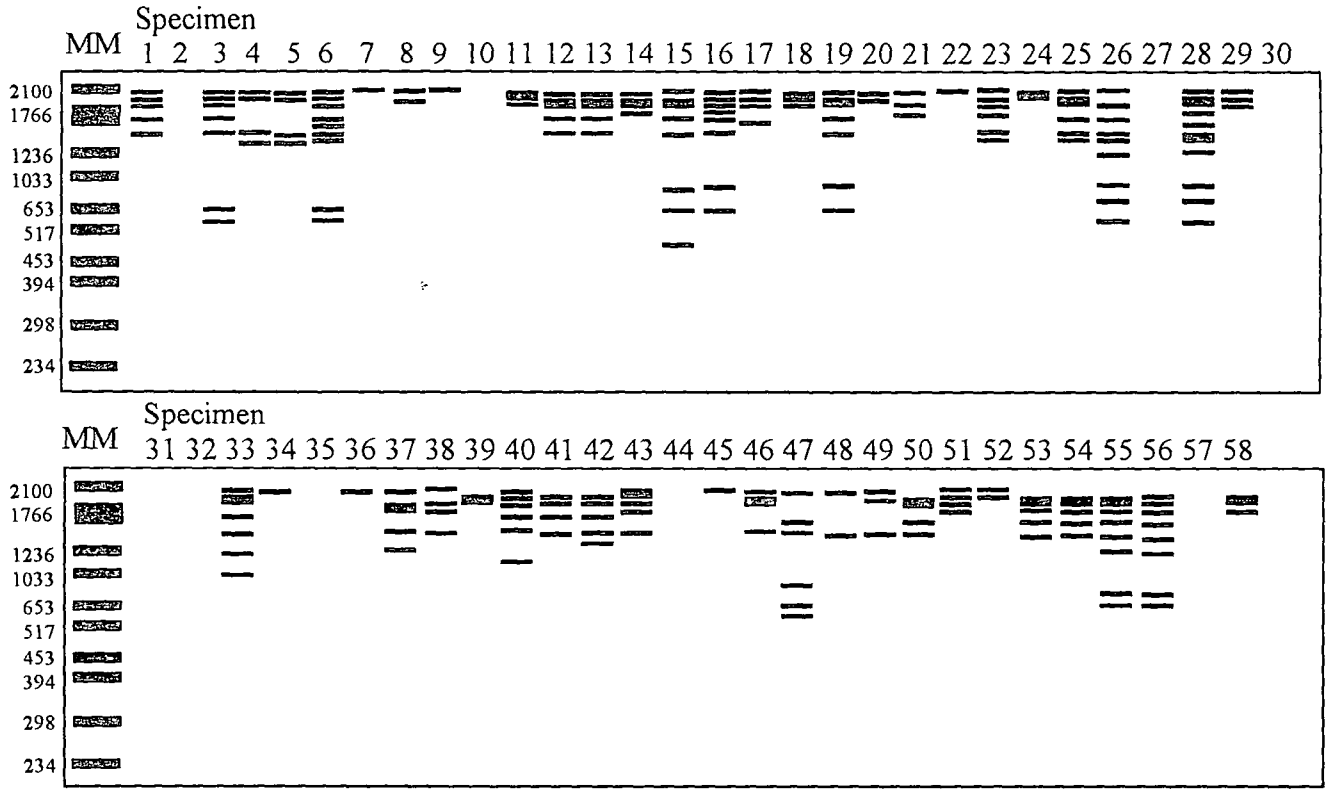


Figure 5.5 A schematic representation of the DAF profiles obtained after amplification with primer DAF 10. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

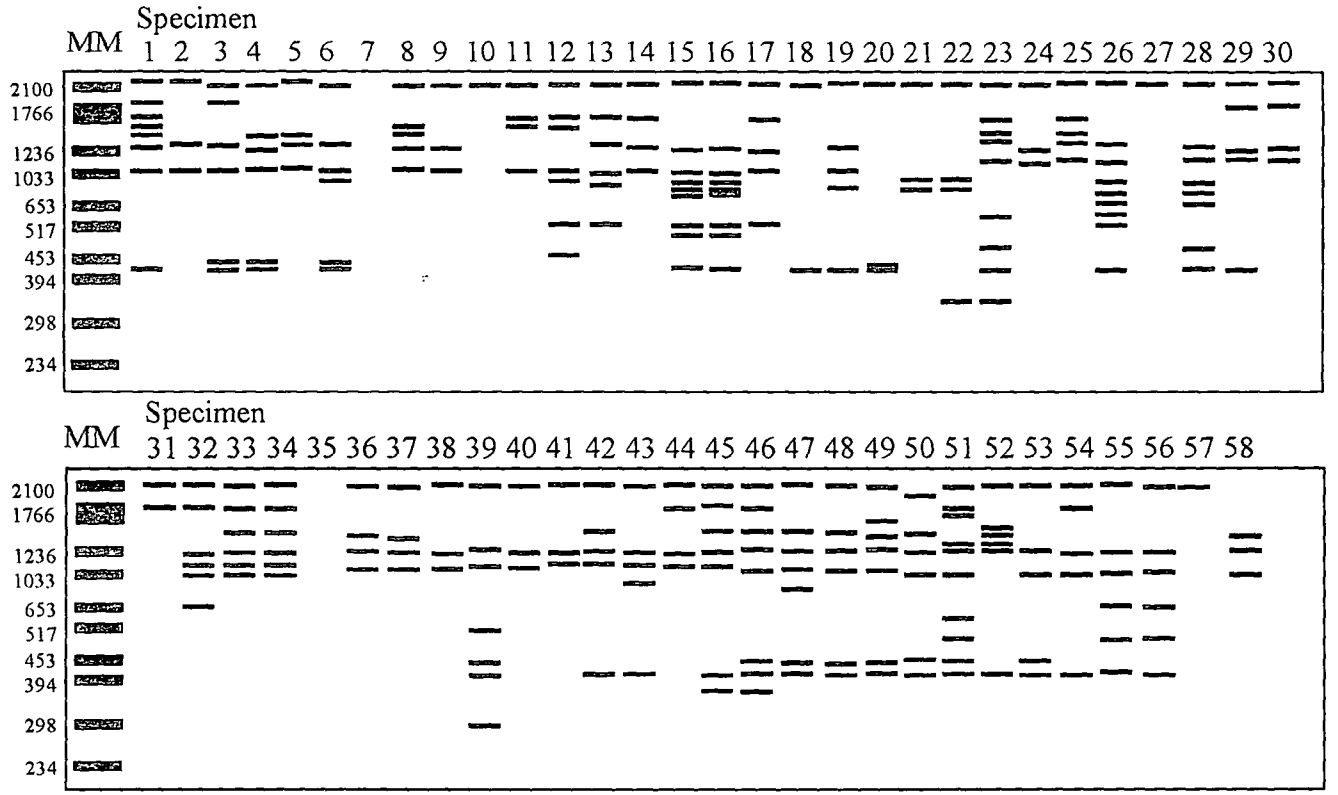


Figure 5.6 A schematic representation of the DAF profiles obtained after amplification with primer K1. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

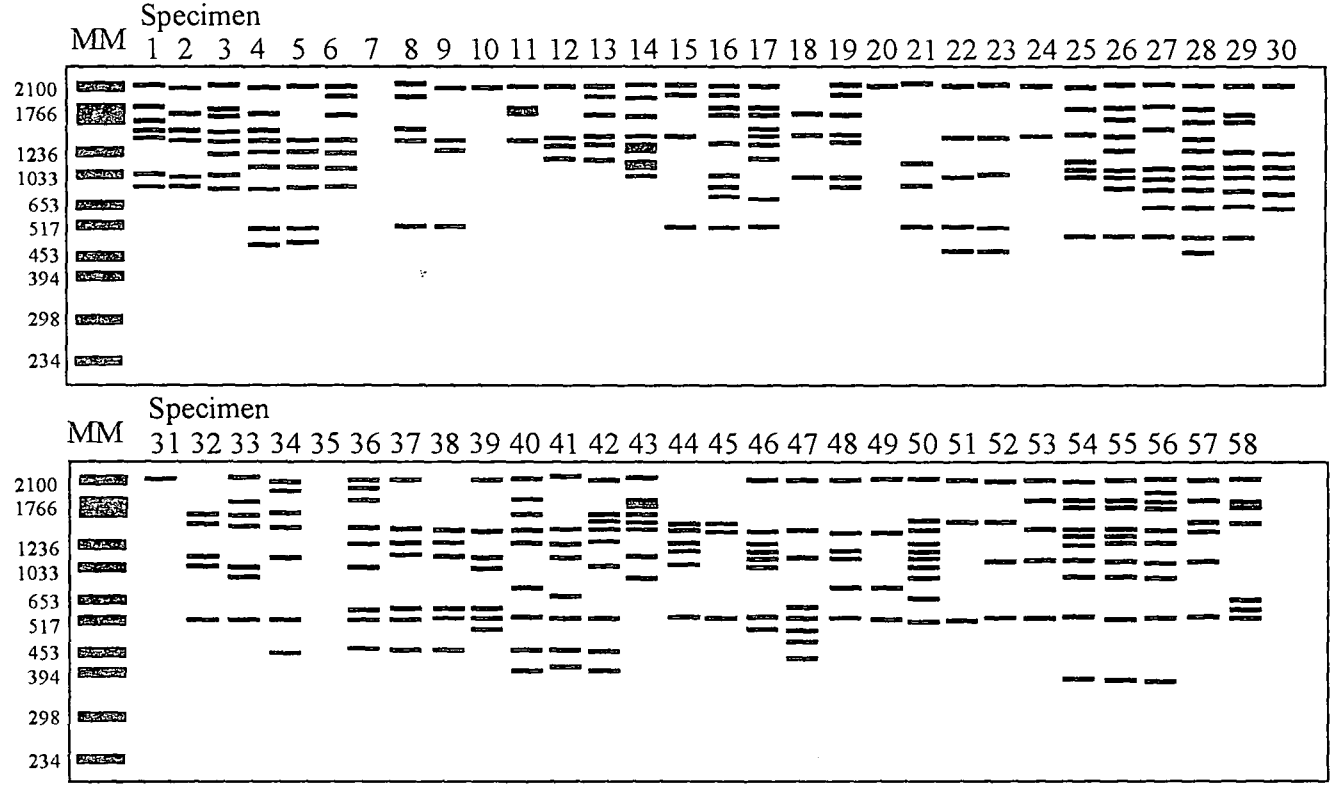


Figure 5.7 A schematic representation of the DAF profiles obtained after amplification with primer K2. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

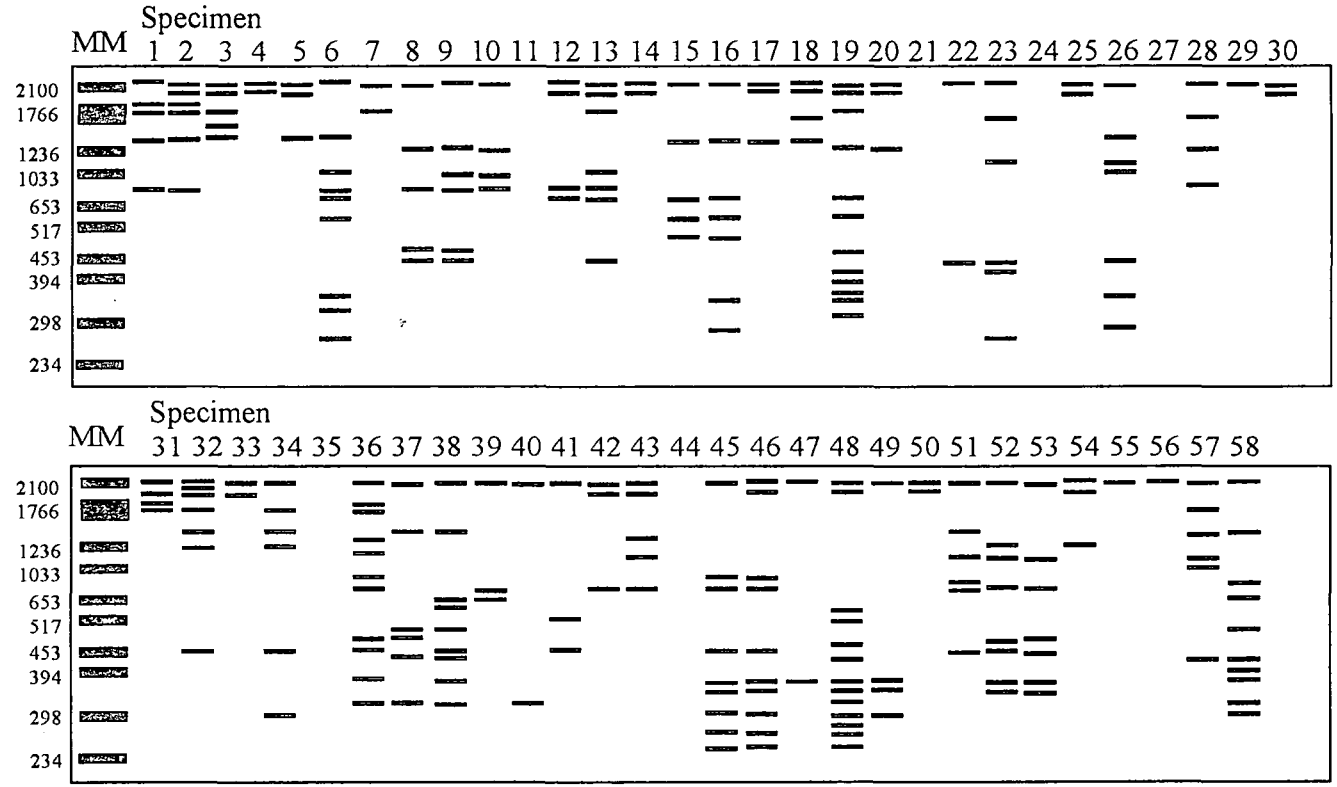


Figure 5.8 A schematic representation of the DAF profiles obtained after amplification with primer K3. 1 – *Encephalartos aemulans* (*Vorster 229*). 2 – *E. altensteinii* (*FS 01*). 3 – *E. altensteinii* (*PRE 06*). 4 – *E. aplanatus* (*Vorster 422*). 5 – *E. aplanatus* (*Vorster 312*). 6 – *E. arenarius* (*PRE 08*). 7 – *E. arenarius* (*Vorster 998*). 8 – *E. bubalinus* (*Vorster 978*). 9 – *E. caffer* (*Vorster 289*). 10 – *E.* cf.*chimanimaniensis* (*Vorster 160*). 11 – *E. cupidus* (*SBG 01*). 12 – *E.* sp. (Didinga) (*Vorster 1101*). 13 – E. eugene-maraisii (PRE 21). 14 – *E. eugene-maraisii* (*Vorster 268*). 15 – *E. ferox* (*Vorster s.n.*). 16 – *E. ferox* (*Vorster s.n.*). 17 – *E. friderici-guilielmi* (*SBG 02*). 18 – *E. ghellinckii* (*Vorster 322*). 19 – *E. gratus* (*Vorster 834*). 20 – *E. heenanii* (*Vorster 622*). 21 – *E. hildebrandtii* (*Vorster 507*). 22 – *E. horridus* (*PRE 12*). 23 – *E. horridus* (*Vorster 362*). 24 – *E. humilis* (*PRE 15*). 25 – *E. inopinus* (*PRE 05*). 26 – *E. laevifolius* (*PRE 19*). 27 – *E. laevifolius* (*Vorster 260*). 28 – *E. laevifolius* (*Vorster 258*). 29 – *E. lanatus* (*PRE 17*). 30 – *E. lanatus* (*Vorster s.n.*). 31 – *E. laurentianus* (Vorster 1000). 32 – *E. lebomboensis* (*PRE 14*). 33 – *E. lebomboensis* (*Vorster 313*). 34 – *E. lebomboensis* (*PRE 01*). 35 – *E. lehmannii* (*PRE 11*). 36 – *E. lehmannii* (*Vorster s.n.*). 37 – *E. macrostrobilus* (*Vorster 1073*). 38 – *E.* sp. (Moyo) (*Vorster 1094*). 39 – *E. munchii* (*Vorster s.n.*). 40 – *E. natalensis* (*PRE 18*). 41 – *E. msinganus* (*PRE 20*). 42 – *E. senticosus* (*Vorster 273*). 43 – *E. trispinosus* (*Vorster s.n.*). 44 – *E. turneri* (*Vorster 1110*). 45 – *E. umbeluziensis* (*PRE 23*). 46 – *E.* umbeluziensis (*Vorster 267*). 47 – *E. villosus* (*PRE 02*). 48 – *E. villosus* (*Vorster 265*). 49 – *E. villosus* (*SBG 03*). 50 – *E.* cf. *woodii* (*Vorster* 682). 51 – E. *senticosus x E. trispinosus* (*Vorster 502*). 52 – *E. umbeluziensis x E. lehmannii* (*Vorster 438*). 53 – *E. umbeluziensis x E. villosus* (Vorster 416a). 54 – *E. altensteinii x E. trispinosus* (*Vorster 701*). 55 – *E. trispinosus x E. altensteinii* (*Vorster 506*). 56 – *E. trispinosus x E.* ferox (Vorster 504). 57 – *E. transvenosus x E. woodii* (*Vorster 787*). 58 – *Stangeria eriopus* (*Vorster 337*).

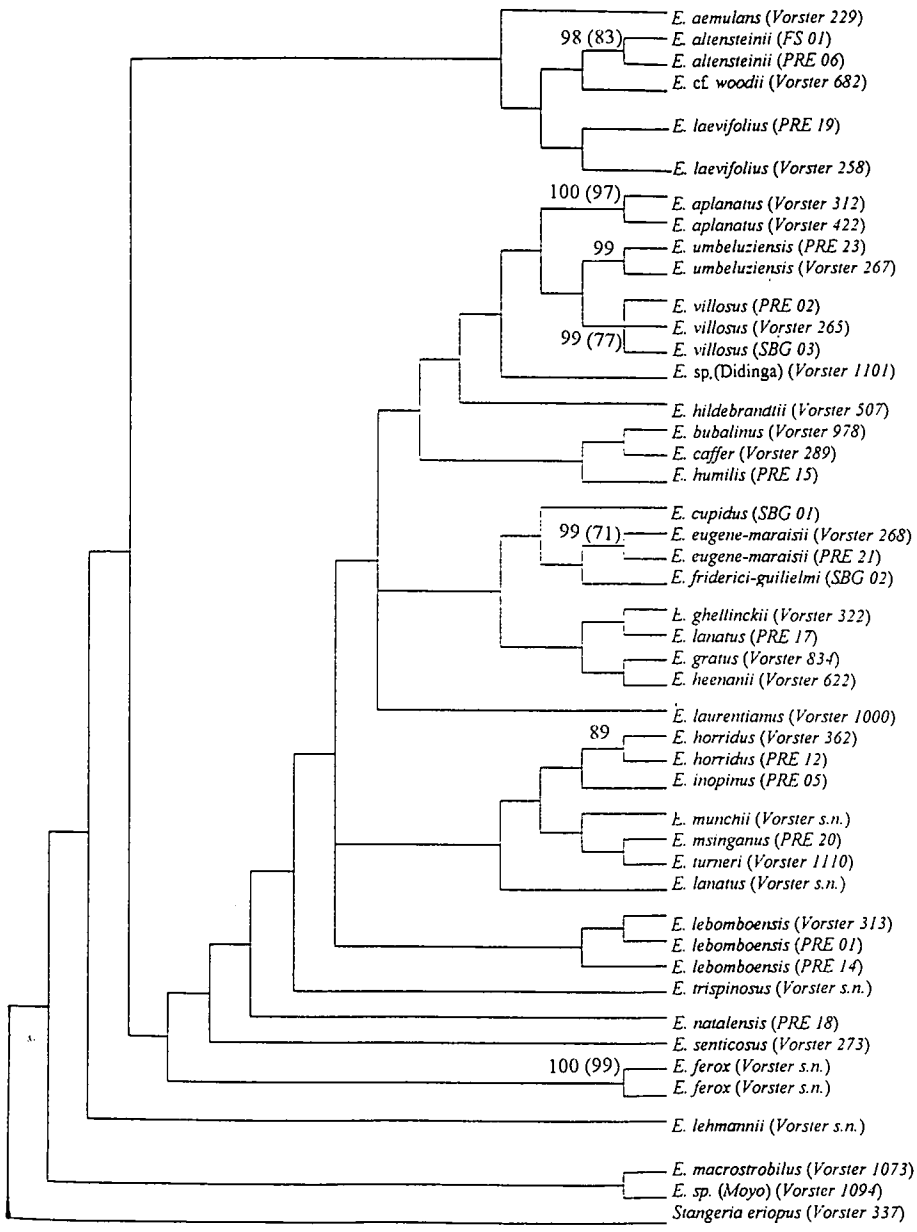


Figure 5.9 A Strict consensus cladogram from 1092 equally parsimonious cladograms of 734 steps, CI of 0.18 and RI of 0.53. The bootstrap values and jackknife values (in brackets) are indicated on the relevant branches.

Siddall (1995) suggested that the jackknife monophyly index is the preferred technique, because, unlike the bootstrap, it is not biased in its value by the influences of uninformative characters. Both of the methods were included in the phylogenetic analysis. The bootstrap values were higher than the jackknife values, but to make the variance among jackknife estimates as large as that among bootstrap estimates, one would have to engage in an extrapolation to make their variance larger (Felsenstein 1985). In this study the basal clades were not supported, but the bootstrap and jackknife monophyly procedures supported infraspecific relationships of specimens allocated to *E. altensteinii*, *E. eugene-maraisii* and *E. ferox*.

The infraspecific genetic distance within species such as *E. altensteinii* (*PRE 06* and *FS 01*), *E. horridus* (*PRE 12* and *Vorster 362*) and *E. laevifolius* (*PRE 19* and *Vorster 260*) is small and clustered together (Figure 5.9). The interspecific genetic distance ranges from 0.02 {*E. laevifolius* (*Vorster* 258) versus *E. lanatus* (*PRE 17*)} to 1.67 {*E. ghellinchii* (*Vorster 322*) versus *E. lehmannii* (*PRE 11*)}. For this study *Stangeria eriopus* (*Vorster 337*) was again used as outgroup. The average genetic distance between *Stangeria eriopus* and the ingroup is 0.63, and the former constitutes a sister group to the monophyletic *Encephalartos* ingroup.

The number of specimens per species were to small to do a detailed analyses of the infraspecific relationship in species. An increase in primers used will also assist in this problem. In conclusion, the DAF techniques were successfully applied in the generating of fingerprints for phylogenetic and genetic distance analysis of the genus *Encephalartos*.

CHAPTER SIX

PHYLOGENY

6.1 Introduction

Phylogeny is the branching history of routes of inheritance (Maddison & Maddison 1992). There are many methods for reconstructing evolution, but according to Lipscomb (1998) they can be divided into three types, (1) overall similarity or phenetic methods, (2) evolutionary methods and (3) cladistic or parsimony methods. In this study cladistics or parsimony, also known as phylogenetic systematics were used.

Two categories of data were used in this study: discrete characters, and similarities or distance data. Discrete characters provided data about individual specimens and the computer program PAUP (Swofford 1998) was used for the analysis thereof (Swofford & Olsen 1990). A similarity or distance value is a comparison of two specimens and describes the pairwise relationship (Swofford & Olsen 1990), and the method of Nei & Li (1979) was use to calculate these values.

The aim of this study was to estimate the phylogenetic relationships for the specimens used when combining RAPD and DAF data, using the method of parsimony and genetic distance analysis. The impact of hybrids on cladistic analysis was also investigated, including whether or not hybrids behave differently in cladistic analysis than species, such that they might be detectable.

6.2 Results

The combined data for 58 specimens consist of 231 characters generated from a total of 13 primers. The average genetic distances between the different species were determined (Appendix I). A heuristic search, using only specimens that amplified in both molecular techniques were used in a phylogenetic study. Initially all hybrids were excluded from the study, and *Stangeria eriopus* (*Vorster 337*) was used as an outgroup. Uninformative characters as well as characters with a RI:CI ratio less than 2 (Lipscomb 1998), were excluded from the study. The analysis produced 18 equally parsimonious cladograms of 1034 steps, CI 0.17 and RI of 0.52. The removal of excessively homoplasic characters after consulting a preliminary parsimony analysis, does not necessarily increase the consistency index in large studies (Sanderson & Donoghue 1989). High bootstrap and jackknife values of the terminal taxa indicate that the cladogram is well supported (Figure 6.1).

The impact of hybrids on cladistic topology was studied in two ways. First, the 'stability' of cladistic results was analysed by comparing the cladogram topology when the number of hybrids (one to seven) are increased in the analysis (Figure 6.2 a-f). The inclusion of the seven hybrids in a heuristic search produced 135 parsimonious cladograms, with a tree length of 1179 and CI 0.15 (Figure 6.2f)

A second method of studying the effect of hybrids on cladogram topology was done quantitatively by comparing the cladistic results of analysis, including hybrids versus species, by means of a heuristic search of a data set containing 44 specimens. This produced 9 equally parsimonious cladograms of 935 steps and CI 0.186. Six hybrids were then successively added to the data set and the number of parsimonious cladograms, tree length, and CI were calculated. To compare the results, the same analysis was done by adding six specimens to the same data set of 44 OTUs (Table 6.1).

Table 6.1 The quantitative comparison of the cladistic analysis with the inclusion of hybrids versus species to a data set.

Number of hybrid specimens	Tree length	Number of parsimonious cladograms	Consistency index
1	966	9	0.180
2	980	9	0.178
3	1008	9	0.173
4	1039	71	0.168
5	1053	9	0.165
6	1062	18	0.164
Number of non-hybrid specimens			
1	956	27	0.180
2	980	18	0.178
3	983	9	0.177
4	994	90	0.175
5	1012	9	0.172
6	1034	18	0.168

The hybrids were compared to the putative parents, and their placement was inspected in the parsimonious cladograms. The Adams consensus of these cladograms is shown (Figure 6.3).

The hybrids were compared to the putative parents, and their placement was inspected in the parsimonious cladograms. The Adams consensus of these cladograms is shown (Figure 6.3).

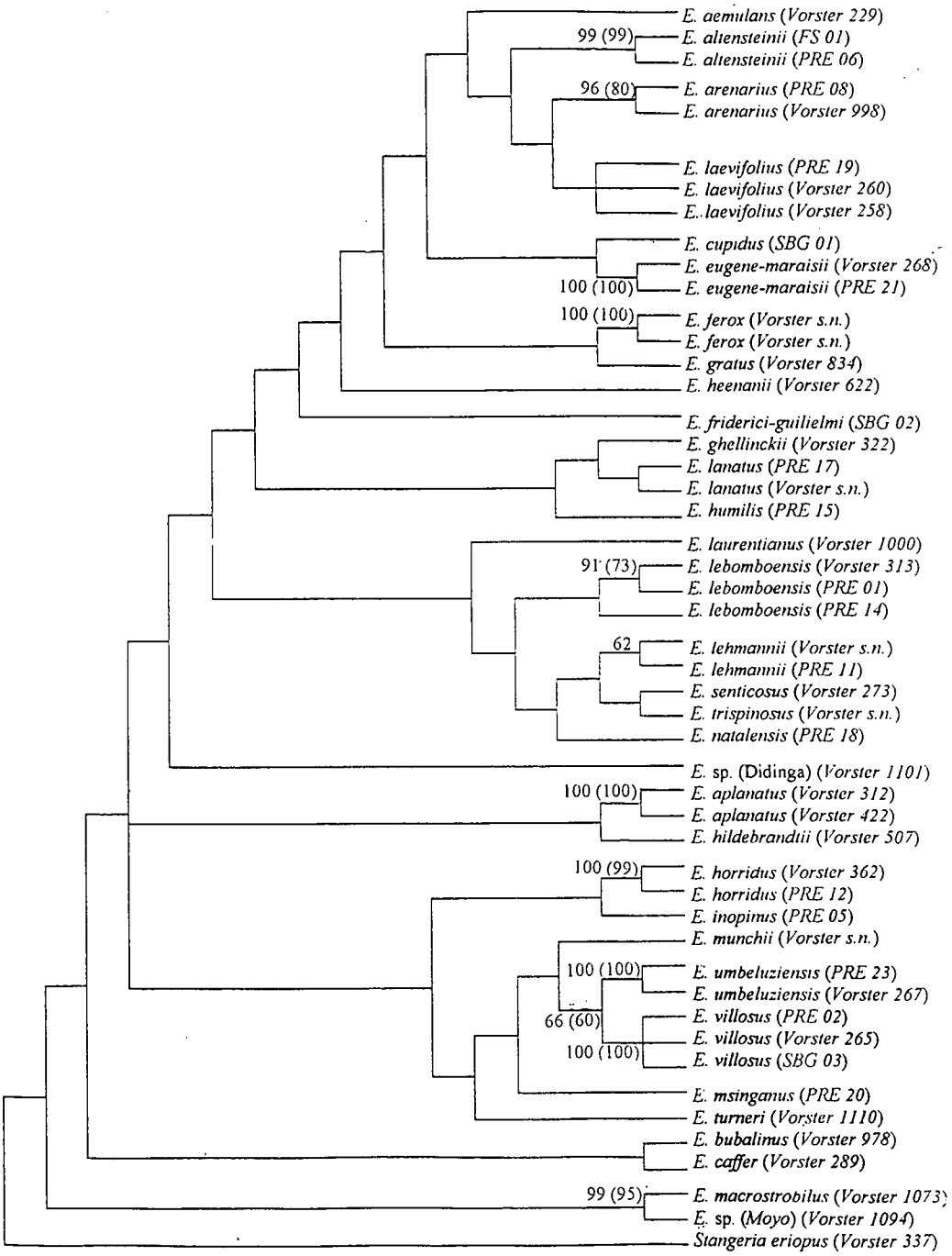


Figure 6.1 A Strict consensus cladogram of 18 equally parsimonious cladograms of 1034 steps, CI of 0.17 and RI of 0.52. This cladogram represents a heuristic search of the combined analysis, with bootstrap values and jackknife values (in brackets) indicated on the relevant branches.

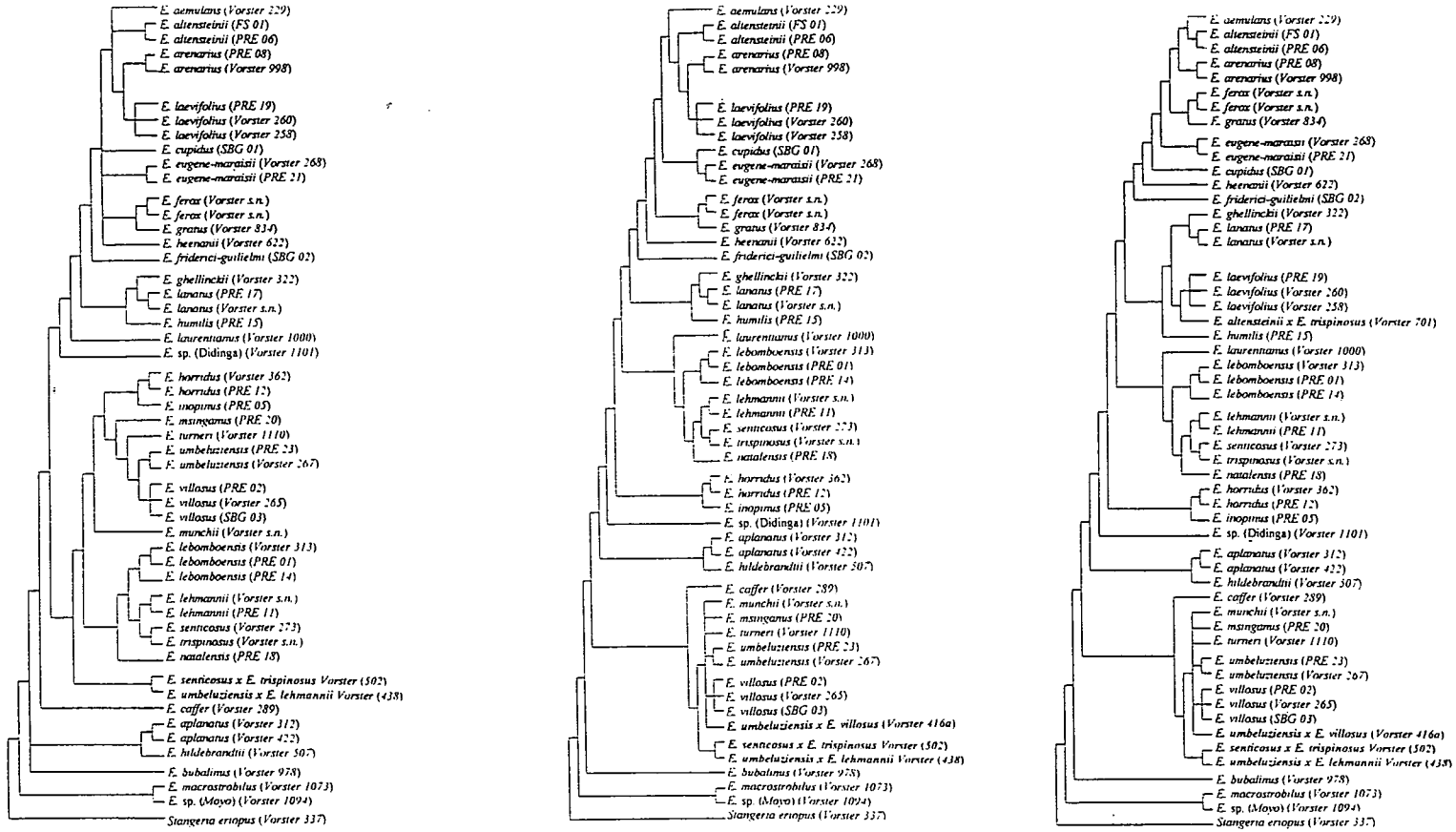


Figure 6.2a Strict consensus cladograms indicating the effect, an increased number of hybrids (two – four hybrids) have on the tree topology during cladistic analysis.

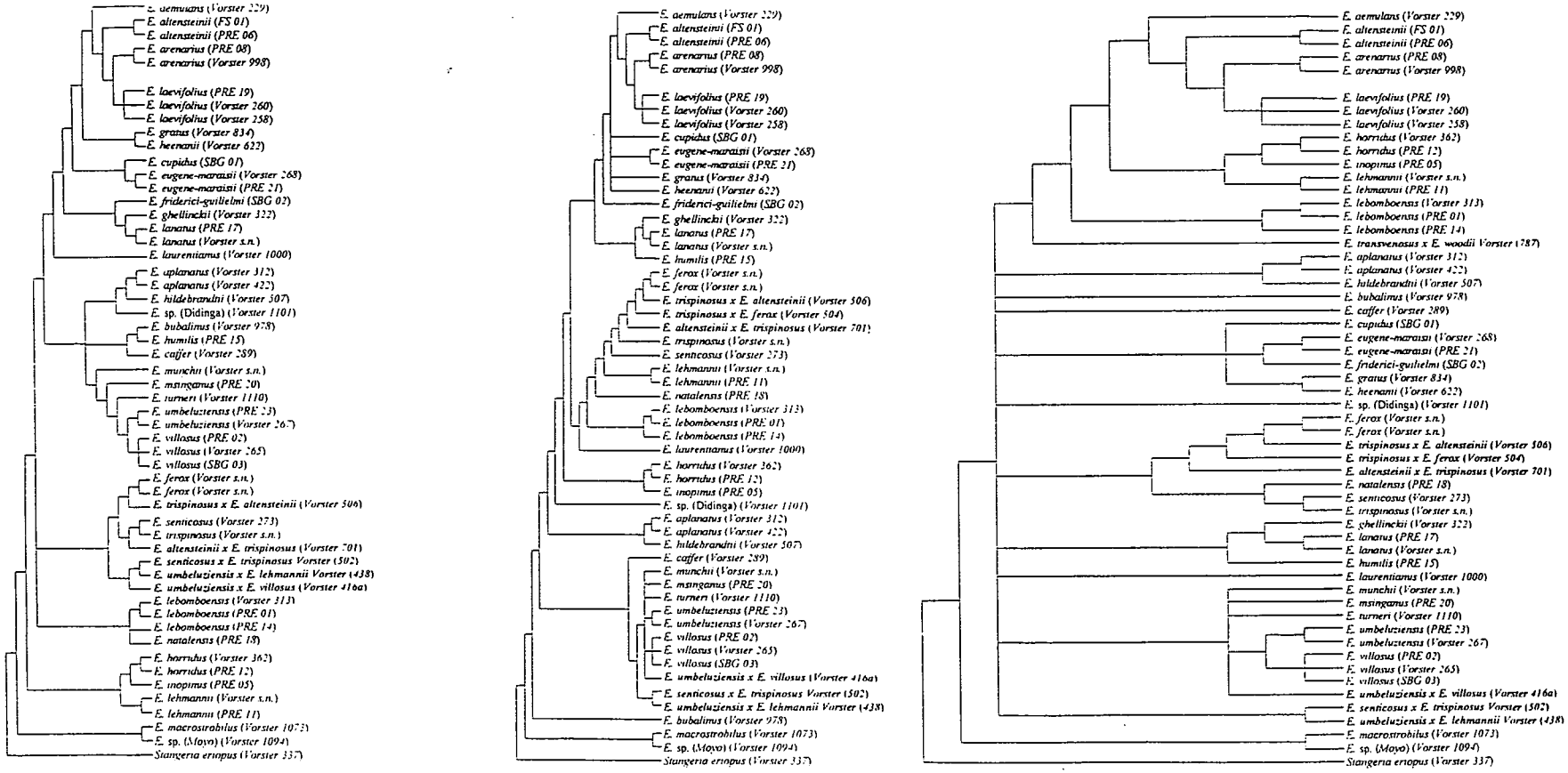


Figure 6.2b (continues) Strict consensus cladograms indicating the effect, an increased number of hybrids (five – seven hybrids) have on the tree topology during cladistic analysis.

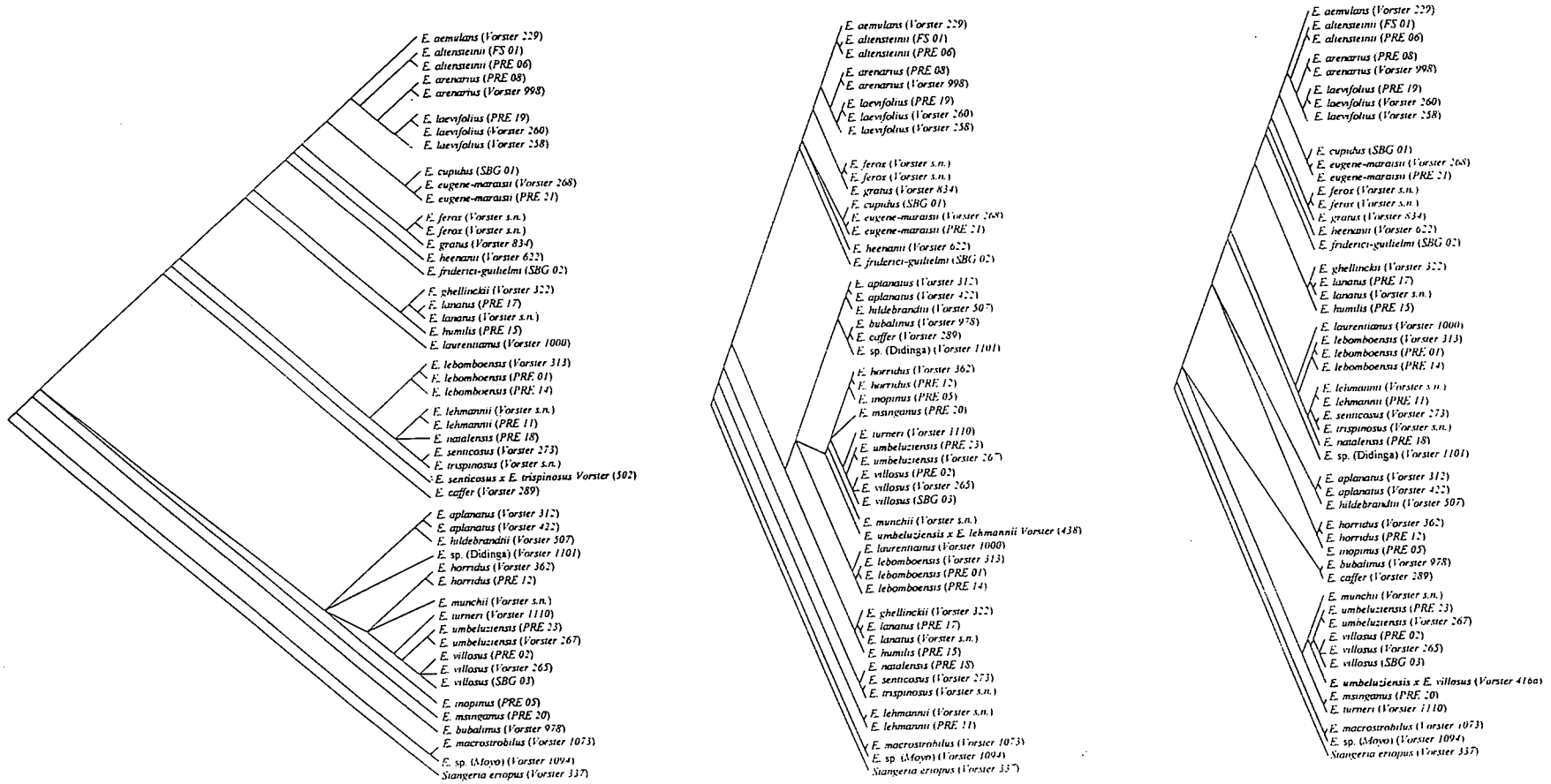


Figure 6.3a Adams consensus cladograms indicating the placement of hybrids (a – c) in a cladistic analysis.

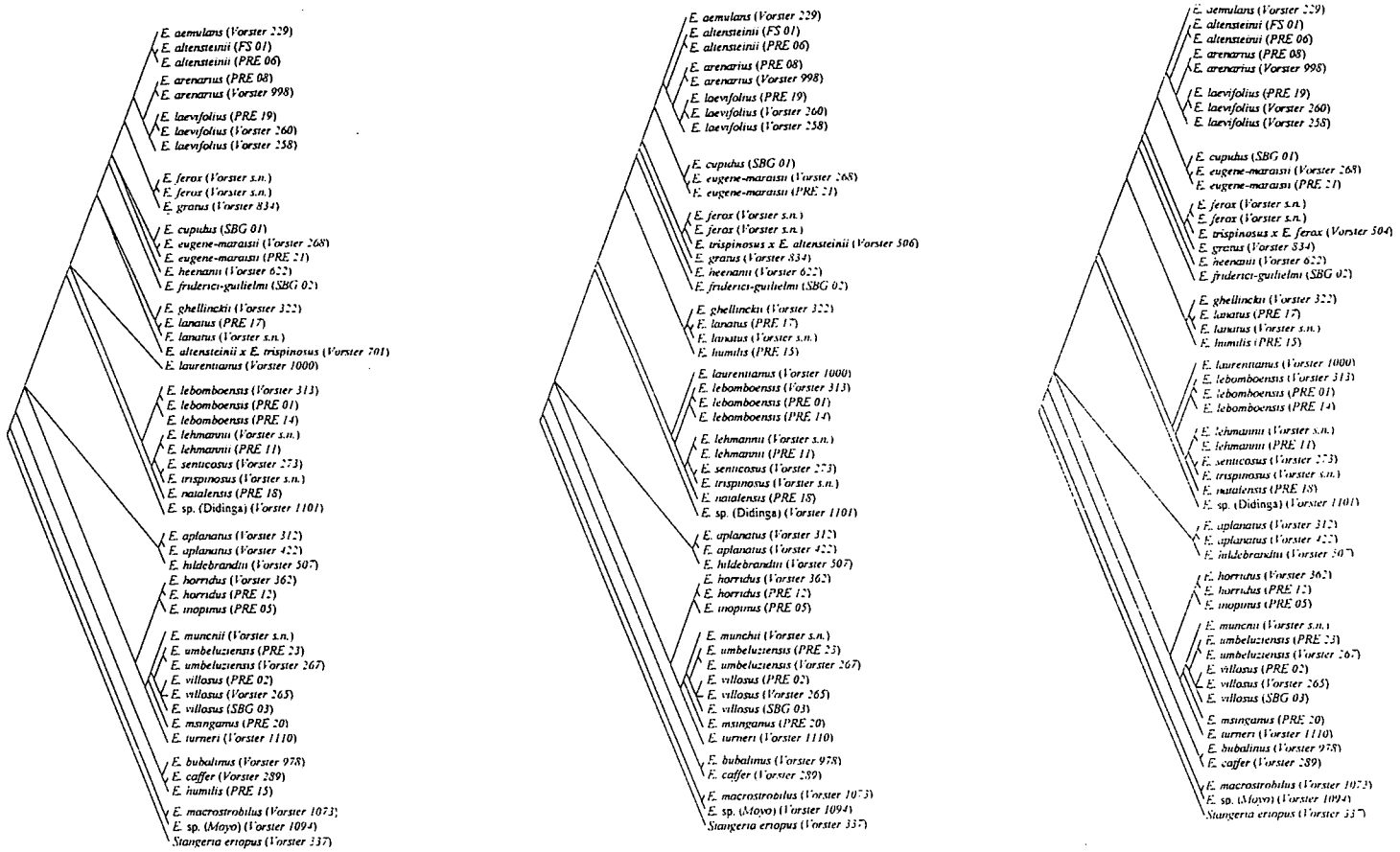


Figure 6.3b (continues) Adams consensus cladograms indicating the placement of hybrids (d – f) in a cladistic analysis.

6.3 Discussion

A combined analysis – Similar cladogram topologies observed with the RAPD, and combined data sets were between the *E. senticosus*, *E. trispinosus* and *E. lehmanni* grouping, with an increased resolution of the combined analysis' cladograms. Several small groupings are observed in the DAF and combined analyses, but this is caused by the larger number of DAF characters in the combined analyses. The *E. villosus* and *E. umbeluziensis* pair are grouped together in all three analyses. Congruence between studies is strong evidence that the underlying pattern has been discovered; because conflict may indicate theoretical or procedural problems in one or both of the analyses (Hillis 1987).

Astolfi *et al.* (1981) suggested that the accuracy of the reconstructed cladograms increases as the number of independent characters used increases, and the estimates of errors involved in reconstructing a cladogram, could be decreased by using a number of characters equal to five times the number of populations. The stability of phylogenies was also obtained with the increased number of primers (13) used for phylogenetic analysis. According to Landry & Lapointe (1996), the stability of a phylogeny can be reached with 12 primers.

The tree length of the combined data sets increased (RAPD: 275 steps, DAF: 734 steps and combined analysis: 1034 steps). However, Sanderson & Donoghue (1989) found that there are no efficient algorithms that always identify the shortest cladogram in large data sets and, therefore, larger studies may result in longer tree lengths. The CI stayed relatively constant in the combined analysis (RAPD: 0.17, DAF: 0.18, combined analysis: 0.16), as well as the RI (RAPD: 0.56, DAF: 0.53, and combined analysis: 0.52). The CI measures the relative number of homoplasy in a data set. The CI in this study is low, which indicates a high occurrence of character homoplasy. The RI measures the relative amount of synapomorphy expected from a data set that is retained as synapomorphy on a cladogram (Lipscomb 1998). A high RI

value gives, therefore, a better description of the cladogram, as cladistic analyses is based on shared synapomorphies between two specimens.

Tree topology – In cladistic analysis the emphasis is on monophyletic groups (groups that have arisen through diversification of a simple ancestor) involving homologous shared and derived characters (Woodland 1991), and the search for sister species or sister groups (Bremer & Wanntorp 1978). The cladogram (Figure 6.1) shows that the ingroup specimens are a monophyletic group and *Stangeria eriopus* a sister group. The average genetic distance between *S. eriopus* and the ingroup specimens is 0.93. Further similarities of the results are observed in collation with other studies (morphology, anatomy, geographical distribution) such as:

- The close molecular relationship between *Encephalartos cupidus* and *E. eugene-maraisii* is well supported by the external morphology of the fronds and cones (Vorster 1986), and the anatomy of the pinnae and geographical distribution (Koeleman *et al.* 1981). These species are also grouped in the EUGENE-MARAISII group by Osborne *et al.* (1990), according to their amphistomatic leaflets with low stomatal densities.

- *Encephalartos humilis*, *E. lanatus,* and *E. ghellinckii* have linear, very narrow leaflets, but species such as *E. friderici-guilielmi* and *E. laevifolius* would also be expected to be in this clade (Vorster 1990).

- *Encephalartos umbeluziensis* and *E. villosus* are presumed to be closely related with similar pinnae anatomy, distribution, and the morphology of the fronds and cones (Koeleman *et al.* 1981, Vorster 1986). Bootstrap (66) and jackknife (60) values, as well as a small genetic distance (0.07), support this sister group relationship. These species, as well as *E. inopinus* and *E. munchii* from the same clade, are grouped in the VILLOSUS group of Osborne *et al.* (1990).

- The close molecular relationship between *Encephalartos friderici-guilielmi* and *E. altensteinii* observed by Van der Bank *et al.* (1998),

was confirmed in this study, but as Van der Bank *et al.* (1998) pointed out, this relationship is not supported by morphological data.

- A sister species relationship between *E. macrostrobilus* and *E.* sp. (Moyo) is well supported by bootstrap (99) and jackknife (95) values.

The plight of hybrids – Theoretically hybrids cause character conflict in cladograms and, therefore, affect the reconstruction of phylogenies (Funk 1985). According to Humphries & Funk (1984) there are three approaches for dealing with hybridisation in phylogenetic reconstruction:

- using the most parsimonious cladogram(s) and leaving homoplasic characters resulting from the presence of hybrids, but sometimes this does not reflect accurate sister-group relationships,

- removing hybrids that have been identified by their intermediancy at the beginning of the analysis, but one can not always identify the hybrids prior to analysis,

- leaving all the taxa in the analysis and examining cladograms for polytomies (nodes with more than two branches) and character conflict that may indicate possible hybrids, but this relies on the belief that hybrids cause polytomies, which does not always happen, according to Humphries (1983).

In this study all three methods were applied but it was found that exclusion of the identified hybrids prior to analysis is the most time effective method.

The effect of hybrids on cladogram topology – Nelson (1983), Wanntorp (1983) and Funk (1985) predicted that hybrids will behave in an unstable manner in cladograms, and should, therefore, have an effect on the cladogram topology. The effect of hybrids on cladogram topology were studied by gradually increasing the number of hybrids (from one to seven) to the data set (Figure 6.2).

Inclusion of one hybrid specimen, *Encephalartos senticosus x E. trispinosus* (*Vorster 502*), had no effect on the cladogram topology. The inclusion of a second hybrid specimen *E. umbeluziensis x E. lehmannii* (*Vorster 438*), placed the specimen with the previous hybrid specimen, as a sister clade with the Lehmannii and Umbeluziensis clade (Figure 6.2a). As the number of hybrids is increased in the analysis, more topology changes take place, but the resolution of the cladograms does not change significantly except in the Umbeluziensis and Lehmannii clades (Figure 6.2b–e).

The effect of hybrids on cladistic analysis – The impact of hybrids on cladistic analysis (CI, tree length, and number of parsimonious cladograms) was also quantitatively tested by successively adding one to six hybrids to the study (Table 6.1). As the number of OTUs (operational taxonomic units) or specimens also effects cladistic analysis (Sanderson & Donoghue 1989), the cladograms obtained from adding hybrids were compared to cladograms with the same number of specimens.

According to Nelson (1983), Wanntorp (1983), and Funk (1985), the inclusion of hybrids in phylogeny will yield

- an increased homoplasy (expressed as poorer fit of the character set to the cladogram),
- an increased number of equally parsimonious cladograms that will collapse to poorly resolved consensus cladograms, and
- distortion of patterns of relationships among related non-hybrid species.

Adding terminals to a cladogram would be expected to produce an increase in length (Farris 1989). The consistency index is highly correlated to the number of taxa included in an analysis, with homoplasy increasing as the number of taxa increases. The consistency index will, therefore, decrease as the number of taxa increase (Sanderson & Donoghue 1989). I found that the inclusion of hybrids did increase the tree length from 966 to 1062, the number of equally parsimonious cladograms was influenced more by the presence of a

hybrid than the number of hybrids, and the CI lowered as the number of hybrids increased. However, the increased number of hybrids was cladistically no more problematic than the inclusion of more specimens to the data set.

Placement of hybrids – The placement of hybrids in cladograms, relative to their parental species, can be categorised into four groups according to McDade (1992). They are:

- hybrids that jump between parents in equally parsimonious cladograms,
- hybrids could be placed as basal members of the clade that include their most derived parent,
- hybrids proximate to the most apomorphic parent, and
- hybrids that group with other hybrids or taxa.

If hybrid specimens were individually included in the data set, and if more than one parsimonious cladogram were obtained, an Adams consensus was determined (Figure 6.3). According to Bremer (1990), Adams cladograms are unsatisfactory for systematic analysis, because they can contain components that are not present in any of the fundamental cladograms; but Funk (1985) and Hillis (1987) pointed out that they may be valuable in pin-pointing taxa that are responsible for incongruence.

McDade (1992) found that most hybrids are placed cladistically as basal members of the lineage that includes their most derived parent. In this study two hybrid specimens clustered with other specimens (Figure 6.3d–e), but most of the hybrids placed cladistically as basal members of the lineage.

Genetic distance analysis – The genetic distances between specimens used in this study range from 0.02 {*E. macrostrobilus* (*Vorster 1073*) versus *E.* sp. (Moyo) (*Vorster 1094)*} to 1.54 {*E. aemulans* (*Vorster 229*) versus *E. ghellinckii* (*Vorster 322*)}. The infraspecific genetic relationships of most

specimens are smaller than the interspecific relationships, as reflected by the consensus cladogram (Figure 6.1).

Genetic distance analysis of hybrids – The average genetic distance between hybrids and putative parents ranges from 0.15 {*E. trispinosus x E. ferox* (*Vorster 504*) versus *E. ferox* } to 0.54 {*E. senticosus x E. trispinosus* (*Vorster 502*) versus *E. senticosus*}. The hybrids are usually closely related to one of the putative parents than to other specimens, but not sufficiently significant that genetic distance analysis could be used to identify putative hybrids and parental relationships.

It is uncertain if cytoplasmic inheritance, whereby a hybrid inherits more characteristics from the maternal plant than the paternal plant, play a roll in the genus *Encephalartos*.

The genetic distance between a hybrid and pollen or seed parent specimen showed no expressive difference. Four of the six hybrids had a closer genetic distance to the male parent than to the female plant. Only the hybrid specimens *E. umbeluziensis* x *E. lehmannii* (*Vorster 438*) and *E. trispinosus* x *E. altensteinii* (*Vorster 506*) were more closely related to the female parent. More specimens of parents and putative hybrids are necessary to determine the effect of parental sex on the genetic distance between hybrids and its parents. In an analysis of *E. natalensis* (female) *x E. woodii* (male) hybrids, De Luca *et al.* (1995), found that the cpDNA genome consistently matches that of the female parent and, therefore, demonstrates the dominant influence of maternal inheritance.

When comparing hybrid specimens with the putative parents, unique fragments (6 – 22 fragments) were observed in the hybrid profiles. The unique fragments were not found in either of the parental specimens, and more of these fragments were observed in the RAPD data than DAF data. These anomalous fragments found in the hybrids could be explained by specimens representative of the parental species being used instead of the original parent individual (Smith *et al.* 1996). According to Smith *et al.* (1996) there are a few other explanations for the unique fragments found in hybrids:

- RAPD markers are not necessarily inherited in an allelic fashion
- there could have been an accumulation of mutations within the hybrid since the hybridisation event
- new primer binding site combinations could have formed as a result of recombination and crossing-over between the parental genomes.

6.4 Conclusion

Phylogenetic relationships – The specimens used in this study are closely related and are clustered in a monophyletic group, which means that the specimens that share derived characters are grouped together, and the specimens that share many derived characters are grouped closer together (Lipscomb 1998). The infrageneric groupings are well supported in this study. The close relationship between *E. macrostrobilus* and *E.* sp. *"Moyo"*, indicate that theses specimens could be similar. Further analyses including more specimens for these two species will confirm whether they should be separated as distinct species.

The results of this study concurs with previous morphological (Norstog 1980, Koeleman *et al.* 1981, Vorster 1986, Osborne 1988) and biochemical studies (Van der Bank *et al.* 1998). Conflict that may occur between morphological and molecular studies of phylogeny could be caused by different assumptions about the evolutionary process or methods of analysis (Hillis 1987).

Hybrids and molecular systematics – Interspecific hybridisation may have a serious impact on the population level for it may disrupt population structure by being more successful than the parents (Arnold *et al.* 1990) and this is most notable in small and rare species populations (Smith *et al.* 1996). According to Funk (1985), the hybridisation of species is uninformative in determining the pattern of relationship, because these individuals are not regarded as evolutionary units, and thus not species. But according to Anderson (1995),

understanding the reproductive biology of plants is important for understanding mechanisms of evolution, and proper interpretation of characters utilised in systematics. In this study the impact of hybrids on cladistic analysis was investigated to determine if hybrids behave differently in cladistic analysis than species, such that they might be detectable.

Although the hybrids did not behave reliably in cladograms, they did not disrupt cladistic analysis in such a way that it could be utilised in hybrid identification or determination of parental relationship. The fact that hybrids are mostly placed as a basal member, suggests that they will not disrupt phylogenies as much as has been predicted, and this also indicates that hybrids will not be easily identified by their behaviour in cladograms. Hybrids complicate cladistic analysis, but they play an important roll in the study of species relationships. Hybrids between distantly related parents are usually sterile or less fertile than the parents. The degree of sterility can be used in hybrid analysis as an indication of how closely related the two parent plants were to each other (Vorster 1986). Genetic distance can not be used in identification of hybrids as no significant similarities or differences were observed between hybrids and parental specimens.

Future studies – This study also helped identify problems that should be addressed in further parsimony analyses of the genus *Encephalartos*:

- Sufficient sampling of individuals per species is necessary.
- A good understanding of species relationship is necessary before hybrid identification and parsimony analysis can be done.
- In hybrid analysis the original parental specimens should be included.
- More extensive analysis of the genus *Encephalartos* are needed before the relationship with other cycad families are analysed.
- The selection of appropriate methods for a phylogenetic analyses in any study is very important. The methods used in this study were

not very successful. In future studies techniques such as AFLPs and sequencing can be used in conjunction with external morphological characters, ecological preferences, geographical distribution, and chromosomal features to fully understand the species relationships of the genus *Encephalartos*.

In conclusion – Phylogenetic analysis is only an estimation procedure of an evolutionary history based on incomplete information using genomic data (Swofford & Olsen 1990). The cladograms of the RAPD and DAF results do correspond, which indicate that the techniques used in this study were suitable in some of the molecular systematic analysis and support part of the results obtained. Although these techniques might be better applied in a population genetic study, this study has provided the ground work for future studies towards building a better understanding of the phylogenetic relationship between *Encephalartos* species.

CHAPTER SEVEN

REFERENCES

ADAMS, E.N. 1972. Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology* 21: 390–397.

ANDERBERG, A. & TEHLER, A. 1990. Consensus trees, a necessity in taxonomic practice. *Cladistics* 6: 399–402.

ANDERSON, E. & HUBRICHT, L. 1938. Hybridization in *Tradescantia*: The evidence for introgressive hybridization. *American Journal of Botany* 25: 396–402.

ANDERSON, G.J. 1995. Systematics and reproductive biology. In: Experimental and Molecular Approaches to Plant Biosystematics, (eds.) P.C. Hoch & A.G. Stephenson. *Monographs in Systematic Botany from Missouri Botanical Garden* 53: 263–272.

ARNOLD, L.A. 1953. Origin and relationship of cycads. *Phytomorphology* 3: 51–65.

ARNOLD, M.L., BENNETT, B.D. & ZIMMER, E.A. 1990. Natural hybridization between *Iris fulva and I. hexagona*: pattern of ribosomal DNA variation. *Evolution* 44: 1512–1521.

ASTOLFI, P., KENNETH, K.K. & CAVALLI-SFORZA, L.L. 1981. A comparison of methods for reconstructing evolutionary trees. *Systematic Zoology* 30: 156–169.

BACKELJAU, T., DE BRUYN, L., DE WOLF, H., JORDAENS, K., VAN DONGEN, S., VERHAGEN, R. & WINNEPENNINCKX, B. 1995. Random amplified polymorphic DNA (RAPD) and parsimony methods. *Cladistics* 11: 119–130.

BALDWIN, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositeae. *Molecular Phylogenetics and Evolution* 1: 3–16.

BALDWIN, B.G. 1993. Molecular phylogenetics of *Calycadenia* (Compositeae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution re-examined. *American Journal of Botany* 80: 222–238.

BALDWIN, B.G., SANDERSON, M.J., PORTER, J.M., WOJCIECHOWSKI, M.F., CAMPBELL, C.S. & DONOGHUE, M.J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247–277.

BARRETT, M., DONOGHUE, M.J. & SOBER, E. 1991. Against consensus. *Systematic Zoology* 40: 486–493.

BASSAM, B.J. & BENTLEY, S. 1994. DNA fingerprinting using arbitrary primer technology (APT): A tool or a torment. *Australasian Biotechnology* 4: 232–236.

BASSAM, B.J., CAETANO-ANOLLÈS, G. & GRESSHOFF, P.M. 1992. DNA amplification fingerprinting of bacteria. *Applied Microbiology and Biotechnology* 38:70–76.

BAUM, T.J., GRESSHOFF, P.M., LEWIS, S.A. & DEAN, R.A. 1994. Characterization and phylogenetic analysis of four root-knot nematode species using DNA amplification fingerprinting and automated gel electrophoresis. *Molecular Plant Microbe Interactions* 7: 39–47.

BOROWSKY R.I., McCLEILAND M., CHENG, R. & WELSH J. 1995. Arbitrarily primed DNA fingerprinting for phylogenetic reconstruction in vertebrates: the *Xilophoprus* model. *Molecular Biology and Evolution* 12: 1022–1032.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

BREMER, K. 1990. Combinable component consensus. *Cladistics* 6: 369–372.

BREMER, K. 1994. Branch support and tree stability. *Cladistics* 10: 295–304.

BREMER K. & WANNTORP H. 1978. Phylogenetic systematics in botany. *Taxon* 27: 317–329.

BROTHERS, D.J. 1978. How pure must a cladistic study be? A response to Nelson on Michener. *Systematic Zoology* 27: 118–122.

BYRNE, R. & MORLEY, B. 1976. Hybridization studies in *Columnea* L. (*Gesneriaceae*). *Botanical Journal of the Linn. Society* 72: 199–210.

CAETANO-ANOLLÈS, G. 1994. MAAP: a versatile and universal tool for genome analysis. *Plant Molecular Biolology* 25:1011–1026.

CAETANO-ANOLLÈS, G. 1997. DNA Amplification Fingerprinting. In: *Fingerprinting methods based on arbitrary primed PCR*, (eds.) M.R. Micheli & R. Bova. pp. 66–80. Springer Laboratory Manual.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Bio/Technology* 9:553–557.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1992a. DNA fingerprinting: MAAPing out a RAPD redefinition. *Bio/Technology* 10: 937.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1992b. Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. *Molecular and General Genetics* 235: 157–165.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1992c. Buffer components tailor DNA amplification with arbitrary primers. *PCR Methods and Applications* 4: 59–61.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1993. Enhanced detection of polymorphic DNA by arbitary amplicon profiling of endonuclease-digested DNA: identification of markers tightly linked to the supernodulation locus in soybean. *Molecular and General Genetics* 241: 57–64.

CAETANO-ANOLLÈS, G., BASSAM, B.J. & GRESSHOFF, P.M. 1994. Multiple arbitrary amplicon profiling using short oligonucleotide primers. In: *Current Topics in Plant Molecular Biology: Plant Genome Analysis*, (ed.) P.M. Gresshoff. pp. 29–46. CRC press, Boca Raton.

CAETANO-ANOLLÈS, G., CALLAHAN, L.M. & GRESSHOFF, P.M. 1996. The origin of bermudagrass (*Cynodon*) off-types inferred by DNA amplification fingerprinting. *Crop Science of America.*

CAETANO-ANOLLES, G., CALLAHAN, L.M., WILLIAMS, P.E., WEAVER, K.R. & GRESSHOFF, P.M. 1995. DNA amplification fingerprinting analysis of bermudagrass (*Cynodon*): genetic relationships between species and interspecific crosses. *Theoretical and Applied Genetics* 91: 228–235.

CAETANO-ANOLLÈS, G. & GRESSHOFF, P.M. 1994a. DNA amplification fingerprinting using arbitrary mini-hairpin oligonucleotide primers. *Bio/Technology* 12: 619–623.

CAETANO-ANOLLÈS, G. & GRESSHOFF, P.M. 1994b. Staining nucleic acids with silver: an alternative to radioisotopic and flourescent labeling. *Promega Notes* 45: 13–22.

CALLAHAN, L.M., WEAVER, K., CAETANO-ANOLLÈS, G.M, BASSAM, B.J. & GRESSHOFF, P.M. 1993. DNA fingerprinting of turfgrasses. In: *International turfgrass Society Research Journal* 7, (eds.) R.N. Carrow, N.E. Christians & R.C. Shearman. pp. 761–767. Intertec Publishing Corporation, Overland Park.

CAPUTO, P., STEVENSON, D.W.& WURTZEL, E. 1991. A phylogenetic analysis of American Zamiaceae (Cycadales) using chloroplast DNA restriction fragment length polymorphisms. *Brittonia* 43: 135–145.

CARPENTER, J.M. 1988. Choosing among multiple equally parsimonious cladograms. *Cladistics* 4: 291–296.

CHALMERS, K.J., WAUGH, R., SPRENT, J.L., SIMONS, A.J. & POWELL, W. 1992. Detection of genetic variation between and within populations of *Gäricidia sepium* and *G. maculata* using RAPD markers. *Heredity* 69:465–472.

CHAMBERLAIN, C.J. 1919. The living Cycads. Hafner, New York.

CHAMBERLAIN, C.J. 1920. The living cycads and the phylogeny of seed plants. *American Journal of Botany* 7:146–153.

CHAMBERLAIN, C.J. 1935. Gymnosperms. In: *Structure and evolution.* University Chicago Press, Chicago.

CITES Secretariat. 1982. *Proceedings of the third meeting of the conference of the parties,* New Dehli. Vol. 2. Secretariat of the Convention. IUCN. Morges.

CLEGG, M.T. & DURBIN, M.L. 1990. Molecular approaches to the study of plant biosystematics. *Australian Systematic Botany* 3: 1–8.

COBB, B.D. & CLARKSON, J.M. 1994. A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Research* 22: 3801–3805.

CRANE, P.R. 1988. Major clads and relationships in the higher gymnosperms. In: *Origin and evolution of the higher Gymnosperms*, (ed.) C.B. Beck. pp. 218–272. Columbia University Press, New York.

CRAWFORD, D.J., BRAUNER, S., COSNER, M.B. & STUESSY, T.F. 1993. Use of RAPD markers to document the origin of the intergeneric hybrid x *Marcgyraceaena skottsbergii* (*Rosaceae*) on the Juan Fernandez Islands. *American Journal of Botany* 80: 89–92.

CRUZAN, M.B. & ARNOLD, M.L. 1993. Ecological and genetic associations in an *Iris* hybrid zone. *Evolution* 47: 1432–1445

DEHGAN, B. & YUEN, C.K.K.H. 1983. Seed morphology in relation to dispersal, evolution and propagation of *Cycas. L. Botanical Gazette* 144: 412–418.

DE LUCA, P., MORETTI, A., SINISCALCO GIGLIANO, G., CAPUTO, P., COZZOLINO, S., GAUDIO, L., STEVENSON, D.W., WURTZEL, E.T. & OSBORNE, R. 1995. Molecular systematics of cycads. In: *Proceedings of the Third International Conference on Cycad Biology*, (ed.) P. Vorster. pp. 131–137. The Cycad Society of South Africa, Stellenbosch.

DEMEKE, T., ADAMS, R.P. & CHIBBAR, R. 1992. Potential taxonomical use of random amplified polymorphic DNA (RAPD): A case study in *Brassica. Theoretical and Applied Genetics* 84: 990–994.

DICKSON, E.E. & WEEDEN, N.F. 1991. Isozymes in North American *Malus* (*Rosaceae*): hybridization and species differentiation. *Systematic Botany* 16: 363–375.

DONALDSON, J.D. 1995. Insect relationship in *Encephalartos*. In: *Proceedings of the Third International Conference on Cycad Biology*, (ed.) P. Vorster. The Cycad Society of South Africa, Stellenbosch.

DYER, R.A. 1965a. New species and notes on type specimens of Southern African *Encephalartos*. *Journal of South Africa Botany* 31: 111–121.

DYER, R.A. 1965b. The cycads of Southern Africa. *Bothalia* 8 pp . Pretoria.

DYER, R.A., VERDOORN, I.A. 1966. Zamiaceae. In: *Flora of South Africa*, (eds.) LE. Codd & B. de Winter. pp. 3–34.

FARRIS, J.S. 1969. A successive approximations approach to character weighting. *Systematic Zoology* 18: 374–385.

FARRIS, J.S. 1982. Outgroups and parsimony. *Systematic Zoology* 31: 328–334.

FARRIS, J.S. 1988. *Hennig86, version 1.5. Program and documentation*. Port Jefferson Station, New York.

FARRIS, J.S. 1989. The retention index and homoplasy excess. *Systematic Zoology* 38: 406–407.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

FELSENSTEIN, J. 1991. *PHYLIP (Phylogeny Inference Package), version 3.4. Program and documentation*. Seattle, Washington.

FUNK V.A. 1985. Phylogenetic patterns of hybridization. *Annals of the Missouri Botanical Garden* 72: 681–715.

GIDDY, C. 1989. *Cycads of South Africa*, (ed.) B. Rycroft & H. Pearson. Purnell, Cape Town.

GILBERT, S. 1984. *Cycads: Status, trade, explotation and protection (1977–1982)*. World Wildlife Fund. Washington, D.C.

GILLES, A.C.M. & ABBOTT, R.J. 1998. Evaluation of random amplified polymorphic DNA for species identification and phylogenetic analysis in *Stylosanthes* (*Fabaceae*). *Plant Sytematics and Evolution* 211: 201–216.

GOODE, D. 1989. *Cycads of Africa*, (ed.) J. Comrie-Greig. The Struik Group, Cape Town.

GOODWIN, P.H., ANNIS, S.L. 1991. Rapid identification of genetic variation and pathotype of *Leptosphaeria maculans* by random polymorphic DNA assay. *Applied and Environmental Microbiology* 57: 2482.

GRANT, V. 1981. *Plant Speciation*.pp. 242–338. Columbia University Press, New York.

GRANT, V. 1998. Primary classification and phylogeny of the Polemoniaceae, with comments on molecular cladistics. *American Journal of Botany* 85: 741–752.

GRESSHOFF, P.M. 1993. Plant genome analysis by single arbitrary primer amplification. *Probe* 4: 32–36.

GRESSHOFF, P.M. 1996. The interface between RFLP techiques, DNA amplification and plant breeding. In: *Biotechnology in Agriculture #13. New diagnostics in Crop Science*, (eds.) J. Skerritt & R. Appels. pp. 101–125. Publishing University Press, Cambridge.

GRESSHOFF, P.M. & MACKENZIE, A.K. 1994. Low experimental variability of DNA profiles generated by arbitrary primer based amplification (DAF) of Soybean. *Chinese Journal of Botany* 6: 1–6.

HARRIS, S.A. 1995. Systematics and randomly amplified polymorphic DNA in the genus *Leucaea* Benth. (*Mimosoideae, Leguminosae*). *Plant Systematics and Evolution* 197: 195–208.

HEISER, C.B. 1949. Study in the evolution of the sunflower species *Helianthus annuus* and *H. bolanderi. Univ. Calif. Publ. Bot* 23: 157–196.

HENNIG, W. 1966. *Phylogenetic systematics*. University of Illinois Press, Urbana.

HEUN, M. & HELENTJARIS, T. 1993. Inheritance of RAPD's in F_1 hybrids of corn. *Theoretical and Applied Genetics* 85: 961–968.

HEUN, M., MURPHY, J.P. & PHILIPS, T.D. 1994. A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. *Theoretical and Applied Genetics* 87: 689–696.

HILLIS, D.M. 1987. Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics* 18: 23–42.

HOEY, B.K., CROWE, K.R., JONES, V.M. & POLANS, N.O. 1996. A phylogenetic analysis of *Pisum* based on morphological characters and allozyme and RAPD markers. *Theortical and Applied Genetics* 92: 92–100.

HOLDER, F. 1999. *A phylogenetic study of some representatives of the genus Pentaschistis.* Unpublished M.Sc. thesis, University of the Orange Free State.

HUMPHRIES, C.J. 1983. Primary data in hybrid analysis. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society*, (eds.) N.I. Platnick & V.A. Funk. pp. 89–103. Columbia University Press, New York.

HUMPHRIES, C.J. & FUNK, V.A. 1984. Cladistic methodology. In: *Current Topics in Plant Taxonomy*, (eds.) H. Heywood & D.M. Moore. pp. 323–361. Academic Press Inc., London.

INNIS, M.A. & GELFAND, D.H. 1990. Optimising of PCRs. In: *PCR protocols: A guide to methods and applications*, (eds.) M.A. Innis, D.H. Gelfand, J.J. Scinsky & T.J. White. Academic Press Inc., London.

INNIS, M.A., MYAMBO, K.B., GELFAND, D.H. & BROW, M.A.D. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of the National Academy of Science, USA* 85: 9436–9440.

JAYARAO, B.M., BASSAM, B.J., CAETANO-ANOLLÈS, G., GRESSHOFF, P.M. & OLIVER, S.P. 1992. Subtyping of *Streptococus uberis* by DNA amplification fingerprinting. *Journal of Clinical Microbiology* 30: 1347–1350.

JOHNSON, L.A.S. 1959. The families of cycads and the Zamiaceae of Australia. *Proceedings of the Linnean Society of New South Wales* 84: 64–117.

JOHNSON, L.A.S. & WILSON, K.L. 1990. General traits of the Cycadales. In: *The families and genera of vascular plants: Pteridophytes and*

gymnosperms, Volume I, (eds.) K. Kramer & P.S. Green. pp 363–368. Springer, New York.

KAEMMER, D., AFZA, R., WEISING, K., KAHL, G. & NOVAK, F.J. 1992. Oligonucleotide and amplification fingerprinting of wild species and cultivars of banana (*Musa* spp.). *Bio/Technology* 10: 1030–1035.

KARP, A., SEBERG, O. & BUIATTI, M. 1996. Molecular techniques in the assessment of botanical diversity. *Annals of Botany* 78: 143–149.

KLEYNHANS, R. 1997. *Genetic variation in Lachenalia bulbifera.* Unpublished M.Sc. thesis, University of the Orange Free State.

KLOPPER, K.C. 1996. *A preliminary phylogenetic study of the genus Pentaschistis (Poaceae, Arundinoideae).* Unpublished M.Sc. thesis, University of the Orange Free State.

KOELEMAN, A., ROBBERTSE, P.J. & EICKER, A. 1981. Die anatomie van die pinnas van die Suid-Afrikaanse spesies van *Encephalartos* Lehm.. *Journal of South African Botany* 47: 247–271.

KOELEMAN, A. & SMALL, J.G.C. 1982. A note on callus formation by stem and root tissue of some *Encephalartos* species. *Journal of South African Botany* 1: 165–166.

KUHNER, M.K. & FELSENSTEIN, J. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molecular Biology and Evolution* 11: 459–468.

LANDRY, P. & LAPOINTE, F. 1996. RAPD problems in phylogenetics. *Zoologica Scripta* 25: 283–290.

LANYON, S.M. 1985. Detecting internal inconsistancies in distance data. *Systematic Zoology* 34: 397–403.

LIN, J-J. & KUO, J. 1995. AFLP™. A novel PCR-based assay for plant and bacterial DNA fingerprinting. *Focus* 17: 66–70.

LIPSCOMB D. 1998. *Basics of cladistic analysis.* George Washington University, Washington D.C.

LISTON, A., ROBINSON, W.A. & OLIPHANT, J.M. 1996. Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering plants. *Systematic Botany* 21:109–120.

LIU, Z. & FURNIER, G.R. 1993. Comparison of allozyme, RFLP and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theoretical and Applied Genetics* 87: 97–105.

LOCONTE, H. & STEVENSON, D.W. 1990. Cladistics of the Spermatophyta. *Brittonia* 42: 197 – 211.

MADDISON, D.R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Systematic Zoology* 40: 315–328.

MADDISON, W.P., DONOGHUE, M.J., MADDISON, D.R. 1984. Outgroup analysis and parsimony. *Systematic Zoology* 33: 83–103.

MADDISON, W.P. & MADDISON, D.R. 1992. *MacClade, version 3: Analysis of phylogeny and character evolution.* Program and documentation. Sinauer Associates Inc., Sunderland.

MARCHANT, C.J. 1968. Chromosome patterns and nuclear phenomena in the cycad families Stangeriaceae and Zamiaceae. *Chromosoma* 24: 100–134.

MARGUSH, T. & MCMORRIS, F.R. 1981. Consensus *n*-trees. *Bulletin of Mathematical Biology* 43: 239–244.

MCDADE, L.A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.

MCDADE, L.A. 1997. Hybrids and phylogenetic systematics III. Comparison with distance methods. *Systematic Botany* 22: 669–683.

MEHRA, P.N. 1986. Letter to the Editor. *Encephalartos* 8: 34–35.

MICHELI, M.R., BOVA, R., CALISSANO, P. & D'AMBROSIO, E. 1993. Randomly amplified polymorphic DNA fingerprinting using combinations of oligonucleotide primers. *Bio/Techniques* 15: 388–390.

MIYAMOTO, M.M. 1985. Consensus cladograms and classifications. *Cladistics* 1: 186–189.

MOGFORD, D.J. 1979. Heterochromatin in *Encephalartos*. *Cytologia* 44: 951-954.

MORETTI, A., CAPUTO, P., COZZOLINO, S., DE LUCA, P., GAUDIO, L., GIGLIANO, G.S. & STEVENSON, D.W. 1993. A phylogenetic analysis of *Dioon* (Zamiaceae). *American Journal of Botany* 80: 204–214.

MULLIS, K.B. 1991. The polymerase chain reaction in an anemic mode: how to avoid cold oligodioxyribonuclear fusion. *PCR methods and applications* 1: 1–4.

NEI, M. 1987. *Molecular evolutionary genetics*. p. 288. Columbia University Press, New York.

NEI, M. & LI, N.H. 1979 Mathematical model for the study of genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science, USA* 74: 5267–5273.

NELSON, G. 1983. Reticulation in cladograms. In: Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society, (eds.) N.I. Platnick & V.A. Funk. pp. 105–111. Columbia University Press, New York.

NIXON, K.C. & CARPENTER, J.M. 1993. On outgroups. *Cladistics* 9: 413–426.

NIXON, K.C., CREPET, W.L., STEVENSON, D.W. & FRIIS, E.M. 1994. A re-evaluation of seed plant phylogeny. *Annuls of the Missouri Botanical Garden* 81: 484–533.

NORSTOG, K. 1980. Chromosome numbers in *Zamia* (Cycadales). *Caryologia* 33: 419–428.

NORSTOG, K. 1990. Studies of cycad reproduction at Fairchild Tropical Gardens. *Memoirs of the New York Botanical Garden* 57: 63–81.

OSBORNE, R. 1985. Sex change in cycads. *Encephalartos* 2: 20–22.

OSBORNE, R. 1986. *Encephalartos woodii*. *Encephalartos* 5: 4–10.

OSBORNE, R. 1988. Cycad seed kernel morphology. *Encephalartos* 13: 26–30

OSBORNE, R., GROBBELAAR, N. & VINCENT, P.L.D. 1990. A numerical phenetic study of the genus *Encephalartos* Lehm. In: *Proceedings of CYCAD '90, the Second Internasional Conference on Cycad Biology*, (eds.) D.W. Stevenson & K.J. Norstog. pp. 279–289. Palm & Cycad Societies of Australia, Queensland.

OSBORNE, R., GROBBELAAR, N. & VORSTER, P. 1988. South African Cycad research: progress and prospects. *Suid Afrikaanse Tydskrif vir Wetenskap* 84: 891–896.

OXELMAN, B. 1996. RAPD patterns, nrDNA ITS sequences and morphological patterns in *Silene* section *Sedoideae* (*Caryophyllaceae*). *Plant Systematics and Evolution* 201: 93–116.

PENNER, G.A., BUSH, A., WISE, R., KIM, W., DOMIER, L., KASHA, K., LAROCHE, A., SCOLES, G., MOLNAR, S. & FEDAK, G. 1993. Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR methods applications* 2: 341.

PIPKIN, S.B. 1972. Introgression between closely related species of *Drosophila* in Panama. *Evolution* 22: 140–156.

RAGOT, M., HOISINGTON, D.A. 1993. Molecular markers for plant breeding comparisons of RFLP and RAPD genotyping costs. *Theortical and Applied Genetics* 86: 975–984.

RAJORA, O.P. 1990. Marker allozyme genes and alleles for differentiation of *Populus deltoides, P. nigra, P. maximowiczii*, and interspecific hybrids. *Canadian Journal of Botany* 68: 990–998.

RAMSER, J., WEISING, K., LOPEZ-PERALTA, C., TERHALLE, W., TERAUCHI, R. & KAHL, G. 1997. Molecular marker based taxonomy and phylogeny of Guinea yam (*Dioscorea rotundata - D. cayenensis*). 1997. *Genome* 40: 903–915.

ROELEFS, R.M. 1979. The reproductive biology of *Cyrtandra grandiflora* (*Gesneriaceae*) on Oahu. *Pacific Sciences* 33: 223–231.

ROGSTAD. S.H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analysis. *Taxon* 41: 701–708.

ROODT, R. 1999. *A phylogenetic study of some South African representatves of the tribe Arundineae.* Unpublished M.Sc.thesis, University of the Orange Free State.

SAMBROOK, J., FRITSCH, E.F. & MANIATIS, T. 1989. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory Press, New York.

SANDERSON, M.J. & DONOGHUE, M.J. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781–1795.

SANG, T.J., CRAWFORD, D.J., KIM, S.C. & STUESSY, T.F. 1994. Radiation of the endemnic *Dendroseris* (*Asteraceae*) on the Juan Fernandez Islands: evidence from sequences of the ITS regions of nuclear ribosomal DNA. *American Journal of Botany* 81: 1494–1501.

SÅSTAD, S.M., STENØIEN, H.K. & FLATBERG, K.J. 1999. Species delimitation and relationships of *Sphagnum recurvum* complex (*Bryophyta*) - as revealed by isozymes and RAPD markers. *Systematic Biology* 24: 95–107.

SCHIERWATER, B. 1995. Arbitrarily amplified DNA systematics and phylogenetics. *Electrophoresis* 15: 1643–1647.

SHARP, P.A., SUGDEN, B. & SAMBROOK, J. 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* 12: 3055.

SHARROCKS, A.D. PCR technologies. The design of primers for PCR. PCR technologies chapter 2 page 5.

SIDDALL, M.E. 1994. *Random Cladistics, version 2.1.1.* Program and documentation. Toronto, Canada.

SIDDALL, M.E. 1995. Another Monophyly Index: revisiting the jackknife. *Cladistics* 11: 33–56.

SMITH, C.A.B. 1977. A note on genetic distance. *Annual Human Genetics* 21: 254–276.

SMITH, J.F., BURKE, C.C. & WAGNER, W.L. 1996. Interspecific hybridization in natural populations of *Cyrtandra* (*Gesneriaceae*) on the Hawaiian Islands: evidence from RAPD makers. *Plant Systematics and Evolution* 200: 61–77.

SOKAL, R. 1994. The continuing search for order. In: *Conceptual issues in evolutionary biology,* (ed.) E. Sober. p. 235. MIT Press, London.

SOKAL, R.R. & ROHLF, F.J. 1981. Taxonomic congruence in Leptopodomorpha re-examined. *Systematic Zoology* 30: 309–325.

SPORNE, K.R. 1967. *The morphology of gymnosperms*, (ed.) A.J. Cain. Hutchinson & Co. Publishers Ltd., London.

STEVENSON, D.W. 1985. A Proposed classification of the Cycadales. *American Journal of Botany.* 72: 971–972.

STEVENSON, D.W. 1990. Morphology and systematics of the Cycadales. *Memoirs of the New York Botanical Garden* 57: 8–55.

STEVENSON, D.W., OSBORNE, R. & HILL, K.D. 1995. The world list of cycads. In: *Proceedings of the Third International Conference on Cycad Biology*, (ed.) P. Vorster. pp. 55–64. The Cycad Society of South Africa, Stellenbosch.

SUH, Y., THIEN, L.B., REEVE, H.E., ZIMMER, E.A. 1993. Molecular evolution and phylogenetic implication of internal transcribed spacer sequences of ribosomal DNA in *Winteraceae. American Journal of Botany* 80: 1042–1055.

SWOFFORD, D.L. 1991. When are phylogeny estimates from morphological and molecular data incongruent? In: *Phylogenetic analysis of DNA sequences*, (eds.) M.M. Miyamoto & J. Cracraft. pp. 295–333. Oxford University Press, New York.

SWOFFORD, D.L. 1993. PAUP: *Phylogenetic analysis using parsimony, version 3.1.1.* Program and documentation distributed by Illinois Natural History Survey. Champaign, Illinois.

SWOFFORD, D.L. 1998. *PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4 beta.* Program and documentation. Sinauer Associates Inc. Publishers, Sunderland.

SWOFFORD, D.L. & MADDISON, W.P. 1987. Reconstructing ancestral character states under Wagner parsimony. *Mathematical Biosciences* 87: 199–229.

SWOFFORD, D.L. & OLSEN, G.J. 1990. Phylogenetic reconstruction. In: *Molecular systematics*, (eds.) D.M. Hillis & C. Moritz. pp. 411–501. Sinauer Associates Inc. Publishers, Sunderland.

TAGUCHI, G. 1986. *Introduction to Quality engineering.* Asian productivity organisation. UNIPUB, New York.

TAGUCHI, G. & WU, Y. 1980. *Introduction to off-line quality control.* Japan Quality Control Organisation. Nagoya, Japan.

THORMANN, C.E., FERREIRA, M.E., CAMARGO, L.E.A., TIVANG, J.G. & OSBORN, T.C. 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. *Theoretical and Applied Genetics* 88: 973–980.

TUSTIN, R.C. 1974. Toxicity and carcinogenicity of some South African cycad (*Encephalartos*) species. *South African Medical Journal* 48: 2369–2372.

VAN DER BANK, F.H., VORSTER, P. & VAN DER BANK, M. 1998. Phylogenetic relationships, based on allozyme data, between six cycad taxa indigenous to South Africa. *South African Journal of Botany* 64: 182–188.

VAN WYK, A.E. & CLAASSEN, M.I. 1981. Sex reversal in *Encephalartos umbeluziensis*. *Veld & Flora* pp. 120–122.

VISSER, N.C. 1997. Genetic variation in *Cenchrus ciliaris* (Poaceae, Panicoideae, Paniceae). Unpublished PhD.thesis, University of the Orange Free State.

VORSTER, P. 1986. Hybridization in *Encephalartos*. *Excelsa* 12: 101–106.

VORSTER, P. 1990. Taxonomy of *Encephalartos* (Zamiaceae): Taxonomically useful external characteristics. In: *Proceedings of CYCAD '90, the Second Internasional Conference on Cycad Biology*, (eds.) D.W. Stevenson & K.J. Norstog. pp. 294–299. Palm & Cycad Societies of Australia, Queensland.

WANNTORP, H.E. 1983. Reticulated cladograms and the identification of hybrid taxa. In: *Advances in Cladistics: Proceedings of the Second Meeting of the Willi Hennig Society*, (eds.) N.I. Platnick & V.A. Funk. pp.81–88. Columbia University Press, New York.

WATSON, J.D., GILMAN, M., WITKOWSKI, J. & ZOLLER, M. 1992. *Recombinant DNA*. W.H. Freeman and Company, New York.

WEAVER, K.R., CALLAHAN, L.M., CAETANO-ANOLLÈS, G. & GRESSHOFF, P.M. 1995. DNA amplification fingerprinting and hybridization of centipedegrass. *CropScience* 35: 881–885.

WEIR, B.S. 1996. *Genetic data analysis.* pp. 190–192. Sinauer Associates Inc. Publishers, Sunderland.

WELSH, J. & McCLELLAND, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18: 7213–7218.

WERTH, C.R. 1991. Isozyme studies on the *Dryopteris "spinulosa"* complex, I: the origin of the log fern *Dryopteris celsa*. *Systematic Botany* 16: 446–461.

WHITELOCK, L.M. 1986. Cycad pollen studies. *Encephalartos* 8: 22–23.

WHITELOCK, L.M. 1995. Cycad conservation in the past and the need for improvements in the future. In: *Proceedings of the Third International Conference on Cycad Biology*, (ed.) P. Vorster. pp. 1–8. The Cycad Society of South Africa, Stellenbosch.

WILKIE S.E., ISAAC P.G. & SLATER R.J. 1993. Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. *Theoretical and Applied Genetics* 86: 497–504.

WILLIAMS, J.G.K., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A. & TINGEY, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531–6535.

WILSON, K. & GAULDING, K.H. 1989. *A biologist's guide to principles and techniques of practical biochemistry.* Edwald Arnold, London.

WOLCIECHOWSKI, M.F., SANDERSON, M.J., BALDWIN, B.G. & DONOGHUE, M.J. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): Evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 80: 711–722.

WOODLAND, D.W. 1991. History and development of classification. In: *Contemporary Plant Systematics.* pp. 400–459. Prentice-Hall Inc., New Jersey.

YU, K.F., VAN DEYNZE, A. & PAULS, K.P. 1993. Random amplified polymorphic DNA (RAPD) analysis. In: *Methods in Plant Molecular Biology and Biotechnology.* pp. 287–301. CRC Press Inc., Florida.

ZHIFENG, G. & THOMAS, B.A. 1989. A review of fossil cycad megasporophylls, with new evidence of *Crossozamia* Pomel and its associated

leaves from the lower Permian of Taiyuan, China. *Reviews of Palaeobotany Palynology* 60: 205–223.

APPENDIX A: A data matrix in the pilot study, representing six specimens, amplified using eight RAPD primers and separating the amplification product on 1.5% agarose (0 – absent, 1 – present, ? – missing fragment).

1. *E. villosus* (*Vorster 265*)	000110101100100010101111011010010000111011100?010110010011
2. *E. altensteinii* (*FS 01*)	010011001100011101001001110010010000111 0?110000010110010010
3. *E. natalensis* (*PRE 18*)	0?00110010 0?0?0?0?00101011001101010101000 110000010110010011
4. *E. ferox* (*Vorster NR*)	000010101110010010 01?00?0?000001010011100 ??000?011111110010
5. *E. lembomboensis* (*Vorster 313*)	000011010110010111001000110110011100110 0111001001011001 0010
6. *Stangeria eriopus* (*Vorster 337*)	101001001001000010100001100010111110101 1101111110111011101

APPENDIX B: A data matrix in the pilot study, representing six specimens, amplified using five DAF primers and separating the amplification product on 8% polyacrylamide medium.

1. *E. villosus* (*Vorster 265*)	1100101110000010111000011110000100000111
2. *E. altensteinii* (*FS 01*)	1110101110110110110100117?1?00000000011011
3. *E. natalensis* (*PRE 18*)	1110101100111111010100010100010010007001
4. *E. ferox* (*Vorster NR*)	0101101100110110111100010110000011110000 1
5. *E. lembomboensis* (*Vorster 313*)	0101101110100111110101010?000000101001101
6. *Stangeria eriopus* (*Vorster 337*)	0001111111100110101111010101101011100 1000

APPENDIX C: A data matrix in the pilot study, representing six specimens. The template was digested with restriction endonuclease prior to amplifcation with the DAF primers (DAF2, DAF3, DAF5, DAF6, DAF10 and DAF12).

1. *E. villosus* (*Vorster 265*)	1100011?111 001000111110110 11100000111 100000?11011010 10000011101 0001100000000??10
2. *E. altensteinii* (*FS 01*)	11000111111 001000111110110 00000000110 100000111011010 10111011000 11100000011111010
3. *E. natalensis* (*PRE 18*)	11000111111 001110111110110 00001011110 101000111011110 10011001000 11100011011111011
4. *E. ferox* (*Vorster NR*)	11100111111 001000111110110 0011101???0 100000011011011 11111011000 11000001011117?10
5. *E. lembomboensis* (*Vorster 313*)	11000101111 001000111111011 00000010110 101101011111011 00111011100 00110001011111010
6. *Stangeria eriopus* (*Vorster 337*)	10111101111 111001111111011 10111111110 010111111011010 00111111110 00001101111111110

APPENDIX D: A data matrix in the pilot study, representing six specimens, amplified using three multiplex DAF primers (K1, K2 and K3).

1. ***E. villosus*** ***(Vorster 265)***	0000000000001111110 0111111001110111 0001100111110
2. ***E. altensteinii*** ***(FS 01)***	0000000000001111110 1111111001110111 0011100111110
3. ***E. natalensis*** ***(PRE 18)***	0000000101111111011 1011011001110111 001?1001?1110
4. ***E. ferox*** ***(Vorster NR)***	0010011110111111010 1011101001110111 0110100111110
5. ***E. lembomboensis*** ***(Vorster 313)***	1110111110110110100 0011111011110111 1010111111111
6. ***Stangeria eriopus*** ***(Vorster 337)***	0011111111110110100 0111111101101111 0010100011110

APPENDIX E: A RAPD data matrix representing 55 specimens, amplified with primers: OPA11, OPA16, OPA20, OPB3 and OPB6.

Encephalartos aemulans (*Vorster 229*)	1010010000011111010001000101011101000010 0101000100111100010100
E. altensteinii (*PRE 06*)	1010110000011111010001?1?10101011100000 1?100010?110100010000
E. altensteinii (*FS 01*)	1010110000?1111010001?1111?0?01011?00000 1110001011101000100000
E. aplanatus (*Vorster 422*)	1110101?000111101011110010000100101017?0? 1110001001101010100
E. aplanatus (*Vorster 312*)	1110101010001117?01011110010000100?010110 11?1000100110101010100
E. arenarius (*Vorster 998*)	0110101000001000010101010101011001110000 1101010101010101010100
E. arenarius (*PRE 08*)	0110101000001000010?0101010101100???00001 1010101010101010100
E. bubalinus (*Vorster 978*)	1011100000001001010101000100011101000000 0001000000010110000100
E. caffer (*Vorster 289*)	1010010000001001010001010101001000110010 0001000000010100001000
E. cf.*chimanimaniensis* (*Vorster 160*)	0001010000001000010001010100000011000000 1101000100000100000001
E. cupidus (*SBG 01*)	1110010000001001010001000101011101010101 ??????????010100000000
E. eugene-maraisii (*Vorster 268*)	0110010000011001010001000100011100010010 0001000100110100100001
E. eugene-maraisii (*PRE 21*)	0110010000?1001010001000100011100010010 0001000?00?10100100001
E. ferox (*Vorster s.n.*)	1110?1000000100?1100101001101000100000000 1011001101011010100001
E. ferox (*Vorster s.n.*)	1110110000001001?1001010010010001000000 ??????????011010100001

E. friderici-guilielmi (*SBG 02*)	00100100000110000101010001010110101101100000000001000000100000
E. ghellinckii (*Vorster 322*)	10000010000000000101010010000100000000001000110001000000000001
E. gratus (*Vorster 834*)	10101000100000010100010001010011001100100000000000000000000001
E. heenanii (*Vorster 622*)	10100100001010001100010101000111001100000000000000010100000000
E. hildebrandtii (*Vorster 507*)	11101100100010010101100001010110110000000101000000010100101100
E. horridus (*PRE 12*)	01100100000001?0101010101110010010000000001000000001010000000?0
E. horridus (*Vorster 362*)	01100100000001010101?0101110010010000000001000000001010000000010
E. humilis (*PRE 15*)	10100010000010000101000101010010100000000001100000011010000000000
E. inopinus (*PRE 05*)	11101100101001010101010101010101011100000100000000000100100000
E. laevifolius (*PRE 19*)	01100010001?10010101010011010100000000000?00110001010101000?0
E. laevifolius (*Vorster 260*)	01100010000?11001010101001?010?00000000000?????????0?0?0?000010
E. laevifolius (*Vorster 258*)	01100010000??10010101010011010?0000000000010011000101010100000?0
E. lanatus (*PRE 17*)	11100010000110000101001011010010100101001000110001000100000000
E. laurentianus (*Vorster 1000*)	10100100000000000101011001000000010000000????????????110100000001
E. lebomboensis (*PRE 14*)	00100010000011110101010001010101001010001000000001101000000001
E. lebomboensis (*PRE 01*)	10110100100011101101010101010001000000001001000000110100000101
E. lehmannii	01100100100011000101010101010010??0?0000

(PRE 11)	10101000001010010010 1
E. lehmannii *(Vorster s.n.)*	???????????011000101 0?0?0101001001101000010?0100000101001 00?01
E. macrostrobilus *(Vorster 1073)*	1010100010000000010001000111010100000000 0001000000000100000001
E. msinganus *(PRE 20)*	1010101000001001010101010110000001000010 0000000010010100000001
E. munchii *(Vorster s.n.)*	1010110010001001010001000101010010000000 0001000000000100000011
E. natalensis *(PRE 18)*	1010010010000000010101100101010100000000 ???????????110101000101
E. senticosus *(Vorster 273)*	0010001000000011010101010100011011001000 ???????????000000000011
E. trispinosus *(Vorster s.n.)*	0110010100000000010101010100011001001001 ???????????010100000001
E. turneri *(Vorster 1110)*	1010110010000001010101000101001010101001 0000000010110100000101
E. umbeluziensis *(PRE 23)*	1010010000001001010101011100010010101001 0000000010110010000101
E. umbeluziensis *(Vorster 267)*	???????????010010101010??1000100?0?0?001?? ????????110010000101
E. villosus *(PRE 02)*	?0100100000010010101000111010 1??0?0010?0 0010000100 1?010100011
E. villosus *(SBG 03)*	1010010000001001010100011101011 10?001001 0001000010011010100011
E. villosus *(Vorster 265)*	1010010000001 00?0101000 11?0?0?1?01001001? ?????????011010100 0?1
E. sp. (Didinga) *(Vorster 1101)*	1110110010000000010101000001011100000000 ???????????000101000001
E. sp. (Moyo) *(Vorster 1094)*	1010101010000000010001000001000110001000 0001000011010100000011
E. senticosus x E. trispinosus (*Vorster 502*)	1010010100001110110100100111000011000000 ???????????010100000001

E. umbeluziensis x E. lehmannii (*Vorster 438*)	10100100000011000101001001010010100000 ??????????010100000001
E. umbeluziensis x E. villosus (*Vorster 416a*)	10101100000010000100111001010100001010010 01010000000101000000011
E. altensteinii x E. trispinosus (*Vorster 701*)	??????????000000100010101010100001010012? ????????010000000001
E. trispinosus x E. altensteinii (*Vorster 506*)	11101100000011010100001101010010010000000 ??????????010100101001
E. trispinosus x E. ferox (*Vorster 504*)	01100101000010010100001001000100000000000 ??????????010100000000
E. transvenosus x E. woodii (*Vorster 787*)	10100101000011010100010000000000000000000 11010000001101000000001
Stangeria eriopus (*Vorster 337*)	10100000010101001010000010000000000010000 ??????????010001100000

APPENDIX F: The average genetic distance (D) calculated between the different specimen profiles amplified with the RAPD primers: OPA11, OPA16, OPA20, OPB3 and OPB6. (# – represents values that could not be calculated, due to division by zero).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1. *E. aemulans*																																		
2. *E. altensteinii*	0.17																																	
3. *E. aplanatus*	0.37	0.24																																
4. *E. arenarius*	0.50	0.37	0.14																															
5. *E. bubalinus*	0.41	0.38	0.24	0.22																														
6. *E. caffer*	0.45	0.36	0.43	0.27	0.41																													
7. *E.* cf. *chimanimaniensis*	0.72	0.62	0.64	0.34	0.63	0.59																												
8. *E. cupidus*	0.41	0.31	0.31	0.30	0.27	0.24	0.73																											
9. *E.* sp. (Didinga)	0.74	0.64	0.56	0.37	0.44	0.76	0.96	0.34																										
10. *E. eugene-maraisii*	0.32	0.30	0.36	0.35	0.41	0.38	0.54	0.25	0.44																									
11. *E. ferox*	0.92	0.94	0.58	0.69	0.89	0.86	0.64	0.89	1.04	0.62																								
12. *E. friderici-guilielmi*	0.69	0.60	0.72	0.35	0.72	0.29	0.85	0.41	0.56	0.38	0.69																							
13. *E. ghellinckii*	1.56	1.73	1.13	0.69	0.99	1.47	1.01	1.22	0.79	1.29	1.01	0.83																						
14. *E. gratus*	0.72	0.52	0.64	0.54	0.51	0.28	1.14	0.41	0.22	0.47	0.73	0.37	0.86																					
15. *E. heenanii*	0.58	0.41	0.59	0.32	0.37	0.18	0.66	0.21	0.41	0.35	0.52	0.34	0.97	0.41																				
16. *E. hildebrandtii*	0.45	0.43	0.30	0.22	0.26	0.46	0.64	0.33	0.35	0.54	0.27	0.46	1.25	0.64	0.59																			
17. *E. horridus*	0.56	0.55	0.38	0.21	0.44	0.31	0.49	0.34	0.37	0.50	0.49	0.52	1.20	0.77	0.44	0.29																		
18. *E. humilis*	0.62	0.62	0.35	0.26	0.41	0.28	0.58	0.51	0.58	0.58	0.45	0.49	0.64	0.73	0.41	0.26	0.30																	
19. *E. inopinus*	0.63	0.28	0.41	0.26	0.46	0.43	0.72	0.31	0.24	0.51	0.72	0.44	0.99	0.32	0.30	0.20	0.38	0.75																
20. *E. laevifolius*	0.81	0.72	0.46	0.15	0.53	0.61	1.06	0.44	0.48	0.59	0.73	0.51	0.36	0.88	0.66	0.46	0.54	0.41	0.37															
21. *E. lanatus*	0.96	0.87	0.43	0.43	0.81	0.66	0.95	0.49	0.66	0.75	0.54	0.29	0.45	0.79	0.86	0.52	0.72	0.29	0.64	0.26														
22. *E. laurentianus*	0.67	0.46	0.37	0.59	0.44	0.52	0.50	0.44	0.36	0.51	0.36	0.69	0.53	0.65	0.45	0.48	0.22	0.41	0.37	0.50	0.45													
23. *E. lebomboensis*	0.57	0.34	0.41	0.33	0.46	0.53	0.76	0.46	0.51	0.43	0.63	0.54	0.54	0.51	0.47	0.59	0.58	0.44	0.32	0.22	0.48	0.51												
24. *E. lehmannii*	0.52	0.58	0.36	0.15	0.49	0.38	0.44	0.49	0.34	0.38	0.44	0.47	0.92	0.66	0.50	0.15	0.16	0.37	0.27	0.34	0.41	0.54	0.44											
25. *E. macrostrobilus*	0.69	0.49	0.72	0.51	0.37	0.56	0.69	0.48	0.09	0.66	0.54	0.73	0.59	0.18	0.49	0.41	0.49	0.58	0.20	0.54	0.65	0.54	0.41	0.44										
26. *E.* sp. (Moyo)	0.89	0.69	0.84	0.53	0.61	0.69	0.85	0.72	0.34	0.78	0.69	0.88	0.61	0.44	0.76	0.53	0.76	0.48	0.54	0.69	0.52	0.69	0.44	0.58	0.24									
27. *E. munchii*	0.74	0.38	0.56	0.47	0.34	0.31	0.37	0.34	0.17	0.41	0.37	0.41	0.69	0.26	0.44	0.13	0.56	0.41	0.19	0.37	0.45	0.49	0.37	0.31	0.08	0.04								
28. *E. natalensis*	0.53	0.43	0.43	0.44	0.41	0.58	0.85	0.41	0.07	0.46	0.56	0.59	0.61	0.44	0.41	0.44	0.51	0.48	0.26	0.34	0.52	0.15	0.26	0.29	0.24	0.30	0.24							
29. *E. nsinganus*	0.69	0.60	0.35	0.35	0.32	0.37	0.56	0.50	0.56	0.56	0.56	0.73	0.61	0.56	0.63	0.44	0.41	0.37	0.34	0.34	0.52	0.34	0.34	0.47	0.44	0.30	0.34	0.54						
30. *E. senticosus*	0.82	0.84	0.53	0.44	0.51	0.73	0.58	0.63	0.58	0.69	0.73	0.62	0.46	0.58	0.66	0.54	0.52	0.49	0.44	0.45	0.54	0.58	0.24	0.59	0.73	0.41	0.33	0.69	0.26					
31. *E. trispinosus*	0.87	0.77	0.59	0.32	0.47	0.54	0.52	0.29	0.30	0.44	0.81	0.56	0.74	0.81	0.37	0.41	0.28	0.56	0.41	0.41	0.62	0.30	0.41	0.26	0.66	0.62	0.52	0.41	0.34	0.17				
32. *E. turneri*	0.61	0.61	0.46	0.46	0.43	0.41	0.98	0.43	0.29	0.57	0.83	0.50	0.96	0.29	0.53	0.31	0.53	0.41	0.30	0.58	0.54	0.37	0.30	0.33	0.47	0.18	0.21	0.22	0.32	0.44	0.26			
33. *E. umbeluziensis*	0.67	0.77	0.60	0.52	0.49	0.46	0.66	0.49	0.66	0.55	0.34	0.47	0.58	0.66	0.50	0.43	0.72	0.47	0.44	0.54	0.76	0.44	0.27	0.46	0.79	0.51	0.34	0.35	0.29	0.31	0.22	0.00		
34. *E. villosus*	0.58	0.83	0.87	0.60	0.41	0.46	0.89	0.41	0.53	0.46	0.35	0.56	0.85	0.64	0.49	0.43	0.59	0.46	0.43	0.53	0.61	0.75	0.43	0.38	0.53	0.41	0.35	0.52	0.46	0.41	0.32	0.28	0.06	
35. *E. senticosus x E. trispinosus*	0.63	0.46	0.46	0.66	0.64	0.61	0.59	0.53	0.59	0.81	0.19	0.63	1.06	0.88	0.54	0.38	0.34	0.41	0.47	0.59	0.34	0.28	0.37	0.41	0.59	0.56	0.37	0.37	0.31	0.56	0.34	0.37	0.44	0.50
36. *E. umbeluziensis x E. lehmannii*	0.65	0.38	0.56	0.47	0.54	0.41	0.62	0.44	0.37	0.61	0.26	0.30	0.69	0.77	0.34	0.29	0.44	0.30	0.28	0.37	0.22	0.26	0.19	0.31	0.49	0.45	0.17	0.19	0.41	0.58	0.33	0.19	0.15	0.31
37. *E. umbeluziensis x E villosus*	0.59	0.30	0.43	0.35	0.49	0.38	0.54	0.41	0.54	0.55	0.54	0.58	0.92	0.54	0.41	0.28	0.72	0.58	0.35	0.54	0.63	0.44	0.35	0.38	0.34	0.22	0.10	0.35	0.47	0.63	0.31	0.19	0.24	0.22
38. *E. altensteinii x E. trispinosus*	1.13	0.86	1.32	0.52	0.96	0.58	0.74	0.62	0.74	0.85	0.97	0.61	0.63	0.56	0.50	0.99	0.65	0.79	0.54	0.56	1.20	0.74	0.29	0.58	0.56	0.51	0.56	0.54	0.61	0.36	0.11	0.29	0.14	0.33
39. *E. trispinosus x E. altensteinii*	0.57	0.41	0.33	0.49	0.56	0.35	0.63	0.30	0.41	0.43	0.22	0.66	1.39	0.51	0.47	0.12	0.37	0.44	0.16	0.51	0.37	0.41	0.50	0.19	0.63	0.59	0.31	0.50	0.26	0.59	0.37	0.32	0.47	0.35
40. *E. trispinosus x E. ferox*	0.75	0.53	0.54	0.56	0.52	0.37	0.79	0.30	0.79	0.37	0.20	0.65	1.39	1.01	0.41	0.34	0.54	0.65	0.45	0.20	0.41	0.45	0.45	0.49	0.79	0.97	0.32	0.45	0.50	0.74	0.27	0.58	0.41	0.49
41. *E. transvenosus x E. woodii*	0.48	0.38	0.38	0.58	0.54	0.41	0.37	0.54	0.62	0.41	0.26	0.99	0.69	0.77	0.56	0.58	0.56	0.66	0.48	0.62	0.73	0.37	0.28	0.31	0.62	0.45	0.26	0.48	0.41	0.92	0.45	0.37	0.34	0.51
42. *Stangeria eriopus*	1.34	1.04	0.79	1.25	1.50	1.18	#	1.22	1.30	1.06	0.79	1.34	1.39	1.30	0.88	1.54	1.39	1.34	0.92	1.01	0.74	1.30	1.43	1.18	1.70	1.25	1.70	1.14	1.34	2.35	1.66	1.43	1.10	0.37

APPENDIX G: DAF data matrix representing 55 specimens amplified with eight primers: DAF2, DAF3, DAF5, DAF6, DAF10, K1, K2 and K3.

Encephalartos aemulans (*Vorster 229*)	011111110101010100000100010000000000000000100001101101010011100000100111110000000100000000000010101010000001001101010000000010101110000110000000011010001010000000000000
E. altensteinii (*PRE 06*)	11101?0???0?0???0?????00?10000000000000000010111110100101011110?0100???0000010100000000??0110110100110000000111010001000000000000000
E. altensteinii (*FS 01*)	1110?10101010111011111001100000000000000000101111101001010111101010001010101001010010111001111101012000111010000010100000000110110100110000000?101000?0000000000000000
E. aplanatus (*Vorster 422*)	1110010100000000000000000110100100101001110101111112?0111011100101000001010000000111000010000111?0?011000??00001010100000001101001101010100011001000100000000000000000000
E. aplanatus (*Vorster 312*)	1110010100000000000000000110?00100101001110101111111011101110010100000101000000011100001010?1100001100000000010101000000011?00?10101010001100100010000000000000000000
E. arenarius (*Vorster 998*)	1111111101010001000111111100000102000000001000111011001100111010001000101000011101100001100011110111100011000000101100000011101001010011000000000101001110100000000110100

E. bubalinus (*Vorster 978*)	??????????????????????0001110000110001000010 1000001011010000011001000001101010010011 1000000000110100000000000000110101000000000 0010001100000000010000000010010000011000 000000
E. caffer (*Vorster 289*)	1000000000000000000000000100000100001000 0001010011000000000110000001000001010010 0111000000000?????????????000000101000000 1000000010100000001000000001011000001100 0000000
E. cupidus (*SBG 01*)	1111100100010101010000?????????????????????? 1010000011001000010000001000100000010001 0000100001111100000000000010100010000000 0001100100000000000000?????????????????????? ???
E. eugene-maraisii (*Vorster 268*)	1111110100011?10101000101000000000000000 0001010100111011000010010001O?0111100010 0110000000001111110010000000100010110000 1000010100111111000000001010000111000001 000000000
E. eugene-maraisii (*PRE 21*)	1111110100011?0101000101000000000000000 0001010100111011000010010001010112?00010 0110000000001111110O?000000010001011000 ?0000101001111110000000010?0000???00000?0 00000000
E. ferox (*Vorster s.n.*)	0011110101101101010111011100000000000000 0001001002?0?0001010110010101111111101110 1111101000111111?1010001100000001011110 1100111000100001000010000000100001010100 000?0?000
E. ferox (*Vorster s.n.*)	0011110101101101010??101110000000000000 0001001001010001010110010101111111?01110 1?1??0?000111111111010001100000010111101

	10011110010000111001000000010000101010000101000
E. friderici-guilielmi (*SBG 02*)	01110101001001010100000001000000000000000001000101010001000011110101010101001000100000101011111001000000010100010100001000000110111110001001000100010000000000000000000
E. ghellinckii (*Vorster 322*)	??????????????????????00001010000000011100110000100000000000010110001??????????????????????????1110000000000100000000000000001001000010000000010011000000000000000000000
E. gratus (*Vorster 834*)	11010101011001010100011001000000000000000000000010001100100011001110101010000000001000100000011110101000110100010101010000001101011000011000000010100100010100101110000
E. heenanii (*Vorster 622*)	??????????????????????10001000000000000000000001000110010001100111010??????????????????????????11000000000000000000000000000011?????????????????????10000100000000000000000
E. hildebrandtii (*Vorster 507*)	??????????????????????00001101010000010111111011111011001001010011101010110110001000100000000111011000000000000000000011000000000000000100010000???????????????????????
E. horridus (*PRE 12*)	??????????????????????00001101110000000001110010011100101001100110027?0000??0000000011000000001??????????????0010?0110000?00101000001000010000101???????????????????????
E. horridus (*Vorster 362*)	00111101010110000000000001101110000000001110010011?10010100110011001000011000000

	0011000000007111110011000000010101100001 0010100000100001000010100001001000000001 100000100
E. humilis (*PRE 15*)	??????????????????????1000100000000101000000 0000001011000000010000000000100010000001 1000000000111000000000000000001010000000 0000000100000000000007???????????????????? ???
E. inopinus (*PRE 05*)	??????????????????????0000101110100000000010 0000001110011011111011001000111100000000 0000000000011110011000000010101010000000 0001000100111000001001000000000000000000 000000
E. laevifolius (*PRE 19*)	??????????????????????0100100000000000000000 0000??1111010107017?0?100101011110011011101 1111111111?1107?1110111000000101111110101 0101010101110?001?0000?0111100000010010001000
E. laevifolius (*Vorster 260*)	??????????????????????0100100000000000000000 0000111?1101010101101100?0???????????????? ??010?0 10?01110100170?????????????????????????
E. laevifolius (*Vorster 258*)	0011111101000000000000100100000000000000 0000000?1111010101011011007010?11001101 1101117??117111110111101110000001011117107 1010101010101110100110000101?11000000100 ?00?000
E. lanatus (*PRE 17*)	??????????????????????0100100000000000000000 0010100010000100011000000000100110000001 0000000000011100000000001000001100000000 0100110001011101001007???????????????????? ???
E. lanatus	0011010101000000000000000100000000000000

(*Vorster s.n.*)	00000000111010000100100100000????????????? ?????????????????????????1000001100000000000 0000000101110100000100000000000000000000 0000
E. laurentianus (*Vorster 1000*)	?????????????????????????0001100000000001000000 1011101010011010010000001110000000000100 0000000001 1?????????????????????????????????10 0000000000000000011100000000000000000000 00
E. lebomboensis (*PRE 14*)	0011???101000000000010001000000000000000 0101???10000?00001?00000?0????????????????? ?????????????????????1000?011100100000101 0101000110000100010101100000000010000000 00
E. lebomboensis (*Vorster 313*)	0001110101000000000001000100000000000000 0001011111000001000011000001000100000001 0011000000011011101010101001000101110010 0000?0101010001100001000107?0??00000000?0 00000000
E. lebomboensis (*PRE 01*)	?????????????????????0000110000000001000011 1010111000000111010010001 0????????????????? ?????????????????????1000101110000000001 00 1010001000001001001011100000000100000100 0
E. lehmannii (*Vorster s.n.*)	0011010101110001110111001100000000000000 0001110010010010000110001001011110000000 0111000000011 0?????????????0000101010000000 0001100010100100011001011010101100001100 1010000
E. macrostrobilus (*Vorster 1073*)	?????????????????????0001110110010001001011 1000111011011011110001011111101111010011 1000000011010110010100001000101010000000 0000000101010000110010000100000000110100

	010000
E. msinganus (*PRE 20*)	10100101110000000100000001011101001010000111110101000010010011010001??????????????????????????01101010000000000001100000000000000101010001010010000000000011001000000000
E. munchii (*Vorster s.n.*)	10101001010000000110000001011101000000000111001011011001010110010001??????????????????????????011000000000000000011000001001100000100011000111000000000001100000000000000
E. natalensis (*PRE 18*)	10111000000100100110010010000000000000000001010011001001100010010001000011111000111100000000110111010100100000000001100000000000010101010000100100??????????????????????????
E. senticosus (*Vorster 273*)	00010101011111100100010001000000000000000010000011000001000011010001000111111010011100000000111011010110000000001011000000000010001110100100001001100000000010000000000000000
E. trispinosus (*Vorster s.n.*)	00000101000000000100010001000000000000000000000011000100100000000000010001111011001111110000011111110010000000000011010000000010111110001010000000100010100100000000000000000
E. turneri (*Vorster 1110*)	????????????????????????000010111010011000101100110100110101001100010010100100000000111100000011??????????????1000001100000000000000110110100001000??????????????????????????
E. umbeluziensis (*PRE 23*)	????????????????????????000010111010010000001170111011?011101111011000??????????????????

	?????????1?1001?000000100010110000000011000011????00001?00?000000011000001001101011
E. umbeluziensis (*Vorster 267*)	??????????????????????0000101110100100000011101101?01??1011?011110110000100110000000111001100101111100110000001000101100000000110000?10111100001100100000001100000100111010111
E. villosus (*PRE 02*)	?10?101?0?0???000?00000001101001001010011111011111011011?011101010001000110111011111011001011107001100011100?01010110000001100000100?100?011111?000000000??0?0?01??????
E. villosus (*SBG 03*)	??01101?010?1?000100000001101001001010011111011?10110111?011?0101000?000??0???0?111101100?01110?00?10001?100?010101?000000110000010011001 0?1???100000000001101010111111
E. villosus (*Vorster 265*)	1?011011010111000100000001101 00?00101001111101117?10?10111101 1?0101000?0001?011101111?01?00?01110100?10001??00101010 1?000001100000100??0010?1????000000000??0?0?01101???
E. cf. *woodii* (*Vorster 682*)	1111000111000101000001000000000000000000010001100010100000101000100000010111010011001111100110110011000000000010101000000011000011011111010100010000000000000000000000000
E. sp. (Didinga) (*Vorster 1101*)	0000100000000000000000000011000010010100101110100000110010001110100011000100000001100000000001111101010000000010100011000100100000011010000000000100000001100000000000000000

E. sp. (Moyo) (*Vorster 1094*)	????????????????????????0001110000000001001011 100101101101101011000100111110101001011 100000001101011001000000000000010100000000 000000010101000011001000010000011010110101 010000
E. senticosus x E. trispinosus (*Vorster 502*)	101101000000000000000000010100010000000 101100001100000110001101000100000000010 101101010000111111000000000110001101000 1 010110000100000000001000000010101100000 1 000000000
E. umbeluziensis x E. lehmannii (*Vorster 438*)	101101000000000000000000010101000000000 100101000100000100001101000100000000010 011101010000111100000000000000111100000 000010000100000100001000000011001100001 1 001110000
E. umbeluziensis x E. villosus (*Vorster 416a*)	110011000000000000000000010101000010000 000101111100001100001000010100001100001 0 011110110000110111011000000000001010000 000110100010000100001000000001001000011 001110000
E. altensteinii x E. trispinosus (*Vorster 701*)	001111010010000001000?????????????????????? 000001101100100001001000100001111101010 1 000000000110111011000000100000101000000 0 010110011100110001000100011000000000000 000000
E. trispinosus x E. altensteinii (*Vorster 506*)	111101010110011001011110000000000000000000 000000001101100100001001000100001110101 0 101000011100110111011010110000000101101 0 010010110011100110001000?????????????????? ???????
E. trispinosus x E. ferox (*Vorster 504*)	111111010100000001011100000000000000000000 000000001101100100001101000100001011100 0 101000000000110111011010110000000101001 0

	0100111100101001100010000000000000000000000
E. transvenosus x E. woodii (*Vorster 787*)	??????????????????????1000100000000000000000 1011111010000000011010000000000000000001 1000000001100000000000000000000000000000 0100110000100001000001011010000000010000 0000
Stangeria eriopus (*Vorster 337*)	1010110110110100010000100100000000000000 1000001010011011010110000001000111010010 1011000000001101110000000000000101010000 0000001100100000001110000000100010100101 110011000

APPENDIX H: The average genetic distance (D) calculated between the different species representing the DAF data sets, obtained after amplification with DAF primers: DAF2, DAF3, DAF5, DAF6, DAF10, K1, K2 and K3. (# – represents values that could not be calculated, due to division by zero).

1. *E. aemulans*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
2. *E. altensteinii*	0.51																																		
3. *E. aplanatus*	0.70	0.30																																	
4. *E. arenarius*	0.47	0.14	0.33																																
5. *E. bubalinus*	0.75	0.65	0.37	1.15																															
6. *E. caffer*	1.20	0.69	0.32	0.52	0.17																														
7. *E. cupidus*	0.72	0.30	0.65	0.42	0.77	0.85																													
8. *E.* sp. (Didinga)	0.98	0.69	0.35	0.51	0.66	0.78	0.66																												
9. *E. eugene-maraisii*	0.63	0.24	0.36	0.14	0.68	0.66	0.26	0.29																											
10. *E. ferox*	0.65	0.28	0.59	0.23	0.85	0.98	0.75	0.58	0.29																										
11. *E. friderici-guilielmi*	0.72	0.19	0.37	0.22	0.65	0.76	0.55	0.56	0.19	0.23																									
12. *E. ghellinckii*	1.52	0.93	0.95	0.98	1.47	1.46	1.31	0.71	0.99	1.13	0.69																								
13. *E. gratus*	0.51	0.15	0.41	0.11	0.66	0.93	0.46	0.57	0.18	0.13	0.24	0.73																							
14. *E. heenanii*	1.31	0.84	0.41	0.66	0.97	1.03	0.97	0.35	0.69	0.76	0.62	0.33	0.30																						
15. *E. hildebrandtii*	1.20	0.61	0.35	0.74	0.68	0.81	0.97	0.36	0.61	0.60	0.57	0.94	0.75	0.84																					
16. *E. horridus*	0.66	0.58	0.36	0.43	0.73	0.72	0.85	0.59	0.13	0.49	0.68	0.89	0.24	1.15	1.04																				
17. *E. humilis*	1.14	0.71	0.38	0.62	0.26	0.41	0.67	0.49	0.65	0.71	0.50	0.81	0.61	0.66	0.37	0.72																			
18. *E. inopinus*	0.77	0.58	0.28	0.62	0.65	0.96	0.85	0.32	0.35	0.68	0.54	0.80	0.61	0.80	0.41	0.29	0.94																		
19. *E. laevifolius*	0.67	0.41	0.83	0.28	0.77	0.88	0.98	0.89	0.34	0.34	0.53	1.28	0.51	1.28	0.79	0.48	1.14	0.64																	
20. *E. lanatus*	0.55	0.05	0.13	0.31	0.77	0.75	0.67	0.92	0.27	0.44	0.27	0.44	0.31	0.66	0.41	0.31	0.50	0.36	0.02																
21. *E. laurentianus*	1.41	0.59	0.80	0.85	0.79	0.78	1.01	0.46	0.67	0.66	0.69	0.98	0.88	0.83	0.28	1.14	1.04	0.89	1.10	1.48															
22. *E. lebomboensis*	0.55	0.37	0.36	0.37	0.76	0.75	0.63	0.53	0.39	0.28	0.39	0.86	0.38	1.06	0.52	0.42	0.81	0.52	0.32	0.47	0.77														
23. *E. lehmannii*	0.57	0.44	0.61	0.37	0.59	0.62	0.75	0.89	0.36	0.23	0.60	1.67	0.42	1.52	0.90	0.49	1.19	1.00	0.63	1.07	0.90	0.62	0.72												
24. *E. macrostrobilus*	0.83	0.63	0.35	0.62	0.39	0.57	0.98	0.56	0.59	0.44	0.56	1.13	0.70	1.22	0.27	0.50	0.99	0.51	0.51	0.98	0.82	0.58	0.52	0.43											
25. *E.* sp. (Moyo)	0.86	0.61	0.31	0.56	0.35	0.54	0.97	0.53	0.57	0.41	0.57	1.24	0.68	1.24	0.29	0.55	0.76	0.57	0.56	0.90	0.73	0.69	0.58	0.36	###										
26. *E. munchii*	0.94	0.69	0.33	0.46	1.04	0.79	0.79	0.44	0.60	0.49	0.66	0.73	0.63	0.82	0.54	0.35	1.18	0.49	0.64	0.54	1.02	0.64	0.62	0.74	0.44	0.36									
27. *E. natalensis*	0.68	0.34	0.51	0.46	0.84	0.62	0.54	0.54	0.29	0.29	0.45	1.33	0.78	1.20	0.59	0.57	1.01	0.77	0.49	0.71	1.08	0.39	0.75	0.52	0.35	0.41	0.44								
28. *E. msinganus*	1.07	0.61	0.39	0.60	0.89	0.66	1.05	0.57	0.56	0.65	0.62	0.75	0.75	1.45	0.69	0.41	1.20	0.51	0.60	0.68	0.94	0.55	0.41	0.69	0.45	0.42	0.18	0.55							
29. *E. senticosus*	0.60	0.33	0.47	0.33	0.64	0.62	0.75	0.17	0.28	0.22	0.27	1.20	0.37	1.20	0.65	0.33	0.95	0.59	0.46	0.49	1.20	0.26	0.86	0.37	0.41	0.46	0.55	0.16	0.52						
30. *E. trispinosus*	0.73	0.50	0.46	0.46	0.92	1.03	0.77	0.66	0.29	0.22	0.45	0.88	0.43	0.97	0.53	0.57	0.92	0.65	0.26	0.54	1.33	0.39	0.99	0.62	0.39	0.45	0.66	0.39	0.84	0.18					
31. *E. turneri*	1.00	0.68	0.44	0.84	0.64	0.47	1.15	0.41	0.69	0.68	0.78	0.92	0.79	1.17	0.22	0.31	0.85	0.52	0.74	1.05	0.81	0.68	0.67	0.45	0.18	0.19	0.30	0.63	0.42	0.58	0.61				
32. *E. umbeluziensis*	0.96	0.63	0.25	0.51	0.60	0.68	1.05	0.32	0.48	1.03	0.56	0.96	0.62	1.33	0.51	0.39	1.07	0.36	0.41	0.39	1.04	0.58	0.72	0.66	0.32	0.37	0.32	0.68	0.45	0.48	0.72	0.22			
33. *E. villosus*	0.84	0.56	0.23	0.43	0.66	0.69	0.61	0.37	0.49	0.29	0.53	1.14	0.34	1.06	0.43	0.37	1.00	0.87	0.15	0.87	0.83	0.59	0.63	0.63	0.21	0.25	0.37	0.52	0.55	0.42	0.59	0.50	0.09		
34. *E.* cf. *woodii*	0.71	0.23	0.39	0.36	0.81	0.74	0.64	0.83	0.28	0.42	0.34	1.46	0.48	1.20	0.84	0.83	1.24	0.81	0.38	0.64	1.61	0.62	1.30	0.66	0.53	0.65	0.65	0.63	0.62	0.28	0.48	0.74	0.29	0.26	
35. *E. senticosus x E. trispinosus*	0.89	0.48	0.45	0.41	0.62	0.53	0.74	0.52	0.52	0.43	0.39	0.92	0.69	1.12	0.57	0.42	1.05	0.80	0.52	0.80	1.12	0.60	0.65	0.69	0.56	0.58	0.42	0.61	0.54	0.52	0.60	0.66	0.25	0.38	0.57
36. *E. umbeluziensis x E. lehmannii*	0.91	0.62	0.55	0.44	0.47	0.42	0.69	0.65	0.48	0.52	0.48	1.06	0.51	1.06	0.71	0.51	1.10	0.90	0.62	0.90	1.06	0.61	0.80	0.46	0.57	0.71	0.53	0.74	0.75	0.39	0.72	0.68	0.21	0.38	0.58
37. *E. umbeluziensis x E villosus*	0.88	0.49	0.38	0.38	0.68	0.49	0.68	0.58	0.37	0.37	0.49	1.08	0.51	1.19	0.58	0.43	1.02	0.63	0.44	1.17	0.90	0.47	0.84	0.47	0.39	0.45	0.44	0.57	0.55	0.44	0.42	0.45	0.07	0.11	0.37
38. *E. altensteinii x E. trispinosus*	0.61	0.32	0.43	0.36	0.79	0.64	0.50	0.56	0.16	0.27	0.09	0.81	0.37	0.99	0.61	0.41	0.93	0.61	0.24	0.45	1.10	0.49	0.92	0.58	0.36	0.47	0.51	0.25	0.71	0.20	0.35	0.71	0.41	0.41	0.30
39. *E. trispinosus x E. altensteinii*	0.66	0.27	0.55	0.17	1.10	0.90	0.53	0.79	0.22	0.17	0.28	1.30	0.34	1.30	0.74	0.56	1.13	0.77	0.23	0.67	1.42	0.56	1.50	0.59	0.61	0.73	0.68	0.41	0.76	0.32	0.60	0.96	0.62	0.30	0.25
40. *E. trispinosus x E. ferox*	0.63	0.28	0.55	0.17	1.00	0.87	0.53	0.69	0.23	0.09	0.33	1.19	0.32	1.30	0.63	0.47	0.93	0.68	0.27	0.53	1.19	0.39	1.18	0.51	0.57	0.25	0.48	0.32	0.61	0.33	0.42	0.86	0.57	0.35	0.26
41. *E. transvenosus x E. woodii*	0.94	0.37	0.55	0.58	0.59	0.38	0.96	0.79	0.66	0.71	0.47	0.92	0.77	0.92	0.36	0.64	0.72	0.96	0.71	1.37	0.46	0.50	0.46	0.56	0.58	0.61	0.69	0.73	0.74	0.59	0.80	0.41	0.45	0.58	0.75
42. *Stangeria eriopus*	0.63	0.43	0.47	0.46	0.63	0.61	0.45	0.69	0.35	0.32	0.38	1.37	0.48	1.37	0.64	0.52	1.00	0.68	0.49	0.92	1.05	0.66	0.98	0.41	0.30	0.46	0.45	0.67	0.72	0.49	0.56	0.78	0.58	0.33	0.54

APPENDIX I: The average genetic distance (D) in a combined analysis of the RAPD and DAF data sets, representing 55 specimens. (# – represents values that could not be calculated, due to division by zero).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1. *E. aemulans*																																		
2. *E. altensteinii*	0.34																																	
3. *E. aplanatus*	0.54	0.27																																
4. *E. arenarius*	0.49	0.25	0.23																															
5. *E. bubalinus*	0.58	0.52	0.31	0.69																														
6. *E. caffer*	0.83	0.52	0.38	0.39	0.29																													
7. *E. cf. chimanimaniensis*	0.56	0.30	0.48	0.36	0.52	0.55																												
8. *E. cupidus*	0.86	0.67	0.46	0.44	0.55	0.77	0.50																											
9. *E.* sp. (Didinga)	0.48	0.27	0.36	0.25	0.54	0.52	0.25	0.36																										
10. *E. eugene-maraisii*	0.78	0.61	0.59	0.46	0.87	0.92	0.82	0.81	0.46																									
11. *E. ferox*	0.71	0.39	0.54	0.29	0.69	0.52	0.48	0.56	0.28	0.46																								
12. *E. friderici-guilielmi*	1.54	1.33	1.04	0.83	1.23	1.46	1.26	0.75	1.14	1.07	0.76																							
13. *E. ghellinckii*	0.61	0.34	0.52	0.33	0.59	0.60	0.43	0.40	0.33	0.43	0.31	0.80																						
14. *E. gratus*	0.95	0.62	0.50	0.49	0.67	0.60	0.59	0.38	0.52	0.64	0.48	0.65	0.35																					
15. *E. heenanii*	0.82	0.52	0.32	0.48	0.47	0.64	0.65	0.35	0.57	0.43	0.52	1.09	0.69	0.71																				
16. *E. hildebrandtii*	0.61	0.56	0.37	0.32	0.59	0.52	0.60	0.48	0.32	0.49	0.60	1.04	0.51	0.79	0.66																			
17. *E. horridus*	0.88	0.66	0.37	0.44	0.33	0.34	0.59	0.54	0.61	0.58	0.49	0.73	0.67	0.53	0.32	0.51																		
18. *E. humilis*	0.70	0.43	0.34	0.44	0.55	0.69	0.58	0.28	0.43	0.70	0.49	0.90	0.46	0.55	0.30	0.33	0.84																	
19. *E. inopinus*	0.74	0.56	0.64	0.21	0.65	0.74	0.71	0.69	0.46	0.53	0.52	0.82	0.70	0.97	0.63	0.51	0.77	0.51																
20. *E. laevifolius*	0.76	0.46	0.28	0.37	0.79	0.71	0.58	0.79	0.51	0.49	0.28	0.44	0.55	0.76	0.46	0.52	0.39	0.50	0.14															
21. *E. lanatus*	1.04	0.53	0.59	0.72	0.62	0.65	0.73	0.41	0.59	0.51	0.69	0.76	0.76	0.64	0.38	0.68	0.72	0.63	0.80	0.97														
22. *E. laurentianus*	0.56	0.35	0.38	0.35	0.61	0.64	0.55	0.52	0.41	0.45	0.47	0.70	0.44	0.77	0.55	0.50	0.62	0.42	0.27	0.47	0.64													
23. *E. lebomboensis*	0.70	0.58	0.53	0.57	0.46	0.42	0.86	0.49	0.60	0.62	0.69	0.61	0.67	0.69	0.32	0.36	0.63	0.44	0.64	0.65	0.41	0.29												
24. *E. lehmannii*	0.76	0.56	0.54	0.57	0.38	0.57	0.73	0.32	0.63	0.49	0.65	0.86	0.44	0.85	0.34	0.49	0.79	0.36	0.52	0.82	0.68	0.49	0.40	0.44										
25. *E. macrostrobilus*	0.88	0.65	0.57	0.55	0.48	0.61	0.85	0.43	0.67	0.55	0.73	0.92	0.56	1.00	0.41	0.66	0.62	0.56	0.63	0.71	0.71	0.57	0.49	0.47	0.02									
26. *E.* sp. (Moyo)	0.84	0.54	0.45	0.46	0.69	0.55	0.57	0.30	0.50	0.43	0.53	0.71	0.44	0.63	0.34	0.45	0.79	0.34	0.50	0.49	0.75	0.50	0.44	0.52	0.26	0.20								
27. *E. munchii*	0.60	0.39	0.47	0.45	0.62	0.60	0.47	0.31	0.38	0.43	0.52	0.97	0.61	0.80	0.52	0.54	0.74	0.51	0.41	0.62	0.62	0.32	0.42	0.40	0.29	0.35	0.34							
28. *E. natalensis*	0.88	0.60	0.37	0.47	0.60	0.52	0.78	0.56	0.56	0.61	0.67	0.68	0.65	1.04	0.56	0.41	0.78	0.42	0.47	0.60	0.64	0.44	0.41	0.58	0.45	0.36	0.26	0.55						
29. *E. nsinganus*	0.71	0.58	0.50	0.38	0.57	0.67	0.69	0.38	0.48	0.48	0.44	0.83	0.47	0.93	0.60	0.43	0.72	0.51	0.45	0.52	0.89	0.25	0.74	0.48	0.57	0.43	0.44	0.43	0.39					
30. *E. senticosus*	0.80	0.64	0.53	0.39	0.69	0.79	0.53	0.48	0.36	0.52	0.50	0.81	0.62	0.67	0.47	0.43	0.74	0.53	0.33	0.58	0.82	0.40	0.78	0.44	0.52	0.54	0.59	0.40	0.59	0.17				
31. *E. trispinosus*	0.81	0.64	0.45	0.65	0.54	0.44	0.79	0.35	0.63	0.75	0.64	0.94	0.54	0.85	0.26	0.42	0.63	0.41	0.66	0.80	0.59	0.49	0.48	0.39	0.33	0.18	0.25	0.43	0.37	0.51	0.43			
32. *E. turneri*	0.82	0.70	0.43	0.52	0.55	0.57	0.77	0.49	0.51	0.69	0.51	0.77	0.64	0.91	0.47	0.56	0.77	0.40	0.47	0.58	0.74	0.43	0.56	0.56	0.56	0.44	0.33	0.51	0.37	0.39	0.47	0.11		
33. *E. umbeluziensis*	0.71	0.70	0.55	0.51	0.53	0.58	0.51	0.45	0.47	0.32	0.55	1.00	0.49	0.77	0.43	0.48	0.73	0.65	0.34	0.74	0.79	0.51	0.52	0.50	0.37	0.33	0.36	0.52	0.51	0.41	0.45	0.39	0.07	
34. *E. villosus*	0.76	0.47	0.45	0.54	0.63	0.57	0.63	0.56	0.67	0.31	0.51	0.99	0.79	0.83	0.47	0.38	0.73	0.64	0.55	0.57	0.70	0.49	0.42	0.55	0.58	0.57	0.40	0.49	0.43	0.54	0.47	0.52	0.35	0.44
35. *E. senticosus x E. trispinosus*	0.78	0.50	0.55	0.46	0.51	0.41	0.57	0.51	0.54	0.39	0.39	0.88	0.64	0.70	0.50	0.48	0.70	0.59	0.49	0.56	0.66	0.40	0.53	0.39	0.53	0.58	0.35	0.46	0.58	0.49	0.53	0.43	0.18	0.34
36. *E. umbeluziensis x E. lehmannii*	0.73	0.39	0.41	0.37	0.59	0.43	0.54	0.56	0.46	0.46	0.53	1.00	0.53	0.80	0.43	0.58	0.80	0.49	0.49	0.90	0.67	0.41	0.62	0.42	0.37	0.34	0.27	0.46	0.51	0.54	0.36	0.32	0.16	0.17
37. *E. umbeluziensis x E villosus*	0.87	0.59	0.88	0.44	0.87	0.61	0.56	0.65	0.51	0.62	0.35	0.72	0.46	0.74	0.80	0.53	0.86	0.57	0.40	0.83	0.92	0.39	0.77	0.58	0.46	0.49	0.53	0.39	0.66	0.28	0.23	0.50	0.27	0.37
38. *E. altensteinii x E. trispinosus*	0.61	0.34	0.44	0.33	0.83	0.63	0.41	0.60	0.33	0.20	0.47	1.34	0.42	0.88	0.43	0.47	0.78	0.47	0.37	0.52	0.91	0.53	0.94	0.39	0.62	0.66	0.50	0.45	0.51	0.46	0.48	0.64	0.54	0.32
39. *E. trispinosus x E. altensteinii*	0.69	0.41	0.55	0.37	0.76	0.62	0.41	0.74	0.30	0.15	0.49	1.29	0.67	0.86	0.48	0.51	0.79	0.56	0.23	0.47	0.82	0.42	0.92	0.50	0.68	0.61	0.40	0.38	0.55	0.54	0.35	0.72	0.49	0.42
40. *E. trispinosus x E. ferox*	0.71	0.38	0.46	0.58	0.57	0.39	0.75	0.70	0.53	0.49	0.73	0.80	0.77	0.74	0.47	0.60	0.69	0.72	0.66	1.05	0.42	0.39	0.45	0.43	0.60	0.53	0.48	0.61	0.57	0.75	0.62	0.39	0.39	0.55
41. *E. transvenosus x E. woodii*	0.98	0.73	0.63	0.86	1.07	0.89	0.83	1.00	0.71	0.55	0.86	1.38	0.89	1.12	1.09	0.95	1.17	0.80	0.75	0.83	1.18	1.04	0.98	0.79	1.00	0.86	1.08	0.90	1.03	1.42	1.11	1.10	0.84	0.35
42. *Stangeria eriopus*	0.72	0.44	0.62	0.32	0.81	0.88	0.84	0.60	0.59	0.69	0.59	1.29	0.83	1.12	0.60	0.66	0.99	0.72	0.32	0.83	1.09	0.53	0.92	0.65	0.60	0.65	0.79	0.58	0.70	0.69	0.56	0.69	0.64	0.58