A phylogenetic study of South African species of *Rhynchosia* (Phaseoleae, Fabaceae)

by

Tlou Sinah Manyelo

A dissertation submitted in fulfilment of the requirements for the degree

of

Masters in Botany

in

Botany & Plant Biotechnology

Faculty of Science

UNIVERSITY OF JOHANNESBURG

Supervisor: Prof. Annah Ntsamaeeng Moteetee

Co-Supervisor: Dr James Stephen Boatwright

November 2014

DEDICATION

This thesis is dedicated to my mother Kgadi Marriat Manyelo, and my kids Thabo, Tebogo and Ofentse, with love.

ACKNOWLEDGEMENTS

First, I would like to thank Almighty God through Jesus Christ, for strength, protection and wisdom. I am greatly indebted to my supervisor, Prof. Annah N. Moteetee and co-supervisor Dr J. Stephen Boatwright for their insightful, intellectual guidance and detailed constructive comments. I would also like to extend my sincere gratitude to the following:

Prof. B-E van Wyk, Prof. P. Tinley and Dr A. Magee: for their advice as well as their constructive comments.

Dr O. Maurin and Mr. B. Daru: for their help with phylogenetic analysis

Prof. M. van der Bank: for laboratory facilities, as well as the staff and students of the African Centre for DNA Barcoding, University of Johannesburg, for their helpful discussions.

The National Research Foundation for funding towards the study.

The South African National Biodiversity Institute for study material.

My family for their love, support and encouragements, I love you all.

ABSTRACT

Rhynchosia is a genus within the tribe Phaseoleae, a group of papilionoid legumes that are economically important both for human and animal consumption because of their high protein content. The genus is pantropically distributed and comprises ca. 230 species. South African species were last studied by Baker in 1923. He identified five sections based on their morphological structures. No molecular studies have been carried out on the species to support the taxonomic classification. A systematic study of the relationship within the species was undertaken using five markers; *matK, rbcL,* trnH-*psbA,* rpl32-*trnL* and Internal Transcribed Spacer (ITS) as well as morphological characters, to investigate phylogenetic relationships in the genus *Rhynchosia*. Results obtained from a combined analysis produced four clades, two of which (clade three and four) received strong support from Bayesian Inference analyses, two of which received low BI support and therefore referred to as groups. Sections *Chrysoscias*, *Polytropia*, *Arcyphyllum* and *Cyanospermum* are embedded within a paraphyletic type section *Rhynchosia*. The results further suggest that sections *Polytropia* and *Chrysoscias,* which possess a high density of glands on the leaves, are very closely related forming a clade with species from section *Rhynchosia*. The morphologically distinct species *R. monophylla* came out as sister to all four clades.

TABLE OF CONTENTS

INDEX TO TABLES

Chapter 1

Chapter 2

Chapter 3

INDEX TO FIGURES

Chapter 1

Chapter 2

Chapter 3

LIST OF ABBREVIATIONS

ABI	Applied Biosystems, Inc.
AIC	Akaike Information Criterion
BBH	Bulbous based hairs
BI	Bayesian inference
BP	bootstrap percentage
Bp	base pair
BSA	bovine serum albumin
Ca.	Approximately
CCBD	Canadian Centre for DNA Barcoding
CI	consistency index
Cp	Chloroplast
CTAB	hexadecyltrimethylammonium bromide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
etOH	ethanol / ethyl alcohol
F	forward primer
ITS	internal transcribed spacer
JRAU	University of Johannesburg (UJ) herbarium Johannesburg, South Africa
LSC	large single copy-region
matK	maturase K
MCMC	(Bayesian) Markov Chain Monte Carlo
MP	maximum parsimony
MulTrees	multiple equally parsimonious trees

NaCl	Sodium chloride
NE	North East
No.	Number
NRF	National Research Foundation (South Africa)
PAUP	phylogenetic analysis using parsimony software program
PCR	polymerase chain reaction
PBS	partitioned Bremer support
PIC	percentage informative character
PP	posterior probabilities
PRE	National Herbarium, Pretoria (PRE)
PVP	Polyvinyl pyrolidone
R	reverse primer
rbcL	riboluse-bisphosphate carboxylase gene
RI	retention index
S	South
SANBI	South African National Biodiversity Institute
sect.	Section
s.l.	sensu lato
s.s.	sensu stricto
SSC	small single copy-region
SW	South West
TBR	tree-bisection-reconstruction
TIM+I+G	Transitional model + Inverse + Gamma
TL	tree length
Trn+G	Tamura-Nei plus Gamma

trnH-psbA	spacer between *trnH* and *psbA* genes
rpl32- trnL	spacer between *rpl32* and *trnL*
Tvm+G	transversion model plus gamma
W	West
WC	Western Cape
µl	microlitre(s)

Chapter 1

General introduction

1.1 Family Leguminosae/ Fabaceae

Fabaceae is the third largest angiosperm family following the daisies (family Asteraceae) and orchids (family Orchidaceae) with more than 19,500 species in about 751 genera (Lewis et al., 2005; LPWG, 2013). It also ranks second to grasses (family Poaceae) in economic, agricultural and ecological importance (Wojciechowski et al., 2004). Legumes are used agriculturally for human consumption (grain legumes) and animal feed (forage and pasture legumes) (Graham and Vance, 2003) because they have high protein content (Murthy and Rao, 2008).

Legumes are predominantly diverse in seasonally dry tropical forests and in temperate shrublands adapted to xeric climates. On the other hand they are poorly represented in mesic temperate habitats (Wojciechowski et al., 2004). All legumes play an important role in the terrestrial nitrogen cycle irrespective of whether or not they form root nodules (Sprent, 2001), and their habitat preference is driven mostly by their nitrogen demanding metabolism (McKey, 1994).

The family is divided into three subfamilies based mainly on floral structure. These are **(i)** the **Caesalpinioideae** which have more or less zygomorphic flowers with the adaxial petal overlapped by the adjacent lateral petals; **(ii)** the **Mimosoideae** which have actinomorphic flowers often with an increased numbers

stamen and **(iii)** the **Papilionoideae** which have zygomorphic flowers, with the adaxial petal outside the adjacent lateral petals (Käss and Wink, 1996).

1.1.1 Subfamily Papilionoideae

Papilionoideae is the largest subfamily comprising ca. 478 genera and ca. 13,800 species across 28 tribes and includes most of the familiar “beans and peas” (Lewis et al., 2005). Apart from floral structure, the subfamily Papilionoideae can easily be distinguished from other subfamilies by vegetative and fruit characters (Polhill, 1981; Tucker, 2002). The subfamily Papilionoideae is easily distinguished from the other sub-families by traits that are now considered to be morphological synapomorphies. Examples include vegetative and fruiting characters (Polhill 1981; Tucker, 2002) and especially orientation of the seed hilum and unidirectional initiation of sepals (Doyle et al., 2000). Twenty-eight tribes are recognised in the subfamily (Table 1.1).

Table 1. 1 The tribes recognised in the subfamily Papilionoideae

Tribe	No. of genera	No. of species
Abreae (Wight & Arn. Ex Endl.) Hutch.	1	17
Amorpheaea Boriss	8	245–248
Bossiaeeae (Benth.) Hutch.	6	72
Brongniartieae (Benth.) Hutch.	10	150–152
Cicereae Alef.	1	43
Crotalarieae (Benth.) Hutch.	16	ca. 1208
Dalbergieae Bronn ex DC.	49	1319–1331
Desmodieae (Benth.) Hutch.	30	524–530
Dipterygeae Polhill	3	ca. 22
Euchresteae (Nakai) H. Ohashi	1	4
Fabeae Rchb.	5	327–330
Galegeae (Bronn) Dumort	24	2882–3183
Genisteae (Bronn) Dumort	25	551–572
Hedysareae DC.	12	400–453
Hypocalyptieae (Yakovlev) A.L.Schutte	1	3

Tribe	No. of genera	No. of species
Indigofereae Benth.	7	768
Loteae DC.	22	282
Milletieae Miq.	45	904–914
Mirbelieae (Benth.) Polhill & Crisp	25	686–689
Phaseoleae (Bronn) DC.	89	1554–1580
Podalyrieae Benth.	9	ca. 133
Psoraleeae Lowe	10	ca. 188
Robinieae (Benth.) Hutch.	11	71–72
Sesbanieae (Rydb.) Hutch.	1	Ca. 60
Sophoreae Spreng. ex DC.	48	391–396
Swartzieae DC.	17	ca. 258
Thermopsideae Yakovlev	6	44–47
Trifolieae (Bronn) Endl.	6	485

1.1.2 Tribe Phaseoleae

Phaseoleae are generally dextrorotatory twiners, rarely prostrate to erect or shrubby, sometimes arborescent. In addition they are generally recognised by leaves with asymmetrical lateral leaflet margins, variable leaflet shape including linear, ovate, as well as lobed forms (Lackey, 1981, Doyle and Doyle, 1993), and by the inflorescence ranging from panicles to nodose-pseudoracemes and pseudoracemes.

Phaseoleae, second largest from Galegeae and arguably the most economically important tribe of the subfamily (Wojciechowski et al., 2004). The tribe Phaseoleae is presently considered to comprise about 1580 species in nearly ca. 89 genera. Morphological analyses have revealed the polyphyletic and paraphyletic nature of the tribe, and two major clades are evident: (i) The Millettioid *s.s* group which includes Abreae, the core-Milletieae together with the Phaseoleae subtribes Diocleinae and Ophrestinae. (ii) The Phaseoleae *s.l.* clade includes the subtribes Kennediinae Benth., Cajaninae Benth., Phaseolinae Bronn, Glycininae Burnett and Clitoriinae Benth. as well as the two traditionally independent tribes the Desmodieae

and the Psoraleeae (Figure 1.1; Lewis et al., 2005), also suggested that tribes Desmodieae and Psoraleeae should perhaps be treated as subtribes. Analysis of *matK, rbcL* and ITS in the milletioid group has added more clarity to the relationship in the Old World clade, and it is also evident that both Phaseoleae and Milletieae *s.l* are polyphyletic and await a complete revision at a tribal level (LPWG, 2013).

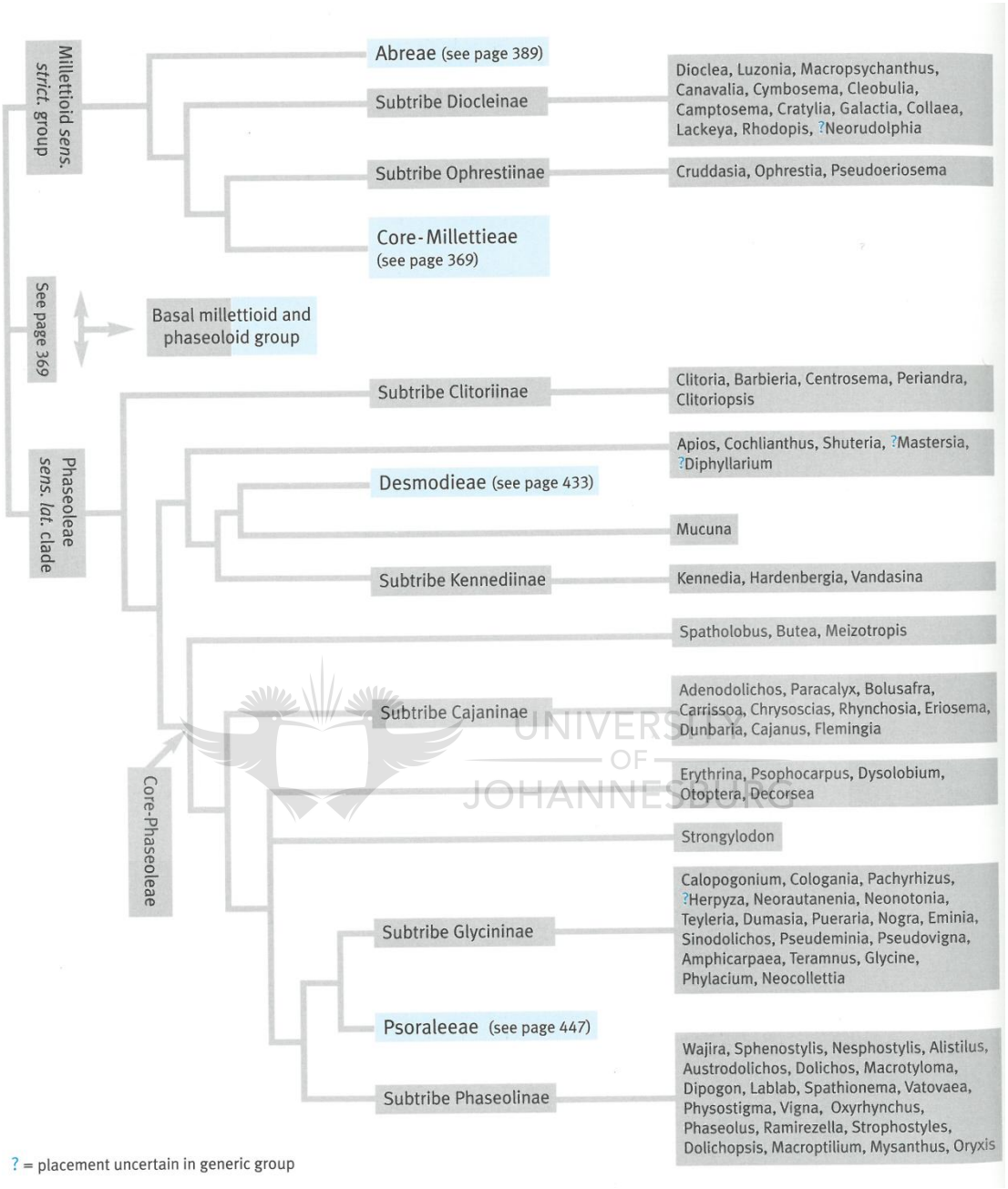


Figure 1.1 A schematic representation of the classification of the phaseoloids based on phylogenetic data from various sources (from Lewis et al., 2005).

1.1.3 Subtribe Cajaninae

Subtribe Cajaninae is the most distinctive group in the tribe Phaseoleae, characterised by the presence of bulbous-based hairs and vesicular glands (Lackey, 1981). These glands, which have been described as “consisting of a squat head of cells contained within a shallow depression of the epidermis”, occur on the calyces, ovaries, stems, lower surfaces of the leaves and fruits (Moteetee and Van Wyk, 2006). In addition, Cajaninae displays a symplesiomorphic character of consistent chromosome number of 11 and a total lack of canavanine (a non-proteinogenic α-amino acid found in certain leguminous plants), the absence of which is viewed as a derived condition (Lackey, 1977). The subtribe is also characterised by trifoliolate leaflets (sometimes 1, and rarely 5) and the absence of bracteoles, with the exception of *Adenodolichos* Harms (Lackey, 1981; Moteetee and Van Wyk, 2006; Kassa et al., 2012).

Although the genus *Adenodolichos* Harms seems to be distorting the perfect description of subtribe Cajaninae, Doyle and Doyle (1993) proved the monophyly of the subtribe in their phylogeny based on chloroplast DNA sequences. The subtribe comprises about 495 species in nine genera, with the exclusion of *Chrysoscias* (Moteetee et al., 2012), the genus is included in *Rhynchosia* Lour. in South African literature (Lewis et al., 2005). The monotypic genus *Carrissoa* Baker f. (represented by *C. angolensis* Baker f.) is also likely to be part of *Rhynchosia* (Lewis et al., 2005, Moteetee et al., 2012). Species of the subtribe Cajaninae reflect a range of habits with individuals occurring as shrubs, subshrubs and herbs. They are also adapted to a range of ecological patterns from dry tropics to subtropics, open forests to grasslands, and woodland thickets (Table 1.2; Lewis et al., 2005).

Table 1. 2 Biogeographical distributions, economic importance and characteristics of genera in subtribe Cajaninae (modified from Lewis et al., 2005)

Genus	No of species	Distribution	Economic importance	Characteristics
Adenodolichos **Harms**	ca. 15-20	Tropical Africa (Guinea, Nigeria, Sudan and Cameroon).	Used for human food (vegetable), medication and cordage.	Herbs or shrubs; seasonally dry tropical woodland, scrubs, wooded grassland as well as grassland; often appearing soon after fire.
Paracalyx **Ali**	6	NE Africa, Indian subcontinent.	Used as a food source, has high protein content. Also used for cancer treatment.	Herbs; seasonally dry tropical forest, thicket bushland and scrubs, often in rocky areas.
Bolusafra **Kuntze**	1	South Africa (Western Cape).	Ornamental, also used as animal food	Herbs or subshrub; Mediterranean montane sclerophyllous shrubland (fynbos), often along stream sides.
Carrissoa **Baker f.**	1	SW Africa (Angola).	Insufficient literature.	Subshrub; seasonally dry tropical open bushland and scrubs, or along watercourses; with woody rootstocks in fire prone environments.
Rhynchosia **Lour.**	ca. 230	Pantropical. (ca. 140 spp. in Africa and Madagascar; ca. 55 endemic spp. in tropical and subtropical America).	Used as famine foods, for making necklaces, also as narcotics.	Herbs, vines or subshrubs; seasonally dry forest; many species are pyrophytes or fire-weeds.
Eriosema **(DC.) Rchba**	ca. 150	(ca. 100–110 spp.) Africa and Madagascar (ca. 40 spp.) N and S America Mexico to N Argentina, (2 spp.) SE Australia and (30 spp.) S America.	Used as famine foods, medicine and fish poison.	Herbs or shrubs; seasonally dry tropical to subtropical forest margins, woodland, grassland and old cultivation as well as waste ground.
Dunbaria **Wight & Arn.**	20	SE Asia, ca. 2 spp. extend to N Australia.	Used for green manure and medicine, also for pigeon pea breeding programmes.	Climbing herbs or subshrub; seasonally dry tropical forest, thicket, scrubs and grassland.
Cajanus **DC.**	34	SE Australia, W Africa, 2 spp. widespread in old world tropics.	Used as food, or fodder crop, considered important for development of improved cultivars.	Herbs or subshrub; seasonally dry tropical open forest to grassland, often in rocky or partly disturbed areas; widely cultivated in the old and new world tropics.
Flemingia **Roxb. ex W.T.Aiton**	ca. 30–35	SE Asia to Australia and Africa.	Used as green manure, medicine, the underground tuber is used for treating human consumption.	Herbs or subshrub; seasonally dry tropical forest, woodland and grassland, sometimes near swamps and streams also in disturbed and waste areas.

1.2 Selecting appropriate phylogenetic markers for the legumes

DNA sequence data have been widely used in phylogenetic studies because of its importance to construct a classification or to infer a phylogeny of plants (Fay et al., 1998), as opposed to morphological character usage which can be difficult?to distinguish because of characters similarity overlaps. Gene choice is based on various factors such as copy numbers, character congruency (low change of homoplasy), ability to resolve relationships, suitable number of parsimony informative characters, and rate of evolution relative to the taxonomic group (Mort et al., 2007). Given that different DNA loci evolve at different rates, it is important to choose the loci for the study group that best fit the mentioned factors (Fay et al., 1998). For plants, the plastid gene *rbcL* has been the most widely used marker (e.g. Chase et al., 1993; Olmstead and Palmer, 1994; Doyle et al., 1997; Doyle et al., 2000, Richardson et al., 2000; Savolainen et al., 2000; Kajita et al., 2001; Sulaiman et al., 2003; Muller et al., 2006). The nuclear gene region ITS as well as *matK* (another chloroplast gene) and *rbcL* are very popular genomic regions for inferring plant phylogenies at different taxonomic levels, due to their structure and pace of molecular variation (Olmstead and Palmer, 1994; Shaw et al., 2005).

Molecular data from DNA sequences began to influence ideas of relationships within legumes in the late 1980s and early 1990s (LPWG, 2013). Such investigations include for example: phylogeny of the subfamily Papilionoideae (Käss and Wink, 1995, 1997b), phylogeny of the three subfamilies based on *rbcL* sequences (Käss and Wink, 1996), phylogeny of the chloroplast gene *rbcL* in the Leguminosae (Doyle et al., 1997), *rbcL* phylogeny with particular reference to Phaseoleae, Millettiae and allies (Kajita et al., 2001), and a phylogeny of the legumes based on analysis of the

plastid *matK* gene sequences (Wojciechowski et al., 2004). This ground-breaking research was followed by the utilization of the different chloroplast gene regions at lower taxonomic levels, including *rbcL*, *matK*, *trnL-F*, *ndhF, rps16* and several others (Davis et al., 2002; Haston et al., 2005; Kassa et al., 2012; Shaw et al., 2007; Steele and Wojciechowski, 2003; Wojciechowski, 2005; Wojciechowski et al., 1999). Figure 1.1 represents a summarised version of the phaseoloid classification based on various phylogenetic studies i.e. *rbcL* for Phaseoleae, Millettiae and close relatives (Kajita et al., 2001), phylogenetic analysis in the tribe Millettiae (Hu et al., 2002) and for the family (Lewis et al., 2005; Wojciechowski et al., 2004).

1.2.1 Review of the genus *Rhynchosia*

The name *Rhynchosia* was first introduced by Loureiro in 1790 to accommodate one species, *Rhynchosia volubilis* Lour. For a period of 60 years from 1899, *Rhynchosia*, commonly known as snout bean (Lackey, 1981), was included as a synonym under the name *Dolicholus* Medikus until 1959, when it was conserved against the latter (Woods and Key, 2009). The genus comprises about 230 species (Lewis et al., 2005) centred mainly in Africa and Madagascar but also extends to warm temperate and tropical Asia, northern Australia, tropical and subtropical America (Grear, 1978, Lewis et al., 2005).

After 1790, several botanists then contributed to the growing list of species, including De Candolle (1825) with 51 species, Ecklon and Zeyher (1836) who recognised 11 species, Meyer (1836) with 19 species, Harvey (1862) as well as Bentham and Hooker (1865) each having recognised 41 species. The South African species of the genus *Rhynchosia* were last studied comprehensively by Baker (1923), who recognised 59 species and separated the species into five sections namely: *Cyanospermum* (Wight and Arn.) Benth. (one spp.), *Chrysoscias* E.Mey. (four spp.), *Polytropia* Presl. (two spp.), *Arcyphyllum* Torrey and Gray (two spp.) and *Rhynchosia* Lour. (referred to as *EuRhynchosia*) (50 spp). Since his revision of the genus, several species have been described in the recent checklist by Germishuizen (2006), and some other species in Boatwright & Moteetee (2014).

Rhynchosia species, particularly those in the type section, are morphologically highly variable, even within species and are often difficult to distinguish from each other. On the realization of these intricacies, Verdcourt (1971), in his treatment of the genus for Flora of Tropical East Africa, divided it into smaller morphological groups and in reference to one of his groups (the *R. minima* group). He indicated that “this is unquestionably the most difficult group within the genus from the point of view of taxonomic complexity”. More recently taxonomic studies have been initiated (for example Boatwright and Moteetee, 2014; Moteetee et al., 2012), to revise the entire genus in South Africa. The present study was aimed at exploring relationships in the genus based on molecular and morphological data.

1.2 Aims of the study

The aims of this study are to produce, for the first time, a molecular phylogeny of the genus *Rhynchosia* to determine:

1. If the genus is monophyletic.
2. If the genus *Chrysoscias* is embedded within *Rhynchosia* or whether it deserves recognition at generic rank.
3. If molecular data support Baker's sectional classification.
4. The relationships between the species of *Rhynchosia.*
5. If the relationships revealed by the molecular data are supported by morphological data?

Chapter 2

Materials and Methods

2.1 DNA regions used in study

Phylogenetic analysis based on a single molecular marker could result in misleading results, since these phylograms represent gene trees and do not necessarily reflect the phylogeny of the corresponding species (Doyle, 1992). In the present study DNA regions from both nuclear (ITS) and chloroplast (*matK, rbcL, trnH-psbA*, and *rpl32-trnL*) genes were sequenced and analysed.

2.1.1 The matK region

The maturase K (*matK*) gene is a protein coding gene of approximately 1570 bp in length, located within domain IV of the intron for the transfer RNA gene of Lysine ($trnK^{uuu}$), except in some ferns where it codes for tRNAlys (uuu) (Neuhaus and Link, 1987). The *matK* gene is ubiquitous in nearly all plants species (Sugita et al., 1985; Neuhaus and Link, 1987, Ems et al., 1995, Steane, 2005, Daniell et al., 2006) with characteristic insertions and deletions (indels) in multiples of three (Barthet and Hilu, 2007), and has recently emerged in plant systematics as the most valuable gene because of its high variability compared to other genes used in phylogenetics (Müller et al., 2006). The *matK* gene has been extensively used in several legume phylogenetic studies ranging from family to intraspecific relationships, e.g. monophyly of Papilionoideae (Wojciechowski et al., 2004), phylogenetic systematics of the tribes Trifolieae and Vicieae (Steele and Wojciechowski, 2003) and tribe

Millettieae (Hu et al., 2000). An evolutionary rates analysis of the family Leguminosae has shown that the core Phaseoleae lineages have the fastest rate of nucleotide substitution for *matK* (Lavin et al., 2005, Delgado-Salinas et al., 2006). Even though frameshift indels as well as premature stop codons could render *matK* non-functional in some plants (Barthet and Hilu, 2007). The gene has been proposed as one of the core barcodes for land plants (CBOL Plant Working Group, 2009), and researcher were advised to build up to existing data by making use of both core plastid markers (*matK* and *rbcL*) in addition to other markers (CBOL Plant Working Group, 2009), hence this region was chosen for this study.

2.1.2 The rbcL region

The slowly evolving gene *rbcL*, has been used widely in investigating relationships between genera, families and orders (Fay et al., 1997). Ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco) is a key enzyme in the photosynthetic carbon metabolism in prokaryotic and eukaryotic autotrophs (Tabita, 1999), and is composed of a large and small subunits (Figure. 2.3; Chase et al., 1993; Wakasugi et al., 2001). The large subunit contains active sites for enzyme activities and is highly conserved antigenically from cyanobacteria to higher plants (Curtis and Glegg, 1984). In a DNA molecule, one copy of the gene coding for large subunit (*rbcL*) is present per chloroplast. The sequences have been utilized successfully in a number of phylogenetic studies in the family Fabaceae such as the phylogeny of the papilionoid legume tribe Phaseoleae (Doyle and Doyle, 1993), the phylogenetic relationships of the three subfamilies of the Leguminosae (Kajita et al., 2001; Käss and Wink, 1996; Sulaiman et al., 2003) and the phylogeography of the genus *Lupinus* L. (Käss and Wink, 1997).

2.1.3 *trnH-psbA* intergenic spacer

This is a non-coding region and one of the most variable in angiosperms; easily amplified across a broad range of land plants, with a size of about 500 bp (Shaw et al., 2007). It consists of two regions differing in their evolutionary characteristics: 1, *psbA* 3'UTR (untranslated region) and 2, *trnH-psbA* intergenic transcriber, which is highly variable in length and composition (Hao et al., 2010).

2.1.4 *rpl32 -trnL* intergenic spacer

This intergenic spacer is located in the small single copy (SSC) region of the chloroplast genome (Shaw et al., 2007) with an average length of 1018 bp, and ranges from 524 to 1417 bp. In Shaw et al. (2007), the region showed high average values in both potentially informative character (PIC) and percentage variability when compared to *ndhF-rpl32* and *ndhA* intron.

2.1.5 The ITS region

The Internal Transcribed Spacers (ITS) are non-coding and known to evolve rapidly, with high copy numbers, availability of universal primers and higher sequence variation than plastid regions making them more suitable for lower taxonomic level studies (Hamby and Zimmer, 1992, Wendel et al., 1995, Álvarez and Wendel, 2003). In particular, ITS1 and ITS2 offer higher levels of species discrimination and is now used extensively with other barcoding regions. ITS has, therefore, been used to study relationships at interspecific and intergeneric levels (Baldwin et al., 1995). ITS sequences have been successfully utilized in several legume phylogenetic studies including for example: *Lupinus* L. (Aïnouche and Bayer, 1999), *Cercis* L. (Davis et

al., 2002), *Stylosanthes* Sw. (Stappen et al., 2002), *Astragalus* L. (Osaloo et al., 2005), *Phaseolus* L. (Delgado-Salinas et al., 2006), *Lotononis* (DC.) Eckl. & Zeyh. (Boatwright et al., 2011).

2.2 Taxon sampling

For this study a dataset of 49 species from various herbaria and localities around South Africa were sampled and examined (Table 2.1). These included 40 species from section *Rhynchosia*, three (3) species from section *Polytropia*, two (2) species from section *Chrysoscias*, and one (1) species each from sections *Arcyphyllum* and *Cyanospermum* as well as two (2) species of *Eriosema* (DC.) G.Don, which were used as outgroups.

2.2.1 Outgroup selection

Some studies (e.g. Doyle and Doyle, 1993; Fortunato, 2000), have shown that within the subtribe Cajaninae, the genus *Eriosema* is the most closely related to *Rhynchosia*, while others (e.g. Kassa, 2012; Figure 2.1) indicate that *Bolusafra* is the most closely related genus. *Eriosema* and *Bolusafra* were both chosen as outgroups, however using *Bolusafra* was abandoned because the relationship within ingroup taxa could not be clearly resolved as it showed less variation. Morphologically *Eriosema*, *Bolusafra* and *Rhynchosia* have been found to be closely similar based on their number of ovules (two) in the ovary (Baudet,1978), but differ in the location of funicular attachment; in *Rhynchosia* the funicle is attached in the middle of the hilum and in *Eriosema* it is attached at the end of the linear hilum (Lackey, 1981), while *Bolusafra* was described according to Lackey (1981) as “viscid *Rhynchosia* – like vine”, but differs from *Rhynchosia* in having prominent seed arils.

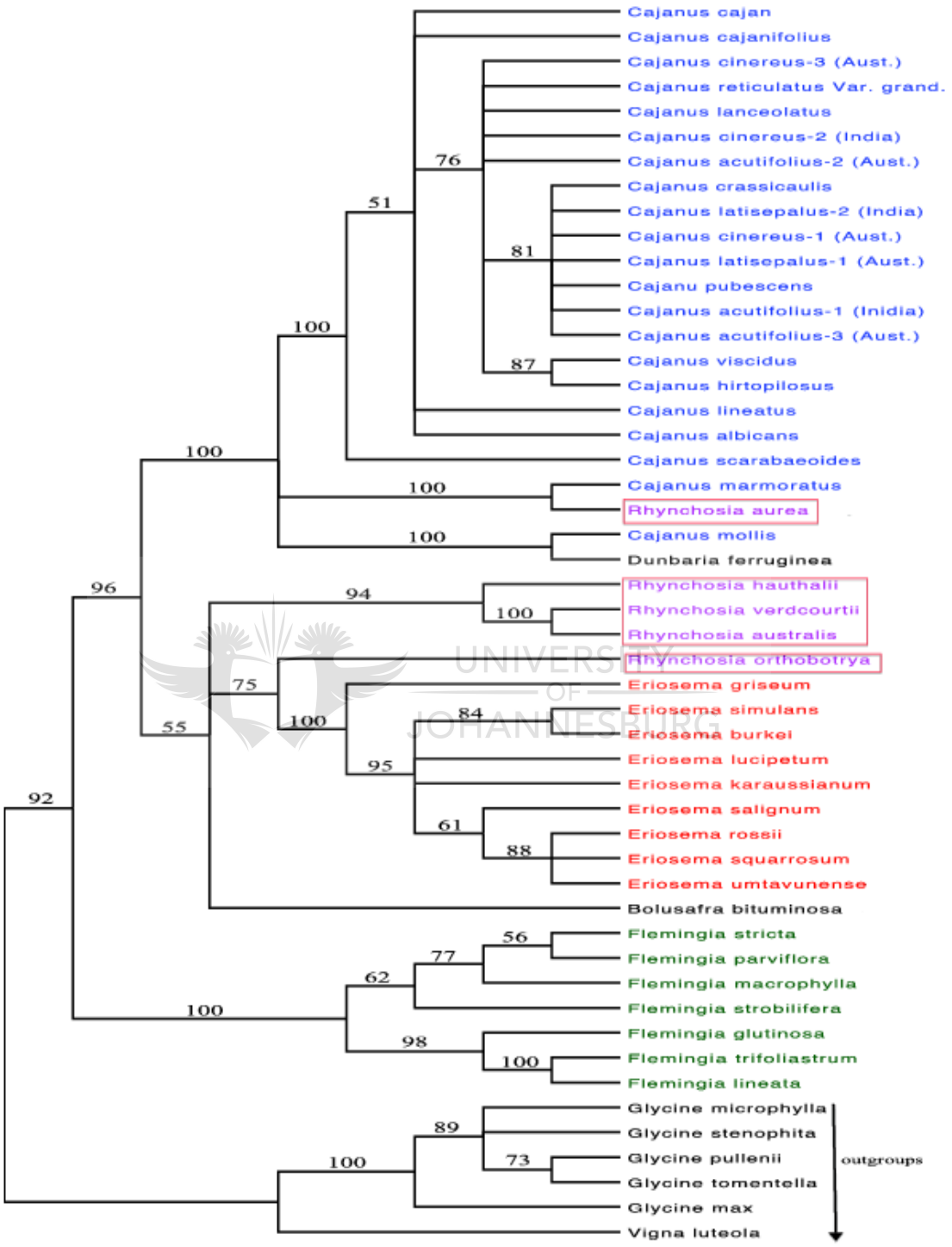


Figure 2. 1 The relationship between *Rhynchosia* and *Eriosema* (adopted from Kassa et al., 2012).

2.3 DNA extraction, amplification and Sequencing

2.3.1 Extraction and purification

Total genomic DNA was extracted from herbarium and silica dried leaf material following the hot cetyltrimethylammonium bromide (CTAB) protocol of (Doyle and Doyle, 1987). Polyvinyl pyrolidone (2 % PVP) was added to samples to reduce the effect of high polysaccharide concentrations. Extracted DNA was purified using QIAquick purification columns (QIAgen, Inc., Hilden, Germany) following the manufacturer's protocol.

2.3.2 Amplification

Polymerase Chain Reaction (PCR) amplifications were performed using either a 9800 Fast Thermal Cycler or a GeneAmp PCR System 9700. Primers used for polymerase chain reaction (PCR) amplification of the cpDNA regions *matK* [kim: ki-Joong-3F-ki-Joong-1-R (CBOL Plant Working Group, 2009), matK472F-matK1248R (Figure 2.2; Chase et al., 2005, Yu et al., 2011)], *rbcL* [rbcLa-F, 1F, rbcLa-R and 724R (CBOL Plant Working Group, 2009)], *trnH-psbA* [*trnH-psbA* (Sang et al., 1997, Tate and Simpson, 2003, Shaw et al., 2007)] and *rpl32-trnL* [rpl32-trnL (Shaw et al., 2007)] are listed in Table 2.2. The region ITS was amplified using primers AB101 and AB102 (external primers) together with ITS2 and ITS 3 (internal primers) (Figure 2.4; White et al., 1990, Sun et al., 1994.). The same primer combinations were also implemented in cycle sequencing (Table 2.2).

PCR amplification for *matK* and *rbcL* was executed at the Canadian Centre for DNA Barcoding (CCDB) and details of the project such as voucher information, pictures and DNA barcodes are available on the BOLD website (http://www.boldsystem.org). I performed additional amplification for *matK* and other

regions at the African Centre for DNA Barcoding (ACDB) at the University of Johannesburg, South Africa. For all reactions I used green premix (Ready Master mix; Advanced Biotechnologies, Epsom, Surrey, UK). Dimethyl sulfoxide (DMSO) was used in ITS reactions to reduce problems encountered in amplification due to the secondary structure of ribosomal DNA (Álvarez and Wendel, 2003) was used in ITS reactions. DMSO is a polar aprotic solvent that is used in PCR to inhibit formation of secondary structures in the DNA template or the DNA primers. It is added before the PCR reaction, to interfere with the self-complementarity of the DNA. Since most of the samples utilized in the study were herbarium material, having a high probability of degraded DNA material (Savolainen et al., 1995), Bovine Serum Albumin (BSA) which is used to stabilize enzymes during DNA digestion, was added to PCR mixtures. In cases where amplification seemed more difficult, 150 µl of total DNA were further purified and concentrated through QIAquick Silica columns (Qiagen, inc) following the protocol for cloning PCR products.

Amplification was carried out using PCR in 25 µl reactions containing 18.1 µl PCR mastermix [50 units/ml of Taq DNA Polymerase (pH 8.5), 400µM each of dTTP, dCTP, dATP, dGTPand 3 mM $MgCl_2$ (Promega Corporation)]; 0.3 µl of both forward and reverse primers (Table 2.2 for primer sequences and references as well as Figure 2.2, 2.3 and 2.4 for diagrams of gene regions); 0.8 µl BSA, 0.5 µl DMSO for nuclear reactions and 5 µl template.

Protocol for *rbcL* amplification consisted of a pre-melt of 3 min at 94°C followed by 28 cycles of denaturation of 1 min at 94°C, annealing of 1 min at 48°C, and extension of 1 min at 72°C, and a final extension at 72°C for 7 min; the *matK* amplification protocol consisted pre-melt of 3 min at 94˚C followed by 30 cycles of denaturation of 1 min at 94°C, annealing for 1 min at 52°C and extension of 2 min 30

sec at 72°C with a final extension at 72°C for 7 min. For regions *trnH-psbA and rpl32-trnL,* a protocol from (Shaw et al., 2007) was utilised. The amplification protocol for ITS followed a pre-melt of 1 min at 94°C, 26 cycles of denaturation of 1 min at 94°C, annealing of 60 sec at 48 °C and extension of 3 min at 72°C then a final extension of 7 min at 72°C. Where amplification was hard to achieve cycles were increased to 35. PCR products were then checked by electrophoresis in a 1.5% agarose gel and visualized with ethidium bromide on a UV transilluminator. Amplified products were purified using Qiaquick columns (QIAgen, Germany) following the manufacturer's protocol.

2.3.3 Cycle sequencing

Cycle sequencing was carried out using BigDye© V3.1 Terminator Mix (Applied Biosystems, Inc., ABI, Washington, Cheshire, UK) in 10 µl reactions using the following reagents: 1.5 µl 5X sequencing buffer, 0.3 µl primer, 0.3 µl big dye terminator, 0.5 µl DMSO (only in ITS), 0.1-1.0 µl template and ddH_2O to make up for the total volume. Cycle sequencing reactions were carried out in 26 cycles for 10 seconds denaturation at 96ºC, 5 seconds annealing at 50ºC and 4 minutes extension at 60ºC. Products were cleaned following the EtOH-NaCl method provided by ABI; and then sequenced on an ABI 3130xl genetic analyser.

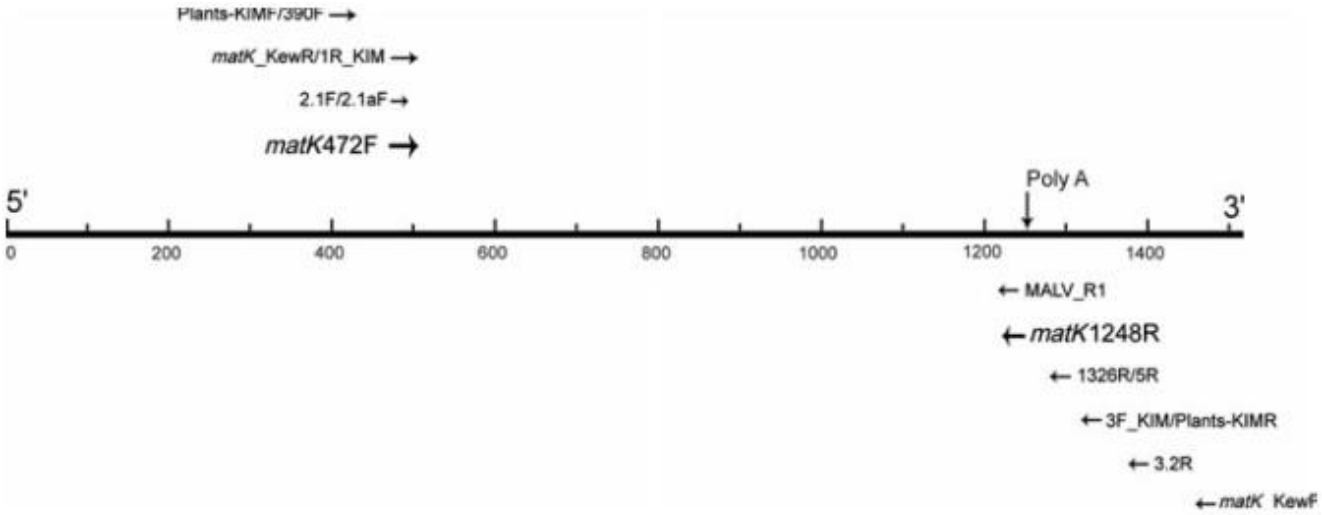


Figure 2. 2 A schematic representation showing pairs of *matK* primers. Pair *427F* and *1248R* was successfully sequenced in the study. Diagram obtained from Yu et al. (2011).

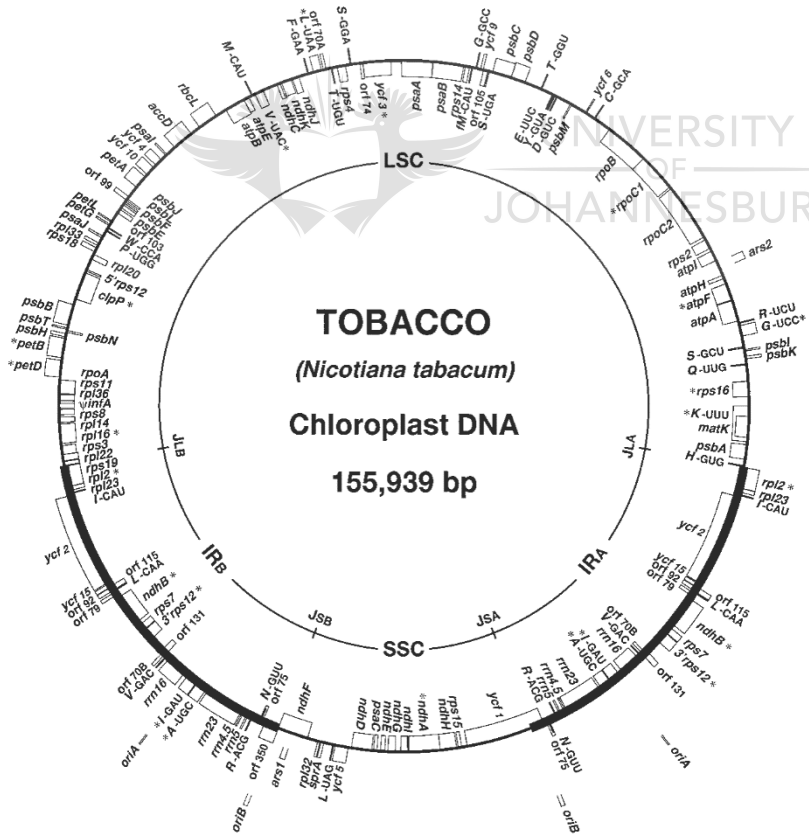


Figure 2. 3 Diagram of the tobacco chloroplast genome (Wakasugi et al., 2001) showing two inverted repeats (IR_A and IR_B), a large single copy (LSC) and small single copy (SSC) region

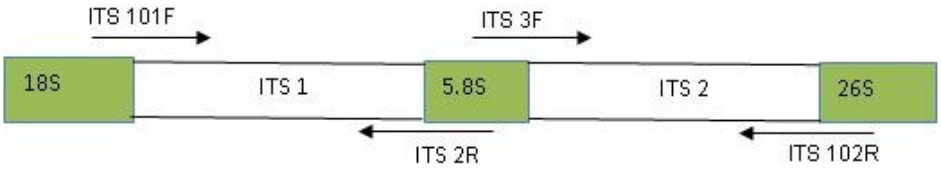


Figure 2. 4 Schematic diagram of the nuclear ribosomal DNA internal transcribed spacer (ITS) region.

2.4 Phylogenetic analysis

Editing and contig assembly of sequences were carried out using Sequencher™ version 4.6 (Gene Codes Corporation), while sequence alignments were done manually using the PAUP version 4.0b1 (Swofford, 2002) programme on a Macintosh G4. From five regions that were tested, three (*matK*, *rbcL* and ITS) yielded good results with amplification across high numbers of species. These results were used in the final analysis. The other two regions *trnH-psbA* and *rpl32-trnL* did not amplify for most samples and results obtained from them will be used towards the next level of the study.

2.5 Parsimony analysis

Maximum parsimony (MP) analyses for all data sets were performed using PAUP* version 4.0b1 (Swofford, 2002). All characters were unordered and weighted equally with gaps treated as missing data. Phylogenetic analyses were carried out independently on sequences from *matK*, *rbcL* and ITS and on a combined data set of all regions (*matK*, *rbcL* and ITS). Heuristic searches included 1000 replicates, of random taxon addition, retaining 10 trees at each step, with tree-bisection-

reconnection (TBR) branch swapping and MulTrees in effect. Clade support was evaluated through bootstrap (Felsenstein, 1985) with the following criterion used as classification of bootstrap support (BP): high/strong (85-100%), moderate (75-84%) and low (50-74%). Congruency among data partitioning was evaluated through TreeRot (Sorenson and Franzosa, 2007) which aids in determination of decay and Bremer support indices (Bremer, 1988). A negative Bremer index is an indication of incongruency between plastid and nuclear regions, while a positive index indicates congruency, no hard incongruence observed therefore allowing for nuclear and plastid gene partitioning. Visual comparison between the topologies and bootstrap values was also used as support to partition the analysis.

2.6 Bayesian analysis

Bayesian inference (BI) of phylogeny using the Markov Chain Monte Carlo (mcmc) method was computed using MrBayes version 3.1.2 (Huelsenbeck and Ronquist, 2001) implementing DNA evolutionary models obtainable from PAUP (Swofford, 2002), through Modeltest version 3.06 (Posada and Crandall, 1988). Appropriate models of DNA evolution were chosen for the individual matrices using Modeltest. The model Trn+G was selected for *rbcL*, Tvm+G for *matK* and TIM+I+G for ITS. The dataset was analysed for 6 000 000 generations with a sample frequency of 500. Using tracer log, likelihood scores were plotted to determine the plot of stationarity. All other trees (1200; 10%) prior to this point were discarded as the burn-in phase. The majority consensus tree was interpreted using Fig tree version. 1.4 and the frequency of posterior probability (PP) is indicated on the tree. The following scale was used to evaluate the PPs: weakly supported (below 0.95) and strongly supported (0.95 -1.0). The tree was further edited using Adobe Illustrator CS6.

2.7 Morphological characters

A matrix of seven morphological characters was prepared for 49 species (47 species of *Rhynchosia* and two species of *Eriosema*). These data were gained from inspection of herbarium specimens and published studies on *Rhynchosia* including Bentham and Hooker, (1865); Baker (1923); Ramcharan et al., (1973); Germishuizen, (1998); Fortunato, (2000); Lewis et al., (2005); Germishuizen, (2006) and Moteetee et al., (2012, 2014). Characters were studied under a Olympus SZX10 light microscope at 6.3X magnification. Evolutionary patterns of these characters were examined by constructing them on to the majority-rule consensus tree obtained from BI analyses using mesquite version 2.75 (Madison and Madison, 2011). Morphological characters, character states and data matrix are defined in Table 2.4. Traits values are scaled into binary values and represented using coloured symbols, with different colours proportional to the absolute value of each trait.

Table 2. 1 Voucher specimens used for DNA sequencing and morphological analysis. X indicates accession number still outstanding from GenBank.

Species	Locality	Voucher	UJ DNA bank accession no	GenBank accession no				
				matK	***rbcL***	***trnH-psbA***	***rpL32-trnL***	**ITS**
Ingroup (*Rhynchosia*)								
R. adenodes **Eckl. & Zeyh**	Queenstown	*Van Wyk 1589* (**JRAU**)	UJ10036	X	X	X	X	X
R. albissima **Grand.**	Ntondozi mountains	*Germishuizen 8686* (**PRE**)	UJ10069	X	X	X	X	X
R. angulosa **Schinz**	Barberton district	*Bester 8846* (**PRE**)	UJ10068	X	X	X	X	X
R. argentea **(Thunb.) Harv**.	Knysna	*Zantovska 150* (**PRE**)	UJ10065	X	X		X	X
R. arida **C. H. Stirt.**	Ploegberg near top	*Van Wyk 2834* (**JRAU**)	UJ10033	X	X	X	X	X
R. atropurpurea **Germish.**	Pilgrim's Rest district	*Hankey, Turner, Aubrey 1064* (**PRE**)	UJ10064	X	X	X		X
R. calvescens **Meikle**	Without locality	*Smith 5933* (**PRE**)	UJ10061	X	X	X	X	X
R. capensis **(Burn. F.) Schinz**	Kleinmond, Kogelberg Nature Reserve.	*Boatwright et al. 667* (**NBG**)	X	X		X		X
R. caribaea **(Jacq.) DC.**	Zuurberg National Park	*Acocks 20303,* (**PRE**)	UJ10057		X	X	X	X
R. chrysoscias **Benth**	Outeniqua Pass	*Boatwright 597* (**NBG**)	X	X		X		X
R. cilliata **(Thunb.) Schinz**	Port Elizabeth	*Van Wyk 3307* (**JRAU**)	UJ10029	X	X			X
R. cinnamomea **Schinz**	Kromrivier, Rustenburg	*Mogg & Cunliff 36524* (**JRAU**)	UJ10014	X	X			X
R. clivorum **S. Moore**	Polokwane, Blouberg	*Esterhuysen 21485* (**PRE**)	UJ10084	X	X			X
R. coddii **Germish**.	Magaliesberg, Rustenburg	*Germishuizen 399* (**PRE**)	UJ10077	X	X	X	X	X

Species	Locality	Voucher	UJ DNA bank accession no	GenBank accession no				
				matK	*rbcL*	*trnH-psbA*	*rpL32-trnL*	ITS
R. cooperi **(Harv. Ex Baker f.) Burtt Davy**	Zuurberg National Park	*Van Wyk & Van Wyk 374* (**JRAU**)	UJ10013	X	X			X
R. crassifolia **Benth**	Montana tuine	*Bester 7285* (**PRE**)	UJ10032	X	X			X
R. densiflora **(Roth) DC.**	Groblersdal, Mpumalanga	*Hardy 2223* (**PRE**)	X					X
R. dinteri **Schinz**	Namibia	*Vahmeijer & du Preez 2607* (**PRE**)	UJ10087	X	X	X		X
R. emarginata **Germish.**	Spektakel Pass, Springbok	*Boatwright 572* (**NBG**)	X	X		X		X
R. ferulaefolia **Benth. ex Harv**.	Stanford	*Jaca et al. 462* (**NBG**)	X	X		X		X
R. galpinii **Baker f.**	Klein Namakwaland	*Pole Evans 56121* (**PRE**)	UJ10092	X	X	X		X
R. grandifolia **Steud**.	Mpofu game reserve	*Von Staden 19* (**PRE**)	UJ10091	X				X
R. harmsiana **Schltr. ex Zahlbr**.	Utrecht district	*Devenish 1872* (**PRE**)	UJ10093		X			X
R. harveyi **Eckl. & Zeyh**.	Tarka district Western Cape	*Acocks 17959* (**PRE**)	X			X		X
R. hirsuta **Eckl. & Zeyh.**	Without locality	*Van Wyk 3373* (**JRAU**)	UJ10017		X			X
*R. hirta (***Andrews**) **Meikle & Verdc.**	Without locality	*PRE 29020* (**PRE**)	X					X
R. komatiensis **Harms**	Sabie, Mpumalanga	*Deall 1040, 1434* (**PRE**)	UJ10075	X	X			X
R. leucoscias **Benth. ex Harv.**	Kleinmond, Kogelberg Nature Reserve.	*Boatwright 668* (**NBG**)	X	X		X		X
R. minima **(L.) DC.**	Komatipoort, Mpumalanga	*Winter 5728* (**PRE**)	UJ10062	X	X			X
R. monophylla **Schltr.**	Kruger National Park	*Moteetee et al. 8* (**JRAU**)	X					X

Species	Locality	Voucher	UJ DNA bank accession no	GenBank accession no				
				matK	***rbcL***	***trnH-psbA***	***rpL32-trnL***	**ITS**
R. nervosa **Benth. & Harv**.	Klipriviersberg Nature Reserve	*Hartley, Balkwill & Moodley 919* (**JRAU**)	UJ10018	X	X	X	X	X
R. namaensis **Schinz**	Merve farm Namibia	*Germishuizen 9568* (**PRE**)	X			X		X
R. nitens **Benth.**	Skeerpoort Magaliesberg	*Moteetee et al. 3* (**JRAU**)	X				X	X
R. ovata **J.M. Wood & M.S. Evans**	Wolkberg wilderness area	*Venter 11143* (**PRE**)	X			X	X	X
R. orthodanum **Harv**	Louis Trichardt	*Breijer 19460* (**PRE**)	UJ10105		X			X
R. pentheri **Schltr. ex Zahlbr**	Witwatersrand	*Hartley, Balkwill & Moodley 926* (**JRAU)**	UJ10024	X	X			X
R. pinnata **Harv.**	Bredasdorp	*Acocks 23124* (**PRE)**	UJ10096	X	X		X	X
R. reptabunda **N.E.Br**.	Ermelo Mpumalanga	*Story 691* (**PRE**)	UJ10099	X	X			X
R. resinosa **(A.Rich.) Baker**	Blouberg Pietersburg	*Schutte 576* (**PRE**)	UJ10048	X	X		X	X
R. rogersii **Schinz**	Barberton	*Stalmans 1399* (**PRE)**	UJ10100					X
R. sordida **(E. Mey.) Schinz**	Silwer Glen Nature Reserve	*Schutte 461* (**PRE**)	UJ10046	X	X	X	X	X
R. smithiana **A.Moteetee & Boatwr.**	Kleinmond side of Betty's Bay	*Van Wyk 3155* (**JRAU**)	UJ10031	X	X	X	X	X
R. spectabilis **Schinz**	Buffelskloof Nature Reserve	*Burrows 4690* (**JRAU**)	UJ10026	X	X		X	X
R. sublobata **(Schumach.) Meikle**	Tsumeb, Namibia	*Boatwright 251* (**NBG**)	X	X				X
R. totta **(Thunb.) DC.**	Cradock district Eastern Cape	*PRE 30352* (**PRE)**	UJ10058	X	X		X	X

Species	Locality	Voucher	UJ DNA bank accession no	GenBank accession no				
				matK	***rbcL***	***trnH-psbA***	***rpL32-trnL***	**ITS**
R. vendae **C.H. Stirt**.	Punda Milia 3 km from KNP	*Grobbelaar 02336* (**PRE**)	UJ10060		X			X
R. villosa **(Meisn.) Druce**	Weza Kwa-Zulu Natal	*Stirton 8110* (**PRE**)	UJ10085	X	X			X
Outgroup (*Eriosema*)								
E. gunniae **C.H.Stirt.**	Doornrug farm Witbank	*Nkonki 89* (**PRE**)	UJ10095	X	X	X		X
E. burkei **Harv.**	Sabie Mpumalanga	*Moteetee et al. 9* (**JRAU**)	X	X				XX

X: not yet accessioned

Table 2. 2 Primer sequences and references of gene regions used in the study.

Name		Sequence	Reference
Plastid regions			
matK	472F	CCC RTY CAT CTG GAA ATC TTG GTT C	Chase et al., 2005; Kress et al., 2005; 2007, YU. et al., 2011
	1248R	GCT RTR ATA ATG AGA AAG ATT TCT GC	
	3F KIM	CGT ACA GTA CTT TTG TGT TTA CGA G	CBOL Plant Working Group, 2009
	1R KIM	ACC CAG TCC ATC TCG AAA TCT TGG TTC	
rbcL	1F	ATG TCA CCA CAA ACA GAA AC	Olmstead et al., 1992; Fay et al., 1997
	724R	TCG CAT GTA CCT GCA GTA GC	
	rbcLa-F	ATG TCA CCA CAA ACA GAG ACT AAA GC	CBOL Plant Working Group, 2009
	rbcLa-R	GTA AAA TCA AGT CCA CCY CG	
trnH-psbA	*trnH*	CGC GCA TGG TGG ATT CAC AAT CC	Tate and Simpson, 2003; Shaw et al., 2007
	PsbA	GTT ATG CAT GAA CGT AAT GCT C	Sang et al., 1997; Shaw et al., 2007
rpl32-trnL	*trnL*	CTG CTT CCT AAG AGC AGC G	Shaw et al., 2007
	rpl32	CAG TTC CAA AA A AAC GTA CTT C	
Nuclear region			
ITS	AB101F	ACG AAT TCA TGG TCC GGT GAA GTG TTC G	White et al., 1990; Sun et al., 1994
	AB102R	TAG AAT TCC CCG GTT CGC TCG CCG TTA C	
	ITS2R	GCT GCG TTC TTC ATC GAT GC	
	ITS3F	GCA TCG ATG AAG AAC GCA GC	

Table 2. 3 PCR amplification parameters for the individual regions *matK, rbcL, trnH-psbA, rpl32-trnL* and ITS

Gene	Pre-melt	Cycles	Denaturation	Annealing	Extension	Final extension
Plastid region						
matK	94˚C (3 min)	30	94°C (1 min)	52°C (1 min)	72°C (2 min 30 sec)	72°C (7 min)
rbcL	94°C (3 min)	28	94°C (1 min)	48°C (1 min)	72°C (1 min)	72°C (7 min)
trnH-psbA	80°C (5 min)	30	95°C (1 min)	50°C (1 min)	65°C (4 min)	65°C (5 min)
rpL32-trnL	80°C (5 min)	30	95°C (1 min)	50°C (1 min)	65°C (4 min)	65°C (5 min)
Nuclear region						
ITS	94°C (1 min)	26	94°C (1 min)	48°C (30 sec)	72°C (3 min)	72°C (7 min)

Table 2. 4 List of morphological character states used in the morphological analysis of *Rhynchosia*. Characters and character states are explained at the end of the table.

Taxon	**Morphological characters**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
R. adenodes	0	0	0	0	0	0	0
R. albissima	1	0	1	0	0	0	0
R. angulosa	0	0	0	0	0	0	0
R. argentea	0	1	0	0	0	0	0
R. arida	1	0	0	0	1	0	0
R. atropurpurea	1	1	0	1	1	0	0
R. calvescens	0	1	0	0	1	?	0
R. capensis	0	1	1	0	0	0/2	0/1
R. caribaea	0	1	0	0	0	0	0
R. chrysoscias	0	1	0	0	1	1/ 2	0
R. ciliata	0	0	0	1	0	0	0
R. cinnamomea	0	1	0	1	0	?	0
R. clivorum	0	0	1	1	1	0	0
R. coddii	0	0	0	1	1	0	0
R. cooperi	0	1	0	0	0	0	0
R. crassifolia	0	1	0	1	0	0	0
R. densiflora	0	1	?	0	0	2	0
R. dinteri	?	?	?	1	0	?	0
R. emarginata	1	0	0	0	0	0/1	0
R. ferulifolia	0	0	0	0	1	0	0
R. galpinii	0	0	1	1	0	0	0
R. grandifolia	0	1	0	1	1	0	0
R. harmsiana	0	1	0	1	0	0	0
R. harveyi	0	1	?	1	1	0	0
R. hirsuta	0	1	0	1	0	0	0
R. hirta	0	1	?	1	0	0	0
R. komatiensis	1	0	?	1	0	0	0
R. leucoscias	1	1	0	0	0	1	0
R. minima	0	1	0	0	0	0	0
R. monophylla	0	1	0	1	0	0	0
R. namaensis	?	?	0	1	0	0	0
R. nervosa	0	1	0	1	0	0	0
R. nitens	1	1	1	1	0	0	0
R. orthodanum	1	0	0	1	0	0	0
R. ovata	0	?	0	1	0	0	0
R. pentheri	0	1	0	0	0	0	1
R. pinnata	0	0	0	0	0	0	0
R. reptabunda	0	1	0	0	1	0	0
R. resinosa	0	1	0	0	0	0	0
R. rogersii	1	?	0	1	0	0	0
R. smithiana	0	0	0	0	0	0	0

Taxon		Morphological characters						
		1	**2**	**3**	**4**	**5**	**6**	**7**
R.	*sordida*	0/1	?	?	1	0	0	0
R.	*spectabilis*	1	?	?	0	0	0	0
R.	*sublobata*	0	1	0	0	0	0	0
R.	*totta*	0	0/1	1	1	0	0	0
R.	*vendae*	0	1	?	1	0	0	0
R.	*villosa*	0	1	?	1	0	0	0
E.	*gunniae*	0	0	0	?	0	0	0
E.	*burkei*	0	0	0	0	0	0	0

(1) Habit: herbs or suffrutices = 0, shrubs =1. **(2) Branches:** non-twining = 0, twining = 1. **(3) Leaves:** venation prominent = 0, venation not prominent = 1. **(4) Leaves:** glands present = 0, glands absent = 1. **(5) Leaves:** bulbous based hairs absent = 0, bulbous based hairs present = 1 **(6) Inflorescence:** racemose = 0, umbellate = 1; solitary = 2. **(7) Pods:** pubescent = 0, glabrous = 1

Chapter 3
Results

3.1 Statistics

Details of statistics from the MP analyses of the *matK, rbcL,* ITS and the combined datasets are provided in Table 3.1. ITS has a higher number of variable sites (45%) than *matK* (12.42%) and *rbcL* (3.26%). A high number of parsimony informative sites was also observed in ITS (25.38%), followed by *matK* (3.96%) then *rbcL* (1.81%). Analyses also revealed that ITS evolved much faster than the other two regions with 2.4 changes per variable site and has a lower consistency index (CI). Two additional gene regions (*trnH-psbA* and *rpL32-trnL* intergenic spacers) were investigated but not pursued due to sequencing difficulties. The statistics for analysis of a small sample of species for these two regions are indicated in Table 3.1 and topologies obtained from these analyses are shown in Figures 3.6 and 3.7.

3.2 Molecular markers

3.2.1 matK

Parsimony analysis consisted of 757 characters from 37 samples, in which 30 characters were parsimony informative. Analysis yielded 9,240 most parsimonious trees bootstrap consensus tree is presented in Figure 3.1A. The trees had a length (TL) of 106 steps, CI of 0.92 and RI of 0.91 (Table 3.1). TVM+G was selected as the

best evolutionary method for this gene according to AIC. Monophyly of *Rhynchosia* received a strong support of 97BP. Sections *Chrysoscias* and *Polytropia* are embedded within section *Rhynchosia*.

3.2.2 *rbcL*

Parsimony analysis resulted in a single tree of 19 steps (Figure 3.1B) with a CI of 0.95 and a retention index (RI) of 0.97 (Table 3.1). The analysis consisted of 552 aligned characters from 33 samples, in which 10 characters were parsimony informative. The best model of evolution selected by Akaike Information Criterion (AIC) for this region is TrN+G. Section *Rhynchosia* is paraphyletic, whereas section *Polytropia* which received a support of 64BP is embedded within section *Rhynchosia.*

3.2.3 ITS

Parsimony analysis produced 5,830 equally most parsimonious trees (TL=986; CI= 0.63; RI= 0.64). The matrix included 923 characters, of which 235 were parsimony informative. TIM+I+G was selected by AIC as the most suitable model (Table 3.1). Of all three the regions, ITS proved to be the most informative and the monophyly of *Rhynchosia* is well-supported (98BP). Four major clades were identified (Figure 3.2). Clade one consists mainly of species from section *Rhynchosia*. This clade has two subclades, one of which received a strong support of 100BP, (*R. atropurpurea*, *R. vendae* and *R. hirta*). Its position as sister to the other subclade is poorly supported with 53BP.

Clade two which include sections *Arcyphyllum*, *Polytropia* and *Chrysoscias* received poor support of 66BP. Species from section *Polytropia* form a clade that received a strong support (100BP) while section *Chrysoscias* received moderate support of 75BP.

Clade three and four, which is only made up of species from section *Rhynchosia,* did not receive strong support from both Bayesian and Bootstrap analyses.

Rhynchosia monophylla is morphologically distinct, being the only species in the genus to have unifoliolate leaves in addition to trifoliolate ones (Baker, 1923), and received support of 86BP for its position as sister to the rest of the clades.

3.2.4 Combined molecular analysis

Results of the combined analysis are topologically congruent with those of the ITS analysis. Congruence between ITS and the plastid data sets was evaluated based on bootstrap percentages and Bremer indices. Bremer support indices were negative for nuclear and positive for the plastid data sets; such indices indicate lack of congruence between plastid and nuclear regions, but based on bootstrap consensus there was no hard incongruence between the three data sets. This indicated that the datasets could be combined directly.

The aligned, combined dataset (*matK*, *rbcL* and ITS) included 2,232 characters of which 275 were parsimony informative (12.32%). The MP analysis yielded 2,137 equally most parsimonious trees of 1,131 steps with a CI of 0.65 and a RI of 0.66 (Table 3.1; Figure 3.3). The monophyly of the genus was recovered with high BI support (1.00 PP) and four major clades were retrieved (Figure 3.4).

Clade one comprises species from section *Rhychosia* and *R. hirta* (sect. *Cyanospermum*). The three groups in this clade have received a strong support from the two analyses; the *R. hirta* group (1.00PP; 100BP), the *R. vendae* group (1.00PP; 99BP) and *R. adenodes* group (0.92PP; 86BP).

Clade two (0.99PP; 0BP) consist of species from sections *Polytropia* (1.00PP; 100BP), *Chrysoscias* (0.60PP; 73BP), *R. densiflora* which is forming a clade with *R. arida* and section *Chrysoscias* (0.83PP; 0BP) also species from section *Rhynchosia.* Species from these clades are all located in the Western Cape.

Clade three (0.94PP; 0BP) in which *R. nervosa* is sister to the two groups, comprises of species from section *Rhynchosia.* The sister relation of *R. cinnamomea* to *R. totta* group which received a moderate support (0.84PP; 65BP) received a weak support of (0.73PP; 0BP). The *R. cilliata* group received strong BI (1.00PP) and weak MP support (53BP).

Clade four is divided in to two groups, the *R cooperi* with strong support from both analyses (0.95PP; 87BP) and the *R. spectabilis* also with a strong support (1.00PP; 96BP).

Rhynchosia monophylla is sister to all clades with a strong support of 1.00PP and 89BP.

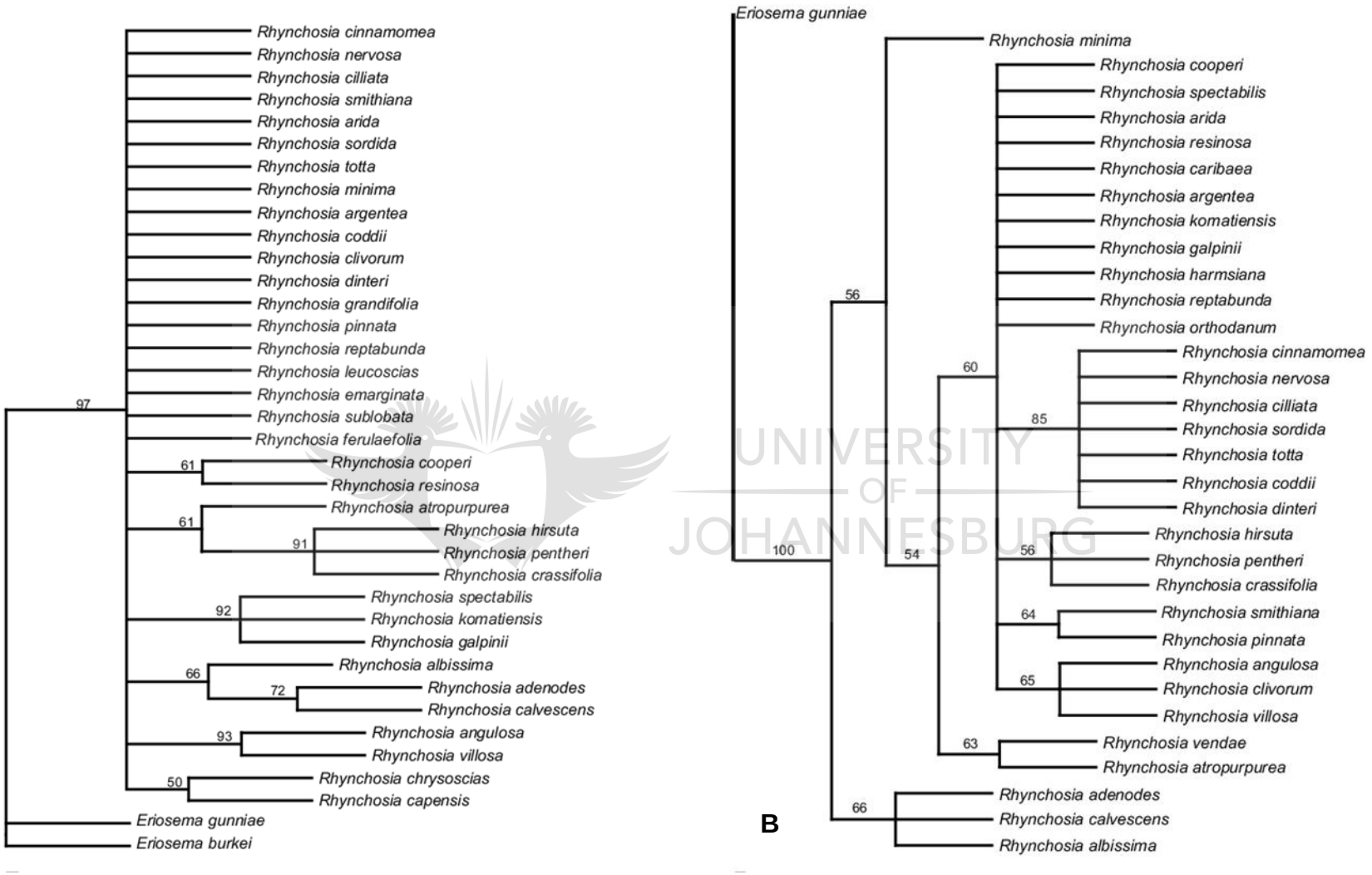


Figure 3. 1 Comparison between the bootstrap consensus trees of A. *matK* and B *rbcL*.

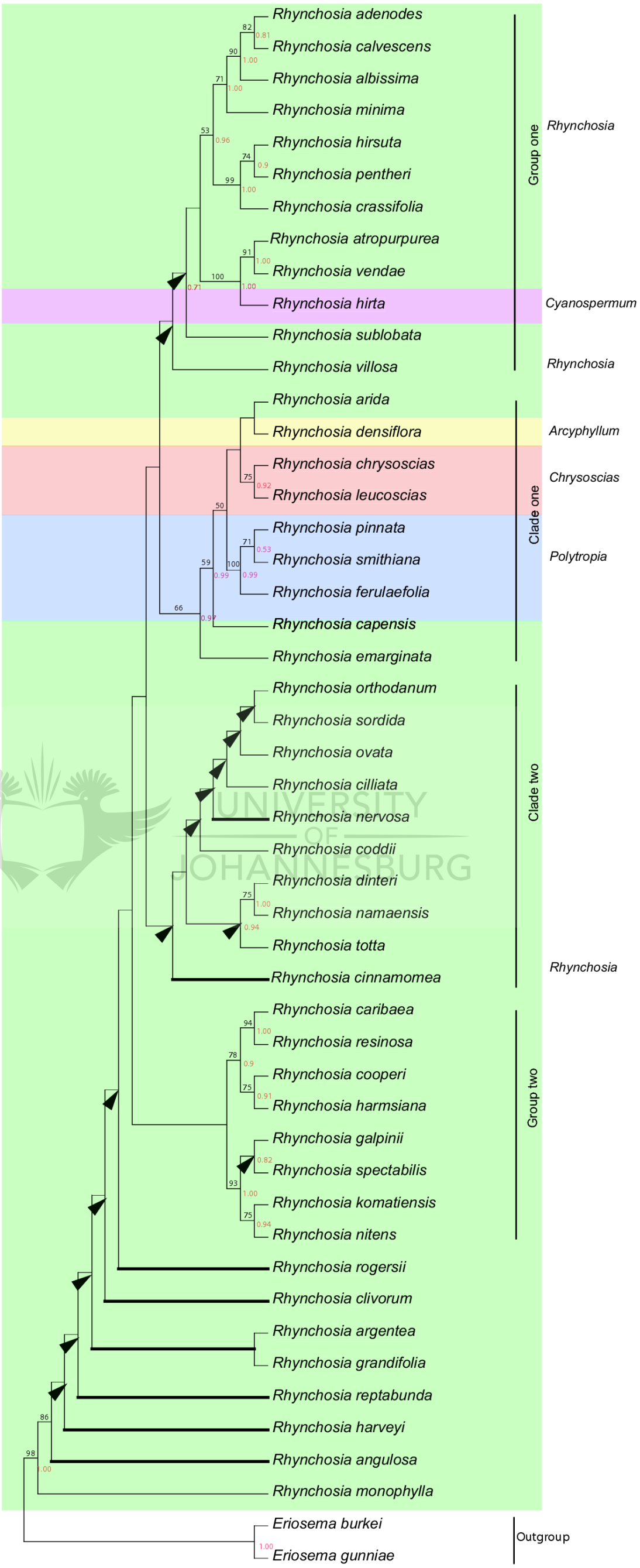


Figure 3. 2 Maximum parsimony analysis of ITS. Parsimonious tree with 50 % bootstrap values indicated above branches and posterior probability values within clades in red. Solid arrows indicate unsupported branches, with bold lines indicating unsupported branches from Bayesian analysis.

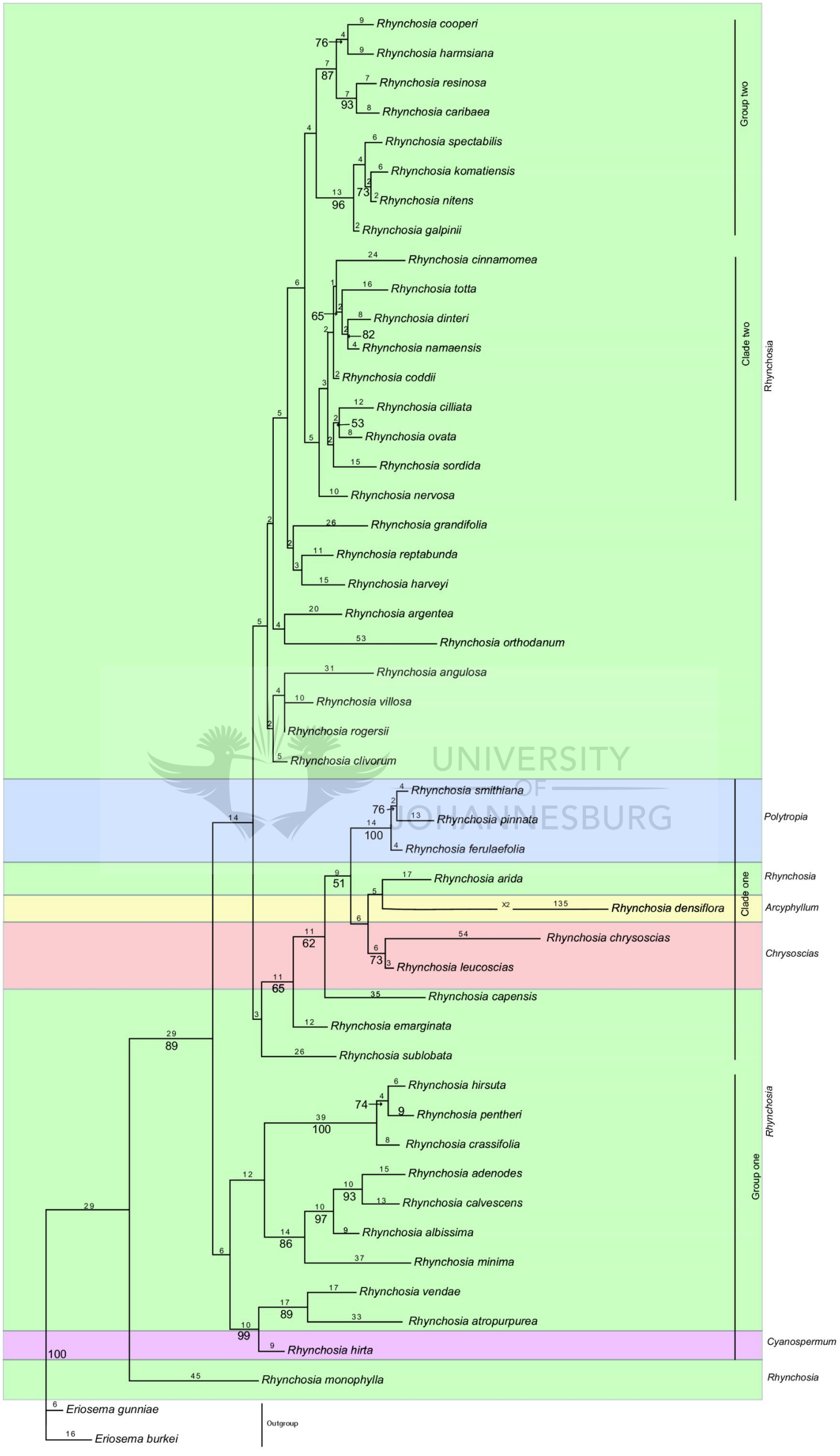


Figure 3. 3 Maximum parsimony tree from analysis of combined data (*matK*, *rbcL* and ITS) indicating branch length above the branches and bootstrap percentages below branches.

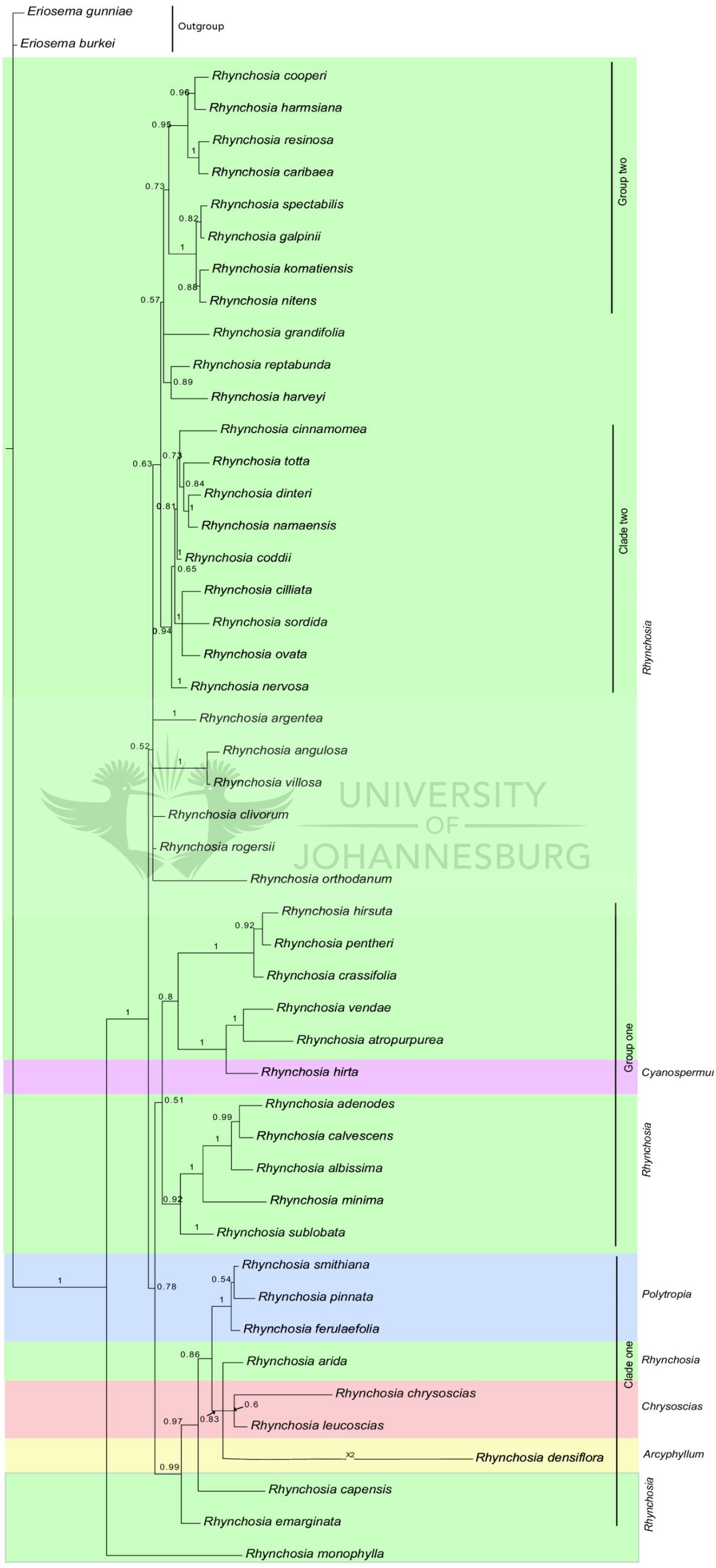


Figure 3. 4 Majority Rule Consensus tree from the Bayesian analysis for the combined data (*matK*, *rbcL* and ITS) with posterior probability values indicated, X2 indicates that the scale has been reduced in half.

Table 3. 1 Statistics and sequence characteristics of the regions *matK*, *rbcL*, ITS and combined analyses.

DNA region	***matK***	***rbcL***	**ITS**	**Combined**
No. of taxa	37	33	49	49
No. of included positions in a matrix	757	552	923	2232
No. of constant characters	663	534	504	1701
No. of parsimony informative characters	30	10	235	275
	3.96%	1.81%	25.38%	12.32%
No. of variable sites	94	18	419	531
	12.42%	3.26%	45.40%	23.79%
No. of trees (Fitch)	9240	1	5830	2137
Tree length (no of steps)	106	19	986	1131
CI	0.92	0.95	0.63	0.65
RI	0.91	0.97	0.64	0.66
Average number of changes per variable site (number of steps/number of variable sites)	1.1	1.1	2.4	2.1
Model selected by Akaike Information Criterion	TVM+G	TrN+G	TIM+I+G	

3.2.5 *trnH-psbA* intergenic spacer

Parsimony analysis of 23 taxa resulted in 40 trees. The bootstrap consensus tree is presented in Figure 3.5. Analysis showed that this region is second most informative in this study with 57 parsimony informative characters and 189 variable sites (Table 3.2). This argument is based on comparison between all regions used in the study (Table 3.1 and 3.2). Sections *Chrysoscias* and *Polytropia* are embedded within section *Rhynchosia*, supporting the results obtained from combined analysis, but *Rhynchosia chrysoscias* and *R. leucoscias* are paraphyletic, but there is no sectional correlation from analysis of *trnH-psbA* as it is observed in topology obtained from combined analyses of *matK, rbcL* and ITS.

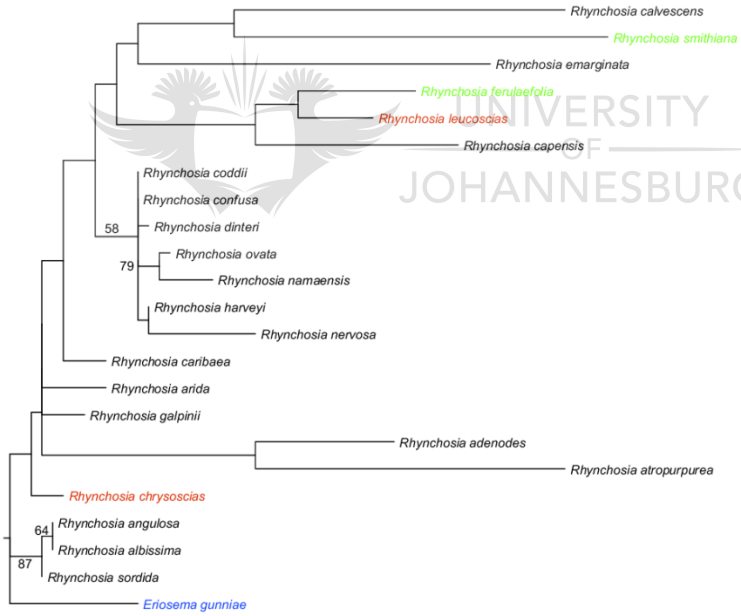


Figure 3. 5 Bootstrap consensus tree from the analysis of *trnH-pbsA*. Different sections indicated with colours, red: section *Chrysoscias*, green: section *Polytropia*, black: section *Rhynchosia* and blue: outgroup.

3.2.6 *rpl32-trnL* intergenic spacer

A parsimony analysis yielded 52 trees of 109 steps (Figure 3.6), with 29 characters that are parsimony informative, and 82 variable sites (Table 3.2). For this analysis, *Eriosema* was not sequenced successfully and the monotypic genus *Bolusafra* was used as outgroup considering that *Bolusafra* is also closely related to *Rhynchosia* (Moteetee and Van Wyk, 2006). Results from this analysis indicate that species from section *Polytropia* is very closely related to *Bolusafra*. *Rhynchosia adenodes* and *R. calvescens* maintained their sister relationship as in combined analysis (*matK, rbcL* and ITS) with a strong support of 86BP.

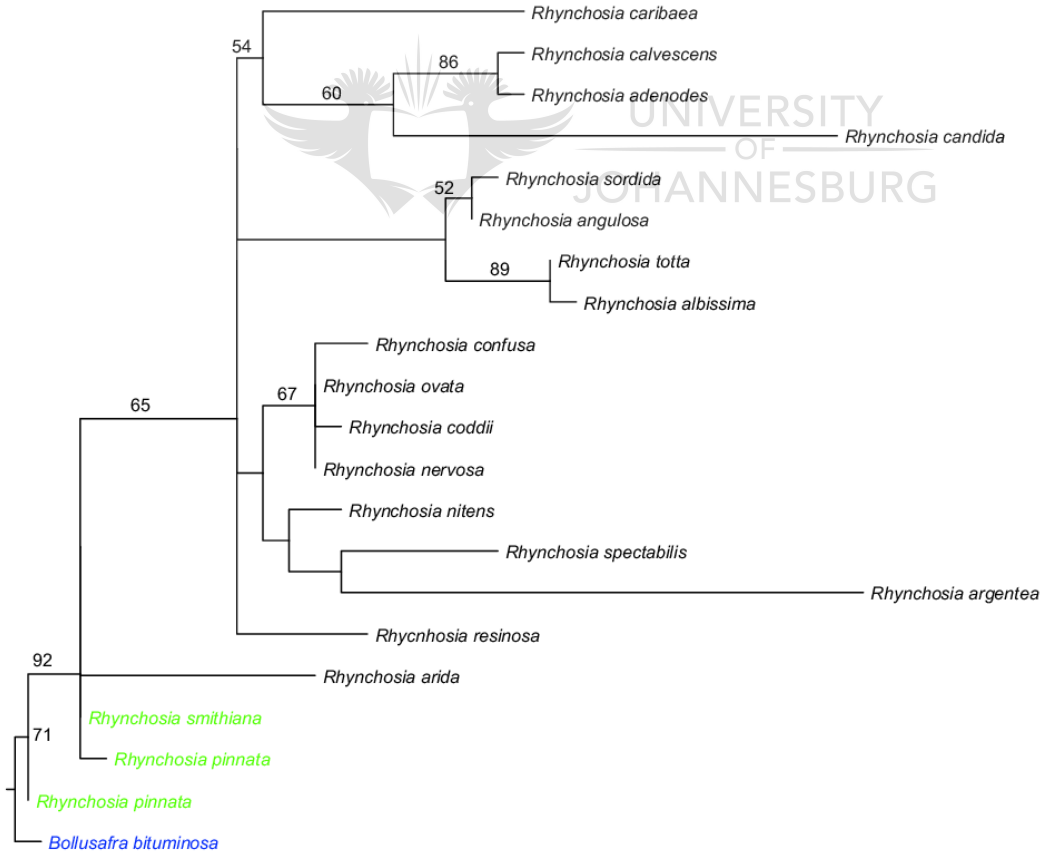


Figure 3. 6 Bootstrap consensus tree from the analysis of *rpl32-trnL* with bootstrap percentage above 50% shown. Different sections indicated with colours, green: section *Polytropia*, black: section *Rhynchosia* and blue: outgroup.

Table 3. 2 MP statistics and sequence characteristics of the regions *trnH-psbA* and *rpl32-trnL*

DNA region	***trnH-psbA***	***rpl32-trnL***
No. of taxa	23	21
No. of included positions in a matrix	460	464
No. of constant characters	271	382
No. of parsimony informative characters	57	29
	12.39%	6.25%
No. of variable sites	189	82
	41%	17.67%
No. of trees (Fitch)	40	52
Tree length (no of steps)	261	109
CI	0.84	0.89
RI	0.66	0.85
Average number of changes per variable site (number of steps/number of variable sites)	1.38	1.32

3.3 Morphological characters

Morphological features were not a major part of the scope of this study, but those were used in order to determine if there could be a correlation between the molecular results and circumscriptions of sections within the genus. Characters such as trichomes (glands and bulbous based hairs), growth habit and branches have shown potential to be informative for defining groups within the genus (Figure 3.7 A, B, C; 3.8 A,B and 3.9 A), but this will have to be pursued in future studies of the genus *Rhynchosia*.

3.3.1 Growth habit

Rhynchosia species are predominantly prostrate to trailing herbs or suffrutices (Baker, 1923). Selected lineages have become woody and are shrubby, e.g. the *R. nitens* group (except for *R. galpini*). Most of the species have a twining habit and grow among surrounding vegetation for support. In some herbaceous lineages, such as the members of section *Polytropia*, taxa display a non-twining habit which is a useful character to distinguish these groups. Species with a shrubby habit tend to not be twining (Figure 3.7 A)

3.3.2 Leaves

Species of *Rhynchosia* have leaflets that are trifoliolate, rarely unifoliolate (*R. monophylla*) and bipinnate to pinnate with 5-9 leaflets (section *Polytropia*), while the shape of leaflets ranges from linear (*R. pauciflora)* through narrowly lanceolate (section *Polytropia*) to ovate-trullate (*Arcyphyllum*, *Cyanospermum* and *Rhynchosia*). The leaf apex is usually acute to somewhat apiculate, and the margins vary from

revolute to non-revolute. Leaflet shape and size is of taxonomic value although some related species have similar shapes (*R. cinnamomea, R. totta*; *R. vendae* and *R. atropurpurea*). Prominent leaf venation is found in most of the species in the genus. The shrubby *R. nitens* group lack prominent venation which might be a character to support the close relationship between these species. Selected species in other parts of the tree also lack prominent venation (Figure 3.7 B,C).

3.3.3 Trichomes

Occurrence, distribution and cellular structure of trichomes have been used by taxonomists largely for identification purposes. For example in research such as: morphology, distribution, and the role of trichomes (Oghiakhe et al., 1992), trichomes as physical barriers for cowpea pod borer *Maruca testulalis* (Veeranna and Hussain, 1997). Other taxonomists have used trichome occurrence to investigate how legumes use trichomes for defence mechanism against insects and other pests (Strauss and Zangerl, 2002; Edwards and Singh, 2006). In terms of type, density and distribution trichomes are very variable in *Rhynchosia* species. Hair distribution on the leaf blade varies greatly, even within species and they are most commonly present on the midrib and frequently on the lower surface of the leaf, calyx and pods (Figure 3.10).

Glandular trichomes have been reported to occur in all other genera in Cajaninae except in *Adenodolichos* (Lackey, 1981, Moteetee and Van Wyk, 2006) as well as in other tribes such as Dalbergieae, Galageae, Sophoreae and Genisteae (Moteetee and Van Wyk, 2006). Some species of *Rhynchosia* are very heavily covered with glandular trichomes (e.g*. R. capensis, R. chrysoscias* and *R. pinnata)*, while in others they seem to be lacking altogether (*R. totta, R. namaensis and R.*

galpinii). In all species glands and hairs occur either together or alternating on the lower surface of the leaf. The pattern of gland presence/absence appears to be quite complex and further investigation and larger sampling of the genus is likely to unravel this pattern and its potential use in the taxonomy of *Rhynchosia* (Figure 3.8 A,B).

3.3.4 Inflorescence

Most species have racemose inflorescences, although section *Chrysoscias* have developed umbellate inflorescences (Figure 3.9 A)

3.3.5 Pod vestiture

Pod pubescence in *Rhynchosia* is very variable ranging from tomentose, pillose, hispid, villous, scarbid to hirsute, with some species being glabrous to glabrescent (Figure 3.9 B).

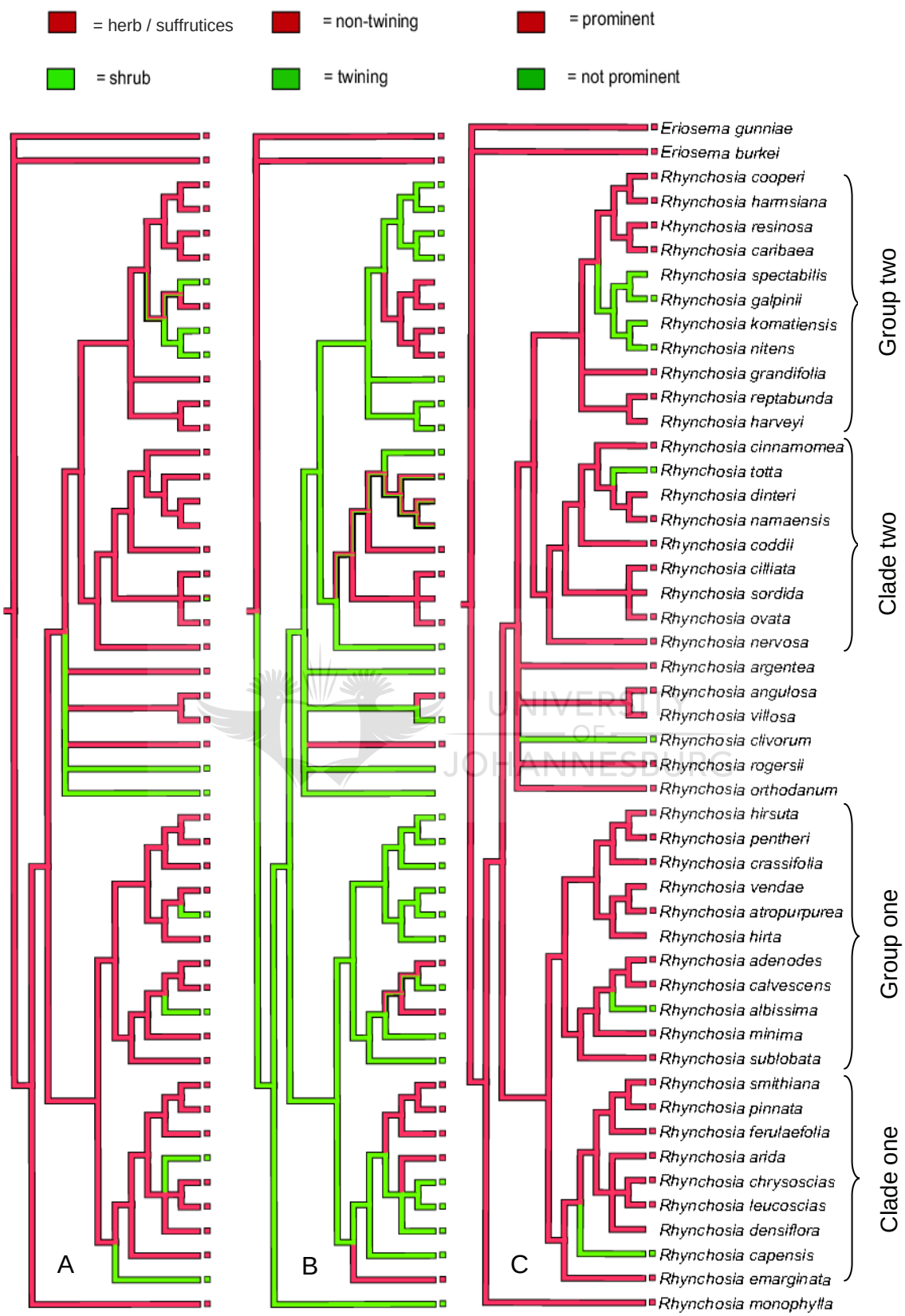


Figure 3. 7 Parsimony-based reconstruction of characters: **A** (habit), **B** (branches), **C** (leaf venation) on the majority rule Bayesian tree of the combined molecular data.

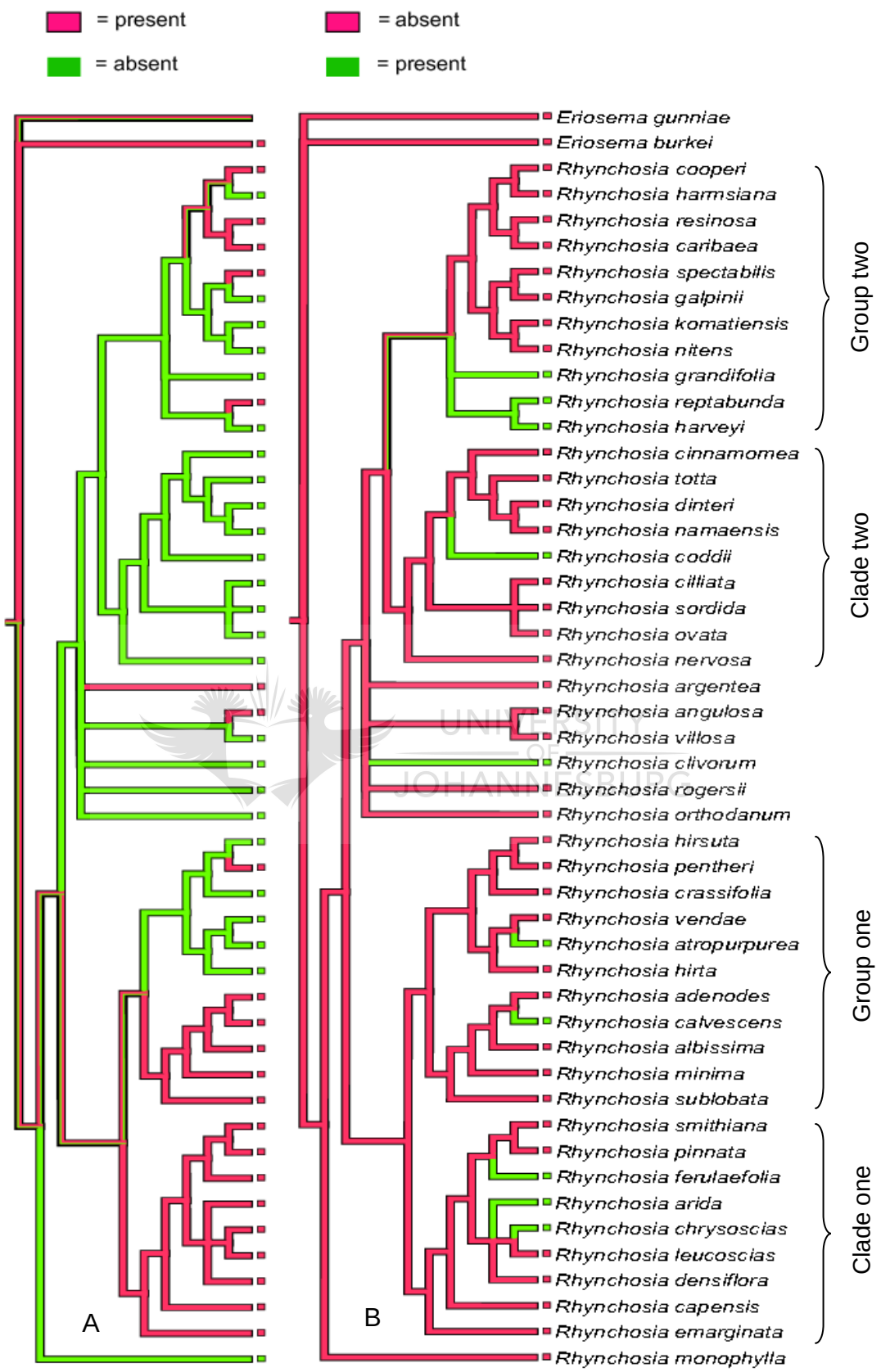


Figure3. 8 Parsimony-based reconstruction of characters: **A** (glands), **B** (bulbous based hairs) on the majority rule Bayesian tree of the combined molecular data.

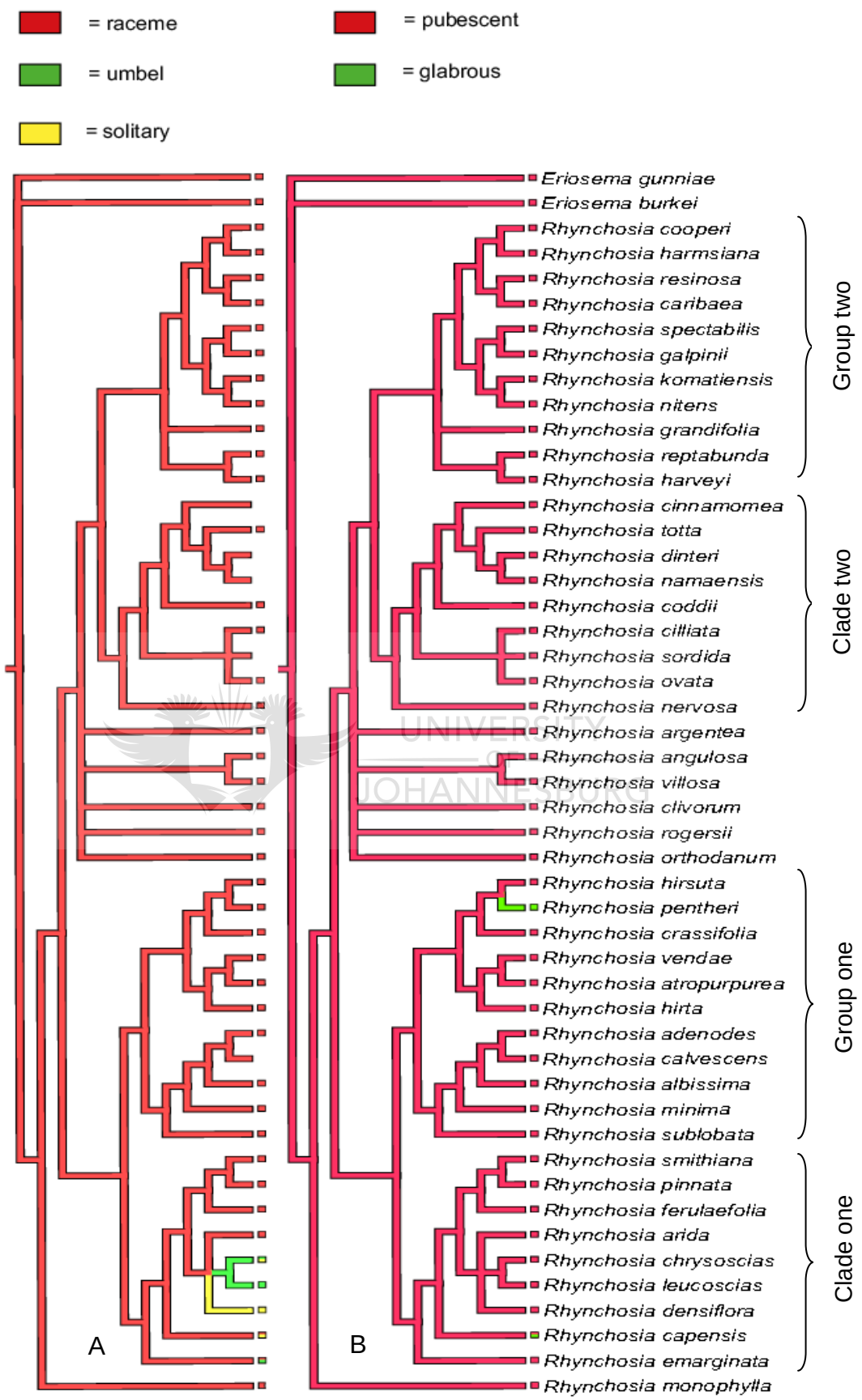
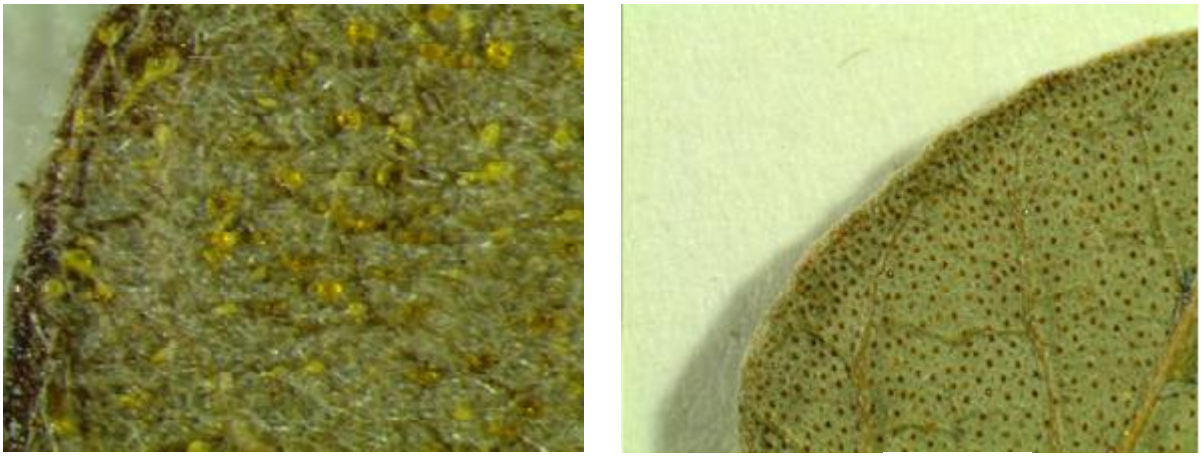


Figure3. 9 Parsimony-based reconstruction of characters: **A** (Inflorescence) and **B** (Pods) on the majority rule Bayesian tree of the combined molecular data.



A B

C

Figure3. 10 Leaves surface showing trichome distribution in *Rhynchosia*: **A**. *R. chrysoscias* with both glands and bulbous based hairs (BBH), **B**: *R*. *capensis* showing only glands and **C***: R. totta* without glands and BBH.

Chapter 4
Discussions

4.1 Discussion

4.1.1 Phylogenetic relationship within the genus

The nuclear gene ITS, has showed to be more informative and has been reported to provide insightful phylogenetic analysis in numerous plant families, including legumes (Hu et al., 2002). In this study it has proved to be the most informative and valuable gene region for the inference of phylogenetic relationships with the highest percentage of parsimoniously informative characters (25.38%) and variable sites at 45.40%.

However, the core of this discussion is based on the results obtained from combined data set (Figures 3.3 and 3.4). The genus is monophyletic, with strong support from BI (1.0PP). Four major clades were identified. Section *Rhynchosia* is not monophyletic, and though sections *Polytropia* (1.0PP; 100BP) and *Chrysoscias* (0.6PP; 73BP) are embedded within section *Rhynchosia,* they appear to be monophyletic. This study represents a first phylogeny for *Rhynchosia* and this will in turn inform further taxonomic studies of the genus. In future studies additional samples, especially from Australia and the Americas, will be added as well as additional gene regions to obtain a global picture of relationships within the genus.

4.1.2 Group one

This group received low support (0.51PP; 0BP). The relationship between *R. hirsuta, R. pentheri* and *R. crassifolia* received high support (1.0PP; 100BP). Other branches also received strong support, including *R. adenodes, R. calvescens, R. minima* and *R. albissima* (1.0PP; 86BP) as well as *R. vendae*, *R. atropurpurea* and *R. hirta* (1.0PP; 99BP). The clade consists mostly of species from section *Rhynchosia* and a member of section *Cyanospermum* (*R. hirta*). Species in this clade are mostly herbs, with twining branches. Taxa from the *adenodes* clade are all glandular, while the *R. vendae* clade is non-glandular, with the exception of *R. pentheri*.

4.1.3 Group two

This group only received weak support from the BI (0.57PP), but from both analysis MP and BI the group contains two distinct groupings. Relationship between the species in these two groupings is mostly well-supported. Group two A consists of *R. cooperi, R. harmsiana, R. resinosa* and *R. caribaea* (0.95PP; 87BP) and group two B of *R. spectabilis, R. galpinii, R. komatiensis* and *R. nitens* (1.00PP; 96BP). Species in group two A are twining herbs, with prominent leave venation, while group two B consists of non-twining shrubs and veins on leaves that are not prominent. Although *R. resinosa* occurs in South Africa, it was not included in Baker's (1923) revision; it was assigned to section *Cyanospermum* by Baker (1923). Interestingly, this topology supports the exclusion of this species by Moteetee et al. (2012) from section *Cyanospermum* as it lacks the typical dark blue seeds that characterise this section.

4.1.4 Clade one

Clade one received no bootstrap support, but strong support from BI (0.9PP). It consists of members from sections *Rhynchosia*, *Polytropia* and *Chrysoscias*. Members from section *Polytropia* (Figures 3.3 and 3.4) received strong support (1.0PP; 100BP) for its monophyletic. These taxa are all non-twining herbs, with high density of glands on the leaves, bulbous based hairs (except in *R. ferulaefolia*) and dense racemes (Moteetee et al., 2014). Another internal branch which comprises *R. arida*, species from section *Chrysoscias* and one species from section *Arcyphyllum* (Figures 3.3 and 3.4), received 0.83 PP and. 0BP Species in this branch are twining herbs, except *R. arida* which is non-twining shrub, and all have glands while *R. arida* and *R. chrysoscias* have bulbous based hairs as well. *Rhynchosia densiflora*, which belongs to section *Arcyphyllum*, has solitary flowers while members of section *Chrysoscias* have umbellate inflorescences. In both topologies the branch leading to *R. densiflora* has a high branch length value. Using the same herbarium specimen, this species was re-sequenced with the same result and it is unclear why this species has so many changes in its sequences. This could be an artefact of under-sampling in this study and adding more samples into the analysis could be important. *Rhynchosia hirta* (group 1) and *R. densiflora* (clade 1) are not closely related as suggested by Baker (1923).

4.1.5 Clade two

This clade received strong support from BI (0.94PP) but no support from the bootstrap analysis. It includes both twining and non-twining herbs without glands. Some of these species are widely distributed and highly variable (e.g. *R. totta*) and warrant further investigation. The relationship of some species is unresolved (e.g. *R.*

argentea, *R. clivorum*, *R. rogersii*) at the base of this clade, additional gene regions may improve the resolution and shed light on the placement of these taxa.

4.2 Depiction from selected morphological characters

Rhynchosia species, particularly those in the type section, are morphologically variable, even within species and are often difficult to distinguish from each other. For analysis, seven morphological characters were incorporated, including characters that have been identified as most distinct for the tribe Phaseoleae (Lackey, 1981), presented in Table 2.4. These were reconstructed on to a majority rule consensus tree obtained from BI. Most characters that were identified are plesiomorphic and the derived characters occur across four clades giving no distinct resolution.

The herbaceous nature is prominent in most species with only few species exhibiting the paraphyletic shrub state across three different clades (Figure 3.7 A). A similar pattern is seen in Figure 3.7 C and E. Prominent leaf veins (Figure 3.7 C) is a common character and shared by majority of species except *R. spectabilis* group, *R.totta*, *R. clivorum*, *R. albissima* and *R. capensis*. Whereas the presence of bulbous based hairs is not common but spread across the clades (Figure 3.7 E).

The non-twining character (Figure 3.7 B) is derived and shared across all clades. *Rhynchosia spectabilis* group in clade four, *R. ovata group*, species from section *Polytropia* as well as *R. emarginata which is sister to species in clade two* exhibits a reversal ancestral character twining character is prominent within the genus.

The presence of glands on leaves is a paraphyletic plesiomorphic character, shared by species in clade two, *R. adenodes* group in clade one, *R. cooperi* group in clade four with exclusion of *R. harmsiana*. This reversal character is also expressed in *R. pentheri, R. angulosa, R. argentea, R. reptabunda* and *R. spectabilis.* (Figure 3.7 D).

Racemose inflorescence character is clearly a plesiomorphic character. Figure 3.7 F indicates a selection of characters for *R. densiflora,* which has solitary flowers, while *R. chrysoscias* and *R. leucoscias* have umbellate inflorescences.

Pods are dominantly pubescent with an exception to *R. pentheri* which exhibit an autapomorphic character for glabrous pods, as shown in Figure 3.7 G.

Given the lack of distinct characters observed in morphological analyses it would not be possible to divide the genus into sections. At the moment I would reserve the splitting or assigning the clades to different sections until at least a large number of species of the genus are represented in the study and perhaps have two or more gene regions added. More robust morphological characters that would be useful in splitting the genus into section would also be paramount for analysis.

Chapter 5
Conclusion

5.1 Conclusion

This study aimed to investigate phylogenetic relationships in the genus *Rhynchosia* based on several gene regions as well as morphological characters that might be useful in circumscribing sections in the genus.

5.1.1 Monophyly of the genus

The monophyly of the genus was supported, based on the individual and combined analyses of plastid (*matK* and *rbcL*) and nuclear (ITS) sequences, though more species could be sampled to further confirm this and give a more comprehensive topology. The support was not very strong for groups one and two with a PP value of 0.51 and 0.57 respectively; the other two clades received a strong Bayesian support of 0.99PP for clade two and 0.94PP for clade three. Adding more samples and gene regions might improve an understanding and definition of the genus.

5.1.2 Sectional classification

It is clear that the sectional classification of Baker (1923) needs to be re-investigated and smaller sections defined for the groups in the large section *Rhynchosia*. The classification system as recognized by Baker, (1923) is not congruent with the

pattern displayed in the molecular phylogeny. The results from this study supported the monophyly of the section *Polytropia* and retrieved moderate support for section *Chrysoscias*. Species such as *R. microscias* (belonging to the latter section), for example, was not successfully sequenced and their inclusion could have aided in better resolved relationships of the section. The study indicates that section *Rhynchosia* is not monophyletic and also found that there is a close relationship between sections *Polytropia* and *Chrysoscias* a correlation which is supported from both molecular and morphological analyses, and has indicated that *R*. *arida, R*. *capensis* and *R. densiflora* are fairly closely related (clade 2). Species in sections *Archyphyllum* and C*yanospermum* were documented morphologically to be closely related (Baker 1923) but have grouped in separate clades. Samples from across the world need to be included in these analyses so that their relationships to the South African species can be assessed and taken into consideration when devising a new sectional classification for the genus.

5.1.3 Utility of ITS

The nuclear gene ITS has proven to be the most informative marker in the inference of the phylogeny of *Rhynchosia*, plastid gene *matK* and *rbcL* were rather conserved, which proves the conserved nature of *matK* in subtribe Cajaninae (Lavin et al., 2005). The species grouping obtained from combined analyses were very similar to results from ITS analyses, but resulted in improved support for the branches from both MP and BI analyses.

5.2 Future research

5.2.1 Wider taxon sampling

Previous study by Kassa, (2012) indicated that not only is *Rhynchosia* paraphyletic it is also polyphyletic as shown by the position of *R. auera* and *R orthobotrya* on the topology (Figure 2.1), implying the need to include wider taxon sampling to gain detailed biogeographic and phylogenetic information for the genus. For this study the two species were not included since the focus was only on South African species. Addition of more species, which occupy wide geographical ranges, will shed more light on the phylogeny of the genus.

5.2.2 Molecular markers

Even though ITS was informative in constructing the phylogeny of *Rhynchosia*, there is a need to add more markers to help improve the resolution. Table 3.2 shows that second to ITS; the regions *trnH-psbA* and *rpl32-trnL* are more informative than other plastid regions used, as also evident from Shaw et al. (2007). The *trnH-psbA* region was found to have higher percentage variability as well as higher potentially informative character (PIC). The sequencing problems such as nonspecific binding of nucleotides and multiple bands encountered with this region need to be addressed in future.

Chapter 6
References

A

AÏNOUCHE, A. & BAYER, R. J. 1999. Phylogenetic relationships in *Lupinus* (Fabaceae, Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. *American Journal of Botany* 86: 590-607.

ÁLVAREZ, I. & WENDEL, J. F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417-434.

B

BAUDET, J. C. 1978. Prodrome d'une classification generique des Papilionaceae-Phaseoleae. *Bulletin du Jardin Botanique National du Belgique* 48: 183-220.

BAKER, E. G. 1923. Revision of South African species of *Rhynchosia*. *Bothalia* 1: 113–138.

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(2): 247-277.

BARTHET, M. M. & HILU, K. W. 2007. Expression of *matK* functional and evolutionary implications. *American Journal of Botany* 94(8): 1402-1412.

BENTHAM, G. & HOOKER, J. D. 1865. Leguminosae. In: Bentham, G. & Hooker, J. D. (editors) *Genera Plantarum, Vol 1(2)*. Lovell Reeve, London. pp 437-562.

BOATWRIGHT, J. S., WINK, M., VAN WYK, B.-E. 2011. The generic concept of *Lotononis* (Crotalariae, Fabaceae): reinstatement of the genera *Euchlora, Leobordea* and *Listia* and the new genus *Ezoloba. Taxon* 60: 161-177.

BOATWRIGHT, J. S. & MOTEETEE, A. N. 2014. *Rhynchosia bolusii* (Phaseoleae, Fabaceae), a new species from the Western Cape Province of South Africa.

Phytotaxa 161(2): 169-172

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42(4) 795-803.

C

CBOL PLANT WORKING GROUP. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences USA* 106: 12794-12797.

CHASE M. W., SOLTIS D. E., OLMSTEAD R. G., MORGAN D., LES D. H., MISHLER B. D., DUVALL M. R., PRICE R. A., HILLS H. G., QIU Y-L., KRON K. A., RETTIG J. H., CONTI E., PALMER J. D., MANHART J. R., SYTSMA K. J., MICHAELS H. J., KRESS W. J., KAROL K. G., CLARK W. D., HEDREN M., GAUT B. S., JANSEN R. K., KIM K-J., WIMPEE C. F., SMITH J. F., FURNIER G. R., STRAUSS S. H., XIANG Q-Y., PLUNKETT G. M., SOLTIS P. S., SWENSEN S., WILLIAMS S. E., GADEK P. A., QUINN C. J., EGUIARTE L. E., GOLENBERG E., LEARN JR. G. H., GRAHAM S. W., BARRETT S. C. H., DAYANANDAN S. AND ALBERT V. A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80(3): 528-580.

CHASE, M. W., SALAMIN, N., WILKINSON, M., DUNWELL, J. M., KESANAKURTHI, R. P., HAIDAR, N. & SAVOLAINEN, V. 2005. Land plants

and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1889-1895.

CURTIS, S. E. & CLEGG, M. T. 1984. Molecular evolution of chloroplast DNA sequences. *Molecular Biology Evolution* 1(4): 291-301.

D

DANIELL, H., LEE, S.-B., GREVICH, J., SASKI, C., QUESADA-VARGAS, T., GUDA, C. B., TOMKINS, J. & JANSEN, R. K. 2006. Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theoretical and Applied Genetics* 112: 1503-1518.

DAVIS, C. C., FRITSCH, P. W., LI, J. & DONOGHUE, M. J. 2002. Phylogeny and biogeography of *Cercis* (Fabaceae): evidence from nuclear ribosomal ITS and chloroplast *ndhF* sequences. *Systematic Botany* 27: 289-302.

DE CANDOLLE, A. P. 1825. Leguminosae. *Prodromus Systematis Naturalis Regni Vegetabilis* 2: 118-120.

DELGADO-SALINAS, A., BIBLER, R. & LAVIN, M. 2006. Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Systematic Botany* 31: 779-791.

DOYLE, J. J. 1992. Gene trees and species trees: Molecular systematic as one-character taxonomy. *Systematic* Botany 17: 144-163.

DOYLE, J. J. & DOYLE, J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

DOYLE, J. J. & DOYLE, J. L. 1993. Chloroplast DNA phylogeny of the papilionoid legume tribe Phaseoleae. *Systematic Botany* 18(2): 309-327.

DOYLE, J. J., DOYLE, J. L., BALLENGER, J. A., DICKSON, E. E., KAJITA, T. & OHASHI, H. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541-554.

DOYLE, J. J., CHAPPILL J. A., BAILEY C D. & KAJITA. T. K. 2000. Towards a comprehensive phylogeny of legumes: evidence from *rbcL* sequences and non-molecular data. In: Herendeen, P. S. & Bruneau, A. (editors). *Advances in Legume Systematics,* Royal Botanic Gardens, Kew*,* pp. 1-20.

E

ECKLON, C. F. & ZEYHER, C. L. P. 1836. Enumaration plantarum africae australis extratropicae. Perthes and Besser, Hamburg, pp. 145-288.

EDWARDS, O. & SINGH, K. B. 2006. Resistance to insect pests: what do legumes have to offer? *Euphytica* 147(1-2): 273-285.

EMS, S. C., MORDEN, C. W., DIXON, C. K., WOLFE, K. H., DEPAMPHILIS, C. W. & PALMER, J. D. 1995. Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. *Plant Molecular Biology* 29: 721-733.

F

FAY, M, F., SWENSEN, S. M. & CHASE, M. W. 1997. Taxonomic affinities of *Medusagyne oppostifolia* (Medusagynaceae). *Kew Bulletin.* 52: 111-120

FAY, M. F., RICHARDSON, J. E., COWAN, R. S. & STRANC, P. A. 1998. The Application of Genetic Fingerprinting Techniques to plant Conservation Problems. *Royal Botanical Gardens, Kew.*

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

Fortunato, R. H. 2000. Systematic relationships in *Rhynchosia* (Cajaninae Phaseoleae–Papilionoideae–Fabaceae) from the neotropics. In: Herendeen, P. S, Bruneau, A. (editors). *Advances in Legume Systematics 9*. The Royal Botanic Gardens, Kew, pp. 339-354.

G

GERMISHUIZEN, G. 1998. Fabaceae: the new species of *Rhynchosia* from South Africa. *Bothalia* 28(1): 58-60

GERMISHUIZEN, G. 2006. *Rhynchosia*. In: Germishuizen, G., Meyer, N.L., Steenkamp, Y. & Keith, M. (editors). *A Checklist of South African Plants.* South African Biodiversity Network Report no. 41, SABONET, Pretoria, pp. 488-490.

GRAHAM, P. H. & VANCE, C. P. 2003. Legumes: importance and constraints to greater use. *Plant Physiology* 131: 872-877.

GREAR, J. W. 1978. A revision of the New World species of *Rhynchosia* (Leguminosae-Faboideae). *Memoirs of the New York Botanical Garden* 31: 1-168.

H

HAMBY, R. K. & ZIMMER, E. A. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis, P. S. Soltis, D. E. Doyle & Doyle, J. J. (editors). *Molecular Systematics of Plants,* Chapman & Hall, New York, pp. 50-91.

HAO, D. C., CHEN, S. L. & XIAO, P. G. 2010. Sequence characteristics and divergent evolution of the chloroplast *psbA-trnH* noncoding region in angiosperm. *Journal of Applied Genetics* 51: 259-273

HARVEY, W. H. 1862. Leguminosae. In: *Harvey, W. H. & Sonder, O. W. (editors). Flora Capensis,Hodges, Smith & Co., Dublin* 2: 1-285.

HASTON, E. M., LEWIS, G. P. & HAWKINS, J. A. 2005. A phylogenetic reappraisal of the *Peltophorum* group (Caesalpinieae: Leguminosae) based on the chloroplast *trnL-F, rbcL,* and *rps*16 sequence data. *American Journal of Botany 92*: 1359-1371.

HU, J.-M., LAVIN, M., WOJCIECHOWSKI, M. F. & SANDERSON, M. J. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trnK/matK* sequences and its implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418-430.

HU, J.-M., LAVIN, M., WOJCIECHOWSKI, M. F. & SANDERSON, M. J. 2002. Phylogenetic analysis of nuclear ribosomal ITS/5.85 sequences in the tribe Millettieae (Fabaceae): *Poecilanthe-Cyclllobium*, the core Millettieae, and the *Callerya* group. *Systematic Botany 27*: 722-733.

HUELSENBECK, J. P. & RONQUIST, F. 2001. Mr bayes: Bayesian inference of phylogenetic trees. *Bioinfomatics* 17: 754-755.

K

KAJITA, T., OHASHI, H., TATEISHI, Y., BAILEY, C. D. & DOYLE, J. J. 2001. *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. *Systematic Botany* 26: 515-536.

KÄSS, E. & WINK, M. 1995. Molecular phylogeny of the Papilionoideae (family Leguminosae): *rbcL* gene sequences versus chemical taxonomy. *Botanica Acta* 108: 149-162.

KÄSS, E. & WINK, M. 1996. Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcL*-sequences. *Biochemical Systematics and Ecology* 24: 365-378.

KÄSS, E. & WINK, M. 1997. Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the *rbcL* gene and ITS 1 + 2 regions on nrDNA. *Plant Systematics and Evolution* 208: 139-167.

KASSA, M. T., PENMETSA, R. V., CARRASQUILLA-GARCIA, N., SARMA, B. K., DATTA, S., UPADHYAYA, H. D., VARSHNEY, R. K., VON WETTBERG, E. J. B. & DOUGLAS, R. C. 2012. Genetic patterns of domestication in pigeonpea (*Cajanus cajan* (L.) Millsp.) and wild *Cajanus* relatives. *PloS One* 7: e39563.

L

LACKEY, J. A. 1977. A revised classification of the tribe Phaseoleae (Leguminosae: Papilionoideae), and its relation to canavanine distribution. *Botanical Journal of the Linnean Society* 74: 163-178.

LACKEY, J. A. 1981. Phaseoleae. In: Polhill, R. M. & Raven P. H. (editors). *Advances in Legume Systematics 1*, Royal Botanic Gardens, Kew, pp. 301-327.

LAVIN, M., HERENDEEN, P. S. & WOJCIECHOWSKI, M. F. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* 54: 575-594.

LEWIS, G. P., SCHRIRE, B., MACKINDER, B. & LOCK, M. 2005. Legumes of the World, Royal Botanic Gardens*, Kew,* pp. 393-432.

LEGUME PHYLOGENY WORKING GROUP 2013. Legume phylogeny and classification in the 21st century: progress, prospects and lessons for other species-rich clades. *Taxon* 62: 217-248.

LOUREIRO, J. De. 1790. *Flora Cochinchinensis* 2. Ulyssipone, Lisbon, pp. 425.

M

MADDISON W. P. & MADDISON D.R. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75, http://mesquiteproject.org.

MCKEY, D. 1994. Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. In: Sprent, J. I. & McKey, D. (editors). *Advances in Legume Systematics 5, the Nitrogen Factor*. Royal Botanic Gardens, Kew, UK, pp. 211-228.

MEYER, E. H. F. 1836. Commentariorium de plantes Africae australis. Leipzig*, Leopoldum Voss 2*: 65-75.

MORT, M. E., ARCHIBALD, J. K., RANDLE, C. P., LEVSEN, N. D., O'LEARY, T. R., TOPALOV, K., WIEGAND, C. M. & CRAWFORD, D. J. 2007. Inferring phylogeny at low taxonomic levels: utility of rapidly evolving cpDNA and nuclear ITS loci. *American Journal of Botany* 94: 173-183.

MOTEETEE, A. & VAN WYK, B.-E. 2006. A revision of the genus *Bolusafra* (tribe Phaseoleae, Fabaceae). *South African Journal of Botany* 72(4): 604-608.

MOTEETEE, A. N., BOATWRIGHT, J. S. & JACA, T. P. 2012. A review of *Rhynchosia* section *Cyanospermum* (Phaseoleae, Fabaceae) in South Africa. *South African Journal of Botany* 81: 124-127.

MOTEETEE, A. N., BOATWRIGHT, J. S. 7 JACA, T. P. 2014. A review of *Rhynchosia* section *Polytropia* (Phaseoleae, Fabaceae) and a new species from Western Cape province, South Africa. *Systematic Botany* 39: 1127-1131.

MÜLLER, K., BORSCH, T. & HILU, K. 2006. Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: contrasting *matK*, *trnT-F*, and *rbcL* in basal angiosperms. *Molecular Phylogenetics and Evolution* 41: 99-117.

MURTHY, K. S. R. & RAO, K. R. S. 2008. Chemical composition and nutritional evaluation of *Paracalyx scariosus* (Roxb.) Ali a wild relative of *Cajanus* from southern peninsular India. *Tropical and Subtropical Agroecosystems* 10: 121-127.

N

NEUHAUS, H. & LINK, G. 1987. The chloroplast *tRNALys*$^{(UUU)}$ gene from mustard (*Sinapsis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. *Current Genetics* 11: 251-257.

O

OGHIAKHE, S., JACKAI, L. E. N., MAKANJUOLA, W. A. & HODGSON, C. J. 1992. Morphology, distribution, and the role of trichomes in cowpea (*Vigna unguiculata*) resistance to the legume pod borer, *Maruca testulalis* (Lepidoptera: Pyralidae). *Bulletin of Entomological Research* 82(4): 449-505.

OLMSTEAD, R. G., MICHAELS, H. J., SCOTT, K. M. & PALMER, J. D. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. *Annals of the Missouri Botanical Garden* 79(2): 249-265.

OLMSTEAD, R. G. & PALMER, J. D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* 81(9): 1205-1224.

OSALOO, S. K., MAASSOUMI, A. & MURAKAMI, N. 2005. Molecular systematics of the Old World *Astragalus* (Fabaceae) as inferred from nrDNA ITS sequence data. *Brittonia* 57: 367-381.

P

POLHILL, R. M. 1981. Papilionoideae. In: Polhill, R. M. & Raven, P. H. (editors). *Advances in Legume Systematics,* Royal Botanic Gardens*, Kew,* pp. 191-208.

POSADA, D. & CRANDALL, K. A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14(9): 817-818

R

RAMCHARAN , E. K., MCANDREWS, J. H & GREAR, J. W. 1973. Pollen morphology of *Rhynchosia* and *Eriosema* (Fabaceae). *Canadian Journal of Botany* 51: 2395-2399.

RICHARDSON, J. E., FAY, M. F., CRONK, Q. C. B., BOWMAN, D. & CHASE, M. W. 2000. A phylogenetic analysis of *Rhamnaceae* using *rbcL* and *trnL-F* plastid DNA sequences. *American Journal of Botany* 87: 1309-1324.

S

SANG, T., CRAWFORD, D. J. & STUESSY, T. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120-1136.

SAVOLAINEN, V., CUÉNOUD, P., SPICHIGER, R., MARTINEZ, M. D., CRÈVECOEUR, M. & MANEN, J.-F. 1995. The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Systematics and Evolution* 197: 87-98.

SAVOLAINEN, V., FAY, M. F., ALBACH, D. C., BACKLUND, A., VAN DER BANK, M., CAMERON, K. M., JOHNSON, S. A., LLEDÓ, M. D., PINTAUD, J.-C., POWELL, M., SHEAHAN, M. C., SOLTIS, D. E., SOLTIS, P. S., WESTON, P., WHITTEN, W. M., WURDACK, K. J. & CHASE, M. W. 2000. Phylogeny of eudicots: a nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bulletin* 55: 257-309.

SHAW, J., LICKEY, E. B., BECK, J. T., FARMER, S. B., LIU, W. S., MILLER, J., SIRIPUN, K. C., WINDER, C. T., SCHILLING, E. E. & SMALL, R. L. 2005. The

tortoise and the hare II: relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142-166.

SHAW, J., LICKEY, E. B., SCHILLING, E. E. & SMALL, R. L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* 94: 275-288.

SORENSON, M. & FRANZOSA, E. 2007. TreeRot, Version 3. Boston University, Boston, MA.

SPRENT, J. I. 2001. *Nodulation in legumes*. Richmond, U.K.:Royal Botanic Gardens, Kew.

STAPPEN, J. V., LOPEZ, U. G., DAVILA, P. & VOLCKAERT, G. 2002. Molecular evidence for the hybrid origin of a new endemic species of *Stylosanthes* Sw. (Fabaceae) from the Mexican Yucatán Peninsula. *Botanical Journal of the Linnean Society* 140: 1-13.

STEANE, D. A. 2005. Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus* (Myrtaceae). *DNA Research* 12: 215-220.

STEELE, K. P. & WOJCIECHOWSKI, M. F. 2003. Phylogenetic analyses of tribes Trifolieae and Vicieae, based on sequences of the plastid gene, *matK* (Papilionoideae: Leguminosae). In: Klitgaard, B. B. & Bruneau, A. (editors). *Advances in legume systematic, part 10, higher level systematics,* Royal Botanic Garden, Kew, pp. 355-370.

STRAUSS, S. Y. & ZANGERL, A. R. 2002. Plant-insect interactions in terrestrial ecosystems. In: Herrera, C. M. & Pellmyr, O. (editors). *Plant-animal*

interactions: an evolutionary approach. Blackwell Publishing Ltd, Oxford, pp. 77-106.

SUGITA, M., SHINOZAKI, K. & SUGIURA, M. 1985. Tobacco chloroplast tRNALys (UUU) gene contains a 2.5-kilobase-pair intron: an open reading frame and a conserved boundary sequence in the intron. *Proceedings of the National Academy of Sciences, USA* 82: 3557-3561.

SULAIMAN, S. F., CULHAM, A. & HARBORNE, J. B. 2003. Molecular phylogeny of Fabaceae based on *rbcL* sequence data: with special emphasis on the tribe Mimoseae (Mimosoideae). *Asia pacific journal of molecular biology and biotechnology* 11(1): 9-35.

SUN, Y., SKINNER, D., LIANG, G. & HULBERT, S. 1994. Phylogenetic analysis of Sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics* 89: 26-32.

SWOFFORD, D. L. 2002. *PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10*. Massachusetts: Sinaues Associates.

T

TABITA, F. R. 1999. Photosynthesis. *Research* 60: 1-28.

TATE, J.A.,& SIMPSON, B. B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* 28: 723–737.

TUCKER, S. C. 2002. Floral ontogeny in *Sophoreae* (Leguminosae: Papilionoideae). III. Radial symmetry and random petal aestivation in *Cadia purpurea*. *American Journal of Botany* 89: 748-757.

V

VARELA, E. S., LIMA, J. P., GALDINO, A. S., PINTO, L. D. S., BEZERRA, W. M., NUNES, E. P., ALVES, M. A. & GRANGEIRO, T. B. 2004. Relationships in

subtribe Diocleinae (Leguminosae; Papilionoideae) inferred from internal transcribed spacer sequences from nuclear ribosomal DNA. *Phytochemistry* 65: 59-69.

VEERANNA, R. & HUSSAIN, M. A. 1997. Trichomes as physical barriers for cowpea pod borer *Maruca testulalis* (Geyer) (Lepidoptera: Pyralidae). *Insect Environment 3(1)*

VERDCOURT, B. 1971. Studies in the Leguminosae Papilionoideae for the 'Flora of Tropical East Africa. *Kew Bulletin* 96.

W

WAKASUGI, T., TSUDZUKI, T. & SUGIURA, M. 2001. The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. *Photosynthesis Research* 70: 107-118.

WENDEL, J. F., SCHNABEL, A. & SEELANAN, T. 1995. An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. *Molecular Phylogenetic Evolution* 4: 298-313.

WHITE, T. J., BRUNS, T., LEE, S. & TAYLOR, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. (editors). *PCR Protocols: a Guide to Methods and Applications.* Academic Press, San Diego.

WOJCIECHOWSKI, M. F. 2005. *Astragalus* (Fabaceae): a molecular phylogenetic perspective. *Brittonia* 57: 382-396.

WOJCIECHOWSKI, M. F., SANDERSON, M. J. & HU, J.-M. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. *Systematic Botany* 24: 409-437.

WOJCIECHOWSKI, M. F., LAVIN, M. & SANDERSON, M. J. 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. *American Journal of Botany* 91: 1846-1862.

WOODS, M. & KEY, J. 2009. The genus *Rhynchosia* (Fabaceae) in Alabama. *Phytologia* 91: 3-17.

Y

YU, J., XUE, J. H. & ZHOU, S. L. 2011. New universal *matK* primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution* 49 (3): 176-181.