Royal Botanic Garden Edinburgh

THE UNIVERSITY OF EDINBURGH

Morphology and Systematics of two African genera of Sapotaceae - *Synsepalum* (A.DC.) Daniell and *Englerophytum* K. Krause

Thesis submitted in partial fulfilment for the MSc in Biodiversity and Taxonomy of Plants

Daniel Borg

August 2013

DECLARATION

I hereby declare that the work contained in this thesis is my own, unless otherwise acknowledged and cited. This thesis has not in whole or part been previously presented for any degree.

Daniel Borg
Edinburgh
22nd August 2013

The image on the title page was taken by Tony Rodd and was obtained online from Flickr at http://www.flickr.com/photos/tony_rodd/4488830439/

ABSTRACT

Englerophytum and *Synsepalum* are two closely-related genera of trees and shrubs found distributed in the African tropics. Previous molecular studies have shown that these genera collectively form a monophyletic clade within the sub-family Chrysophylloideae (Sapotaceae). However, little is known about the inter-relationships of the taxa within the *Englerophytum-Synsepalum* clade. The current generic circumscriptions of *Englerophytum* and *Synsepalum* are based solely on non-molecular evidence.

This study has used sequences from the nuclear ITS region and the chloroplast trnH-psbA region in order to help resolve the phylogeny within the *Englerophytum-Synsepalum* clade. The ITS dataset proved to be more informative than the chloroplast dataset and produced a better resolved tree. No hard incongruence was found between the datasets and thus they were combined into one matrix.

Results from Parsimony and Bayesian methods have shown that the *Englerophytum-Synsepalum* clade consists of six well supported lineages, two composed solely of taxa from the genus *Englerophytum* and four composed of taxa from the genus *Synsepalum*.

A morphological study of the taxa within the *Englerophytum-Synsepalum* clade, indicated that each lineage can be distinguished by suites of vegetative and floral characters. Leaf venation patterns, calyx fusion, style length and staminodal structure were amongst the most useful characters for distinguishing clades.

Some of the sub-clades within the *Englerophytum-Synsepalum* clade were also found to closely fit descriptions of former genera that have now been placed in synonymy with *Englerophytum* and *Synsepalum*.

ACKNOWLEDGEMENTS

Firstly I would like to express my deepest gratitude to my supervisors, James Richardson, David Harris and Barbara MacKinder whose help and support was constant throughout the thesis. They were always there to discuss problems, provide their professional opinion and provide encouragement, sometimes even on weekends. I can never thank them enough.

There are several people at RBGE who have helped me throughout my thesis. I am forever indebted to Laura Forrest and Michelle Hollingsworth who provided the laboratory training and who were always there for help and support whenever something did not work as planned in the laboratory. I would also like to thank Peter Wilkie who has organised weekly Sapotaceae meetings throughout this thesis to help us understand the morphology of the Sapotaceae. A big thank you also goes to Mark Hughes, who has patiently guided me through the post-laboratory phylogenetic work and was always available for help.

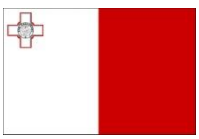
I would also like to thank all the Herbarium staff for help in using BG-Base, especially Erzsebet Gyongy and Suzanne Cubey. Thanks goes as well to the Library staff, especially Graham Hardy who knows the library shelves by heart and was always available to help.

The thesis has also benefited from the help of (i) Syndney Ndolo Ekiba, who was has guided me on the use of BRAHMS (ii) William Hawthorne from FHO who was so kind to provide some of his specimens for inclusion in this project (iii) Xander and Jan from WAG who provided several specimens of *Englerophytum* (iv) Bhaskar Adhikari for help in mapping morphological characters on the tree (v) David Middleton for providing guidance on the EJB referencing style (vi) Louis Ronse De Craene for being an excellent course supervisor (vii) Lakmini Kumarage for her suggestions and help in PCR and sequencing (viii) Carmen Puglisi for help in Paup script

I would also like to thank my family for their unwavering support throughout this thesis.

My final thanks goes to all of my course mates, not only for their academic help but also for their friendship throughout the course. I would particularly like to thank (i) Peter Moonlight for the numerous helpful discussions we had on phylogenetics (ii) James Clugston for his help and suggestions on the presentation of the thesis (iii) Atsuko Shibata for her wonderful company in the lab (iv) Isuru Udayanga Kariyawasam for being supportive in times of need (v) Lorena Villaneuva for her help in mapping (vi) Kimberly Fackler and Bryony Smart for providing good laughs and excellent company (vii) Steve Scott and Fran Culverhouse for their company in the herbarium and (viii) Stephen Mifsud for providing a taste of home – Malta.

The research work disclosed in this publication is partially funded by the Strategic Educational Pathways Scholarship Scheme (Malta).
The scholarship is part-financed by the European Union – European Social Fund.



Operational Programme II – Cohesion Policy 2007-2013
Empowering People for More Jobs and a Better Quality of Life
Scholarship part-financed by the European Union
European Social Fund (ESF)
Co-financing rate: 85% EU Funds; 15% National Funds

Investing in your future

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Appendix 1

Appendix 2

1. INTRODUCTION

1.1 Introduction to *Englerophtyum* and *Synsepalum*

Englerophytum and *Synsepalum* are two closely-related genera of woody trees and shrubs belonging to the family Sapotaceae in the Order Ericales (APG III, 2009). They are distributed across tropical Africa (Pennington, 1991), where they usually inhabit primary forest or savannah, often close to riverine areas (Govaerts *et al.*, 2001).

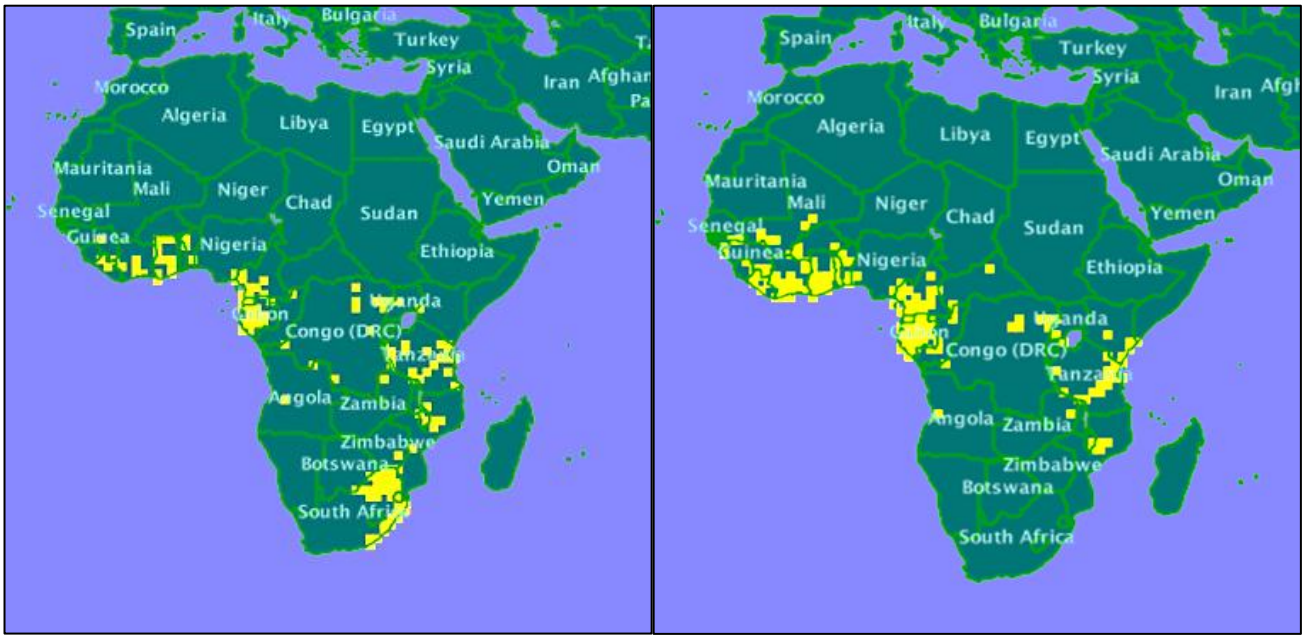


Figure 1.1 – Distribution of *Englerophytum* (left) and *Synsepalum* (right) across tropical Africa

Members of *Englerophytum* and *Synsepalum* are used by local people for several purposes. Species such as *S. afzelii*, *S. dulcificum* and *E. magalismontanum* possess edible fruits, many of which are sweet and suitable for making syrup, jelly and wine (Watt & Breyer-Brandwijk, 1962; Govaerts *et al.*, 2001). Some other species, such as *S. msolo* and *S. brevipes*, possess hard timber and are used in construction of tools, canoes and houses (Bolza & Keating, 1972; Louppe *et al.*, 2008). The species *S. stipulatum* is often harvested to produce charcoal (Keay, 1989). Other uses of plants in these genera include medicines against coughs and colds (*S. cerasiferum*, *S. brevipes*), sweetening agents (*S. dulcificum*) (Ayensu, 1972), landscaping and decoration (*E. natalense*) (Burkill, 2000).

According to the most recent online checklist of the Sapotaceae (WSCP, 2013), the genus *Englerophytum* has 14 accepted species whilst *Synsepalum* has 35. However, in the absence of a recent formal revision, these values are provisional and the exact number of species within the genera and their circumscription still remains unclear.

1.2 Taxonomic position of *Synsepalum* and *Englerophytum* within the Sapotaceae

The position of *Englerophytum* and *Synsepalum* in the Sapotaceae has changed several times over the years. These changes came about as a result of the continuously evolving classification scheme within the Sapotaceae. There has been numerous classification schemes proposed for the Sapotaceae (De Candolle, 1844; Hartog, 1878; Baillon, 1890; Pierre, 1890; Dubard, 1912; Lam, 1939; Aubréville, 1964; Baehni, 1965), however the most recent generic treatment of Pennington (1991), is amongst the most important since it is based on the greatest amount of evidence.

In his classification scheme, Pennington split the Sapotaceae into five tribes - Mimusopeae, Sideroxyleae, Omphalocarpeae, Isonandreae and Chrysophylleae. He placed *Synsepalum* and *Englerophytum* in the Chrysophylleae , a placement which, although based exclusively on non-molecular evidence, is very similar to the placement given to these genera in latest molecular-based classification scheme of the Sapotaceae (Swenson & Anderberg, 2005). The characters of the tribe Chrysophylleae, are provided in Table 1.1

Table 1.1 – Morphological characters of the tribe Chrysophylleae (Pennington, 1991)

1. Calyx in a single whorl of 4-5 sepals	5. Corolla lobes undivided
2. Sepals imbricate or quincuncial	6. Stamens exserted or included
3. Corolla lobes and stamens usually same number as sepals	7. Small staminodes, present or absent
4. Corolla tubular, cyathiform or rotate	8. Seed scar usually adaxial, rarely basiventral

In the current classification (Swenson & Anderberg, 2005), Pennington's tribal classification has been replaced by a sub-familial classification and consists of three sub-families: (i) **Sarcospermatoideae** (ii) **Sapotoideae** and (iii) **Chrysophylloideae.** *Synsepalum* and

Englerophytum have been shown to belong to the sub-family Chrysophylloideae, which is very similar in circumscription to the Tribe Chrysophylleae with minor alterations. Most of the Pennington's characters for the tribe Chrysophylleae (Table 1.1) still apply to the subfamily Chrysophylloideae. Figure 1.2 shows a simplified version of the taxonomic position of the *Englerophtum* and *Synsepalum* within the Sapotaceae. They form a strongly supported monophyletic clade within the Chrysophylloideae and have a sister relationship with the core Chrysophylloid clade which contains the majority of the members of the Chrysophylloideae.

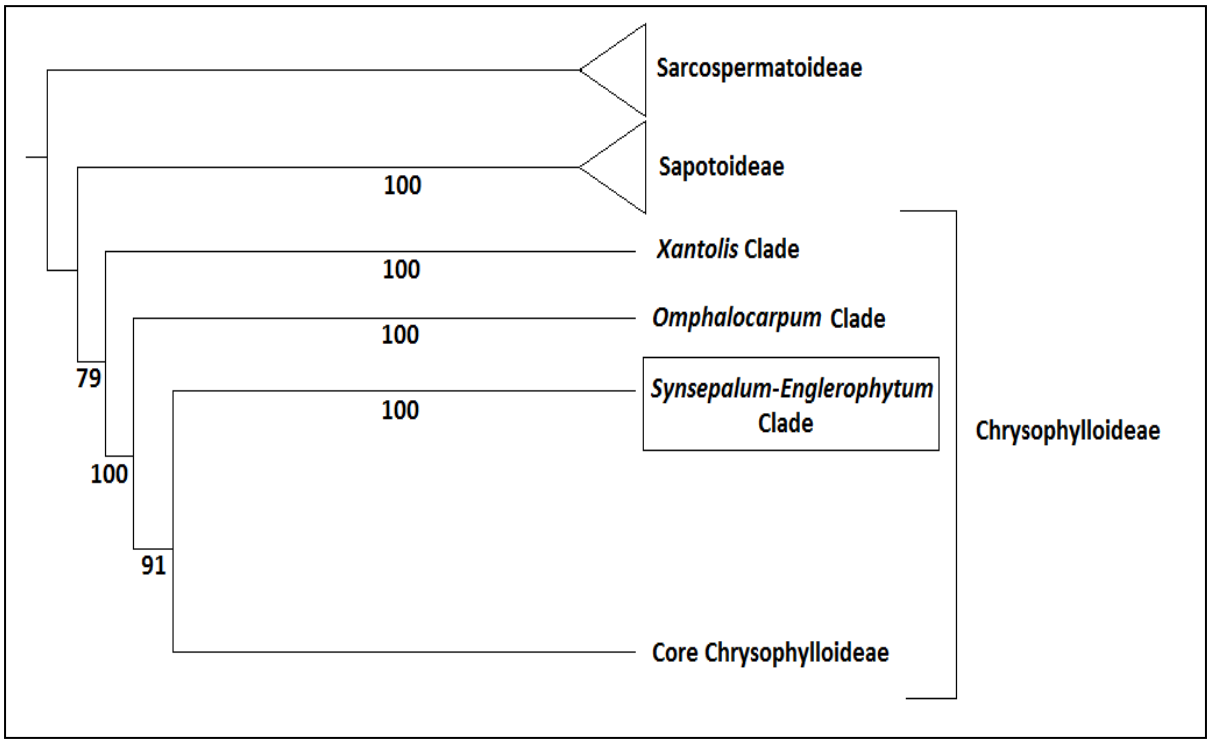


Figure 1.2 - A simplified phylogenetic tree of the Sapotaceae adapted from Swenson & Anderberg (2005) and Swenson et. al. (2008). Values below the branches represent bootstrap support values. *Englerophytum* and *Synsepalum* together form a strongly monophyletic clade within the Chrysophylloideae.

1.3 Generic concepts in the Sapotaceae

Although the close relationship between *Synsepalum* and *Englerophytum* is undisputed, there has always been a difficulty in distinguishing between these genera both morphologically and molecularly. In general, establishing clear morphological boundaries between genera in the Sapotaceae has always posed difficulties to taxonomists.

The first reason for this is that the Sapotaceae possess relatively few useable morphological characters at the generic level when compared to other families (Pennington, 1991). Recent

studies (Swenson *et al.*, 2008) have also shown that many morphological characters in the Sapotaceae show a tendency for homoplasy and hence might lead to wrong assumptions about relationships.

The second reason for the difficulty in defining genera concerns the nature of variation in the Sapotaceae. The Sapotaceae exhibit what is known as polythetic variation (Sneath, 1962).This means that there is considerable overlap of characters between genera and very few genera (only *Argania* and *Magodendron*) can be defined using unique characters. Most genera must therefore be defined by suites of characters known as differential characters (White, 1962) which although exhibiting some degree of overlap between genera, can still be used to separate the majority of the members of one genus from another.

The failure to understand the nature of such variation by previous taxonomists has resulted in classification schemes in which genera are separated based on discontinuities in single character states. A clear example of such a classification scheme is that established by Aubréville, 1964. His classification scheme possessed several narrowly defined genera based on single discontinuities. For instance, in his system, members of *Synsepalum* were scattered in at least six different genera (*Vincentella, Afrosersalisia, Pachystela, Pseudopachystela, Synsepalum, Tulestea*). Pennington (1991) states, that such a taxonomic approach towards Sapotaceae is inconsistent because it leads to fluid classification schemes that can change depending on which character is given more weight.

As a consequence of the numerous previous attempts at classifying the Sapotaceae, extensive synonymy exists within the family, particularly in the genera, whose circumscription and names were repeatedly changed in order to conform to different systems. *Synsepalum* and *Englerophytum*, as presently described, have at least 13 generic synonyms each (Pennington, 1991; Govaerts *et al.*, 2001)

1.4 Taxonomic history of *Synsepalum*

The genus name *Synsepalum* was first validly published by Daniell in 1852 after following the suggestion of De Candolle (1844) who in his Prodromus had tentatively put *Synsepalum* as a section of the genus *Sideroxylon* whilst suggesting that perhaps it deserves to be a

distinct genus in its own right. Since its establishment, the circumscription of *Synsepalum* changed several times.

At the time of its first valid publication, *Synsepalum* was monotypic containing only its type species *Synsepalum dulcificum*. It was initially distinguished from the genus *Sideroxylon* (to which it was formerly thought to belong to) by its partially fused 5-fid calyx and its obtuse anthers (De Candolle, 1844; Daniell, 1852).

Engler (1904) widened the circumscription of *Synsepalum* in order to include two other species *S. stipulatum*, and *S. ulugurense*. Engler considered several characters in his circumscription of the genus, as can be seen from his extensive descriptions; however the ultimate distinguishing character of *Synsepalum* was still the fusion of the sepals.

In 1911-1912, *Synsepalum* and several other genera (incl. *Bakerisideroyxlon, Pachystela, Sersalisia*) were grouped together by Dubard into a single genus - *Bakeriella*. This was done in an attempt to unite genera which were poorly defined, often just by a single character, into a more consistent genus which could be better defined by a robust suite of characters.

However Dubard's classification was mostly ignored, consequently leading to a large number of narrowly circumscribed genera distinguished from each other by single character states or by a series of uncorrelated characters. Whenever a new species failed to conform to a narrow generic circumscription, a new genus was erected. For instance, A. Chevalier (1943) erected a new genus of two species called *Afrosersalisia* because the degree of fusion of its sepals varied from that of *Synsepalum.*

Meanwhile, new species were still being added to *Synsepalum* including (i) *S. subcordatum* and *S. longicuneatum* by De Wildeman (1914) (ii) *S. glycodora* by Wernham (1917) and (iii) *S. attenuatum* and *S. fleureyanum* by Chevalier (1943)

In the 1960s, there was the need to revise the classification of the genera of the Sapotaceae and Aubréville (1964) and Baehni (1965) provided two widely divergent views on the family which complicated matters further. In Aubréville's account 122 genera were recognized and in Baehni's account 63 genera were recognized. Aubréville kept *Synsepalum* as a distinct

genus distinguishing it from other genera on the basis of its fused cup-shaped calyx and staminode length. Baehni (1965) sunk *Synsepalum* into other genera such as *Pouteria* and *Richardiella.*

One of the most recent and thorough generic treatments of Sapotaceae was made by Pennington (1991). In his treatment of *Synsepalum*, which included analysis of morphological, anatomical, cytological and palynological characters, Pennington rejected Baehni's classification stating that it is mostly artificial and recognizes *Synsepalum* as a genus in its own right. However, Pennington refrained from using the narrow genus circumscription used by Aubréville and widened his circumscription of *Synsepalum* to include the following genera - *Vincentella, Afrosersalisia, Pachystela, Pseudopachystela* and *Tulestea*. He argued that characters separating such genera were uncorrelated with each other and instead used a suite of more consistent characters to define *Synsepalum* (Table 1.2.), consequently increasing the number of species to 16.

Table 1.2 – Characters used by Pennington (1991) to define the genus *Synsepalum*

1.Frequent occurrence of large stipules	6. Endosperm absent
2. Eucamptodromous venation	7. Embryo with plano convex cotyledons
5. Pentamerous flowers	8.Seed broad, not laterally compressed and with a broad scar
4. Corolla nearly always rotate, cyathiform of shortly tubular with wide-spreading lobes	9. Stamens fixed at or near the top of the corolla tube
5. Corolla lobe aestivation, imbricate or duplicate valvate	10. Stamens nearly always exserted, with well developed filaments

Since 1991, the genus has grown (WSCP, 2013) and now contains approximately 35 accepted species (Table A1.1). However, as already stated, in the absence of a proper revision, this figure is still provisional.

1.5 Taxonomic History of *Englerophytum*

The genus *Englerophytum* was first validly published by Krause in 1914. He differentiated *Englerophytum* from all other genera in the family via the fusion of the stamens into a staminal tube and its characteristic parallel leaf venation. At the time of its first publication,

Englerophytum was monotypic, containing only the species *Englerophytum stelecantha* (now known as *E. stelecanthum*).

After the publication of *Englerophytum*, De Wildeman (1919) erected another genus called *Bequaertiodendron* in which he defined another species called *Bequaertiodendron congolense* which also has fused stamens and leaves with parallel venation.

Unaware of the establishment of the previous names, Aubréville & Pellegrin (1957) erected a third genus called *Tisserantiodoxa* in order to accommodate yet another species from Oubangui (Congo), which also had the same staminal and leaf characters. They named the species *Tisserantiodoxa oubanguiensis*.

Thus up till 1957, there were three monospecific genera each accomodating one species with fused stamens and parallel leaf venation – *Englerophytum, Bequaertiodendron* and *Tisserantiodoxa*.

In 1960, Heine & Hemsley, after making surveys in different parts of Africa, decided that it was appropriate to merge the genus *Tisserantiodoxa,* into *Bequaertiodendron* but refrained from merging *Englerophytum* into *Bequaertiodendron* since type material of Krause was lacking. However they stated that *Englerophytum* and *Bequartiodendron* are probably the same genus.

It was Aubréville (1961) who, despite lacking type material, decided to sink *Bequaertiodendron* into *Englerophytum*. In his paper, he emphasizes the shared characters of these genera including (i) leaves in clusters at the tips of branches (ii) numerous parallel lateral veins (iii) and hair on the leaf underside which to him seemed to be reliable enough to merge the taxa. By 1961, the genus *Englerophytum* contained six species – *E. hallei* (later synonymized with *E. stelecanthum*), *E. congoense* (now *E. congolense*), *E.oubanguiense, E. vermoesnii, E. le-testui* and *E. kouloungense*.

In 1971, Aubréville discovered type material of *Englerophytum* which further supported his decision to transfer *Bequaertiodendron* into *Englerophytum*. In this paper Aubréville also makes remarks on the strong similarity of the vegetative characters of the genera *Pseudoboivinella, Neoboivinella, Wildemaniodoxa, Zeyherella* and *Englerophytum*.

Pennington (1991) makes similar observations to those of Aubréville and decides to widen the circumscription of *Englerophytum* to include also the genera *Pseudoboivinella, Neoboivinella, Wildemaniodoxa* and *Zeyherella*. He rejected studies made by Liben (1989) which claimed that *Zeyherella* and *Wildemaniodoxa* deserved generic status.

Hence the present circumscription of *Englerophytum* (WSCP, 2013) includes 14 accepted species (Table A1.2).

1.6 Morphological Evidence

The dynamic taxonomic histories of *Englerophytum* and *Synsepalum* clearly show the difficulty encountered by taxonomists to delimit these two taxa based on morphological characters. The similarities between *Englerophytum* and *Synsepalum* are so significant, that there have been cases where taxonomists (Meeuse, 1960; Baehni, 1965) placed species from *Synsepalum* and *Englerophytum* together in the same genus.

Meeuse (1960) , in his account on the Sapotaceae, put *Synsepalum cerasiferum* and *Synsepalum brevipes* together with *Englerophytum magalismontanum* and *Englerophytum natalense* into a single genus *Pouteria* based on similarities in floral characters (e.g. flower and calyx merism) and fruit and seed characters (e.g. broad seed scar, 1-5 seeded fruit). Baehni (1965) also placed *E. natalense* and *S.msolo* together in the same genus *Amorphospermum* based of floral and fruit characters. Such cases clearly show the complex and overlapping nature of morphological characters in these genera.

The understanding of polythetic variation distinguished Pennington from previous taxonomists and allowed him to make more taxonomically sound decisions in his generic delimitations based on robust suites of characters. However, despite taking this approach, Pennington still expresses a degree of uncertainty about his delimitations of *Synsepalum* and *Englerophytum*, especially when discussing members of *Englerophytum* which he terms ‘poorly defined’.

Pennington admits that there exists very strong similarities between the genera, the most prominent being (i) the frequent presence of stipules (ii) 5-merous flower structure (iii) irregular presence of staminodes (iv) similar seed and embryo characters.

He only decided to split these two genera on the basis of two characters which to him appeared consistent enough to keep *Englerophytum* and *Synsepalum* as separate genera. These characters are leaf venation patterns and filament fusion. *Englerophytum* possesses a consistently striate brochidodromous leaf venation pattern with a strong tendency of fusion of the filaments. In contrast *Synsepalum* usually exhibits eucamptodromous leaf venation patterns and has free stamens. Figures 1.3 and 1.4 show representative species from each genus indicating their major defining characters.

Since Pennington's account, there have been no further revisions of *Englerophytum* and *Synsepalum* and the difficulty in defining these genera still stands to this day. However, with the advent of molecular techniques more evidence has been gained on the taxonomic status of *Englerophytum* and *Synsepalum.*

1.7 Molecular evidence

A series of molecular phylogenetic studies using nuclear and plastid DNA have recently been carried out (Bartish *et al.*, 2005, 2011; Schönenberger *et al.*, 2005; Swenson & Anderberg, 2005; Smedmark *et al.*, 2006; Smedmark & Anderberg, 2007; Swenson *et al.*, 2008) in order to help resolve relationships between the main groups of the Sapotaceae.

Although none of the studies were specifically aimed at resolving inter-generic relationships between *Englerophytum* and *Synsepalum*, some of them (Swenson & Anderberg, 2005; Swenson *et al.*, 2008) made use of a very small sample of species (between 4-7) from these genera and have indirectly provided further insight into their taxonomy. A compilation of *Englerophytum-Synsepalum* trees resulting from these studies is shown in Figure 1.5.

The combined monphyly of *Englerophytum* and *Synsepalum* is very evident in these trees and is usually well supported. However, the relationships within this *Englerophytum-Synsepalum* clade are still rather unclear.

Parsimony analysis of combined data from chloroplast DNA, nuclear DNA and morphology (Tree C) shows species from *Synsepalum* and *Englerophytum* segregating into monophyletic sub-clades according to genus, suggesting that perhaps *Englerophytum* and *Synsepalum* are distinct genera.

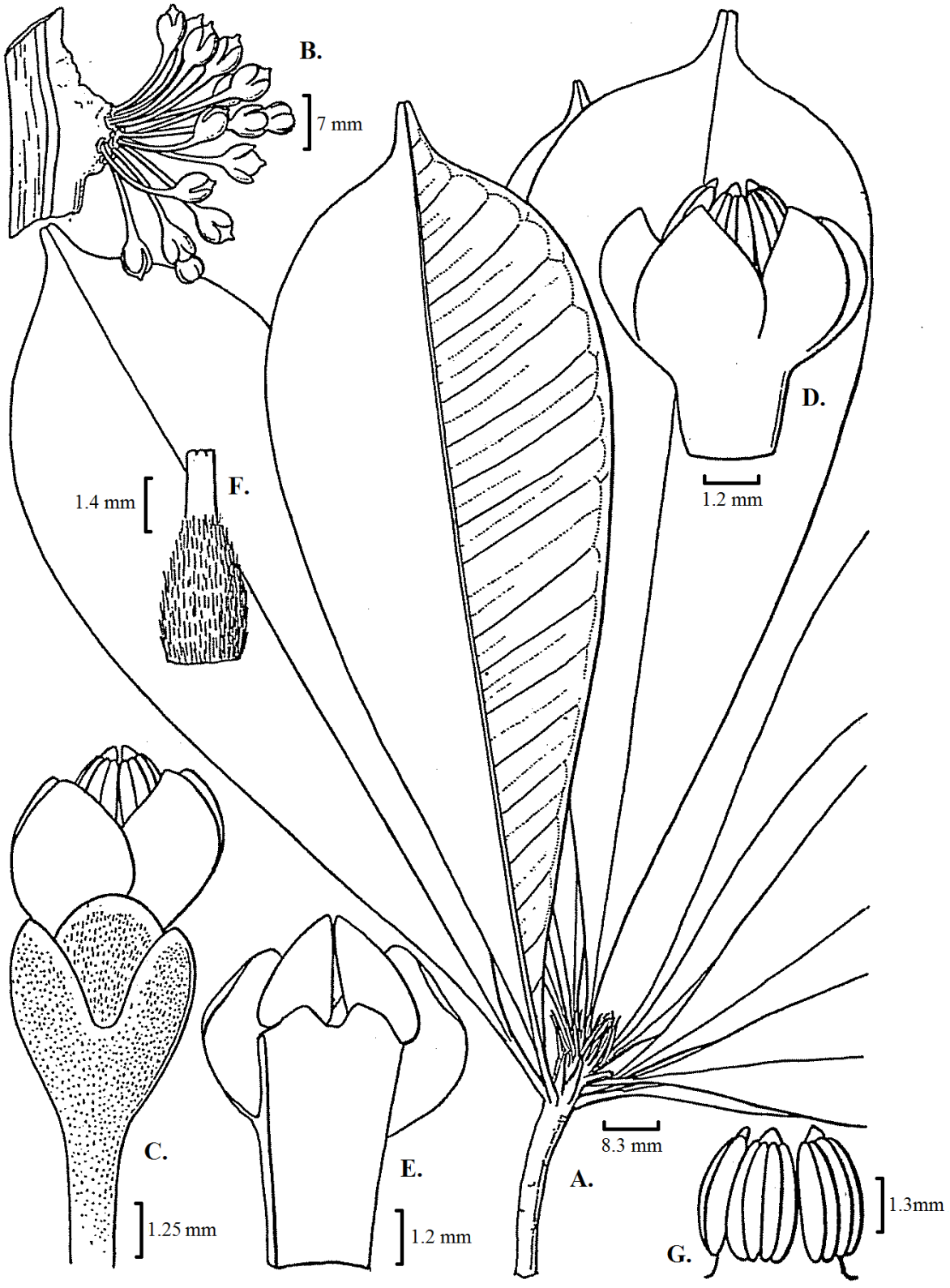


Figure 1.3 - Diagrammatic representation of *Englerophytum stelechanthum* showing the main characteristics of the genus *Englerophytum* (A) Clustered leaves with looped brochiododromous venation (B) Inflorescence (C) Flower (D) Flower with removed calyx (E) Longitudinal section through staminal tube (F) Style and Ovary (G) Extrorse anthers (adapted from Aubréville, 1965)

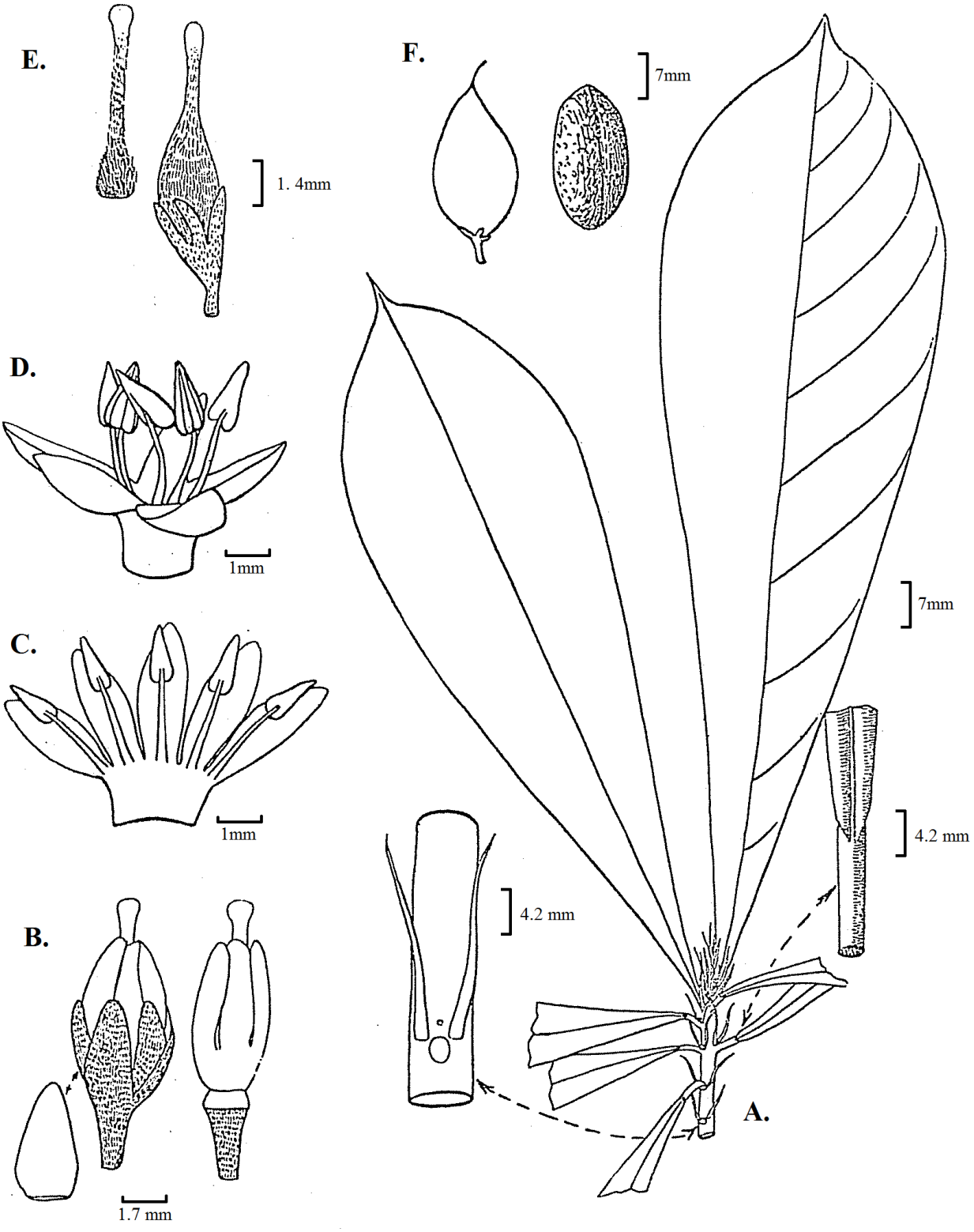


Figure 1.4 - Diagrammatic representation of *Synsepalum brevipes* showing the main morphological characters of the genus *Synsepalum.* (A) Clustered leaves with eucamptodromous venation (B) Flower in bud (C) Dissected stamen-petal tube (D) Flower (E) Style and Ovary (F) Fruit (adapted from Aubréville, 1959)

However, in trees obtained from nuclear DNA (Tree A) and plastid DNA (Tree B and D), this view is not supported. In these trees, *Englerophytum* and *Synsepalum* did not segregate into separate monophyletic generic sub-clades but instead formed a single heterogeneous clade suggesting that perhaps *Synsepalum* and *Englerophytum* are simply one genus.

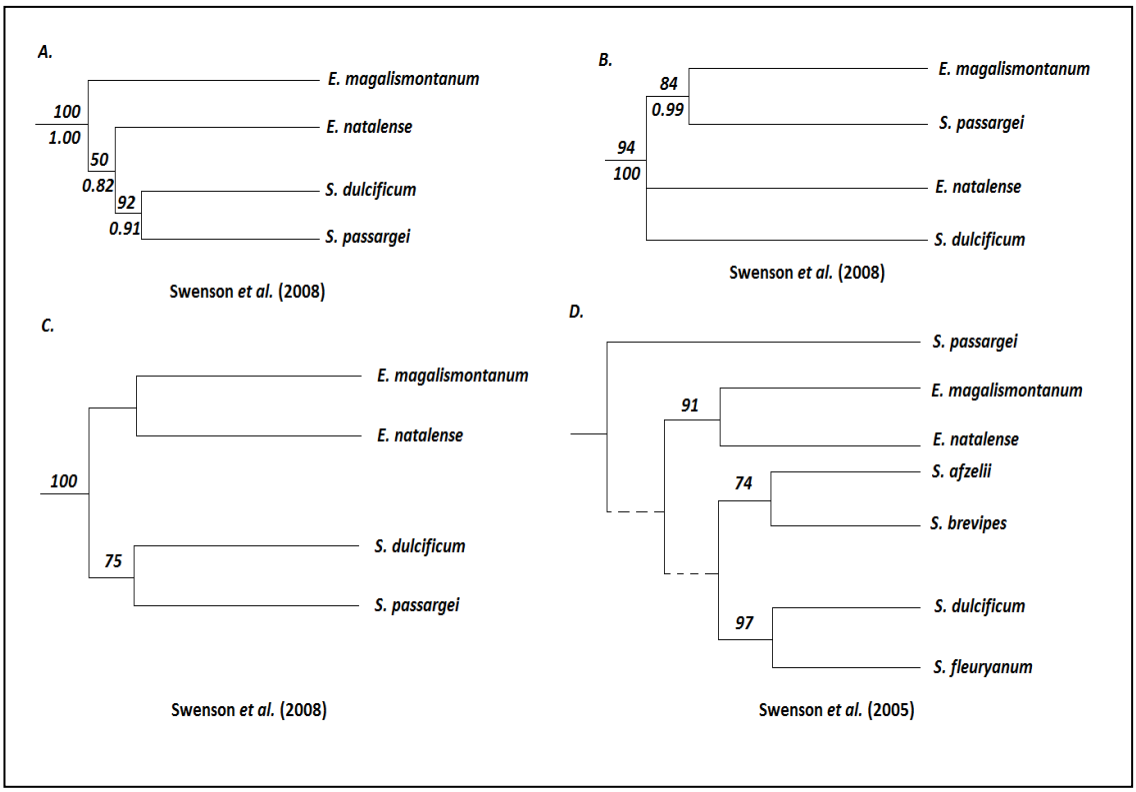


Figure 1.5 - A compilation of trees from previous molecular studies on a small sample of members of the *Englerophytum-Synsepalum* clade. (A) A 50% Bayesian-majority rule consensus tree from a combined matrix of 2 nuclear DNA loci (ITS, Chs1) (B) A 50% majority rule Bayesian consensus tree obtained from a combined matrix of 7 chloroplast loci (C) One of the most parsimonious trees of a combined matrix of 2 nuclear loci, 7 chloroplast loci and morphological data (D) A strict consensus of 19 396 trees obtained from a parsimony analysis of ndhf sequences. Values above branches are jack knife support values whilst values below the branches represent posterior probabilities. Dashed line represent branches that lack jack knife support.

Besides conflict with respect to relationships between the genera there is also some conflict in the relationships between species. An example of this can be shown in the relationships of *S. passargei*. Tree B shows a rather high 0.99 posterior probability supporting the sisterhood of *E. magalisontanum* with *S. passargei*. However, in Tree A the closest sister species to *S. passargei* is *S. dulcificum* rather than *E. magalismontanum.*

Therefore although molecular data confirmed that *Englerophytum* and *Synsepalum* together form a monophyletic clade, it has, as yet, been unable to resolve relationships within this

clade. Part of the reason for this is that the studies made so far were not focused on resolving the relationships between *Englerophytum* and *Synsepalum* and only made use of a very small sample of the clade. Therefore there seems to be the need of further investigation of this clade.

1.8 Aims of the project

This main aims of the project are:

- To elucidate the relationships within the *Englerophytum-Synsepalum* clade through the molecular phylogenetic analysis of the nuclear ITS region and the plastid psbA-trnH region.
- To confirm the monophyly of the *Englerophytum-Synsepalum* clade by using a wider sample of species.
- To determine whether molecular data is consistent with the current generic delimitations of *Englerophytum* and *Synsepalum*
- To gain insight on the morphological synapomorphies shared by members of the *Englerophytum-Synsepalum* clade and its sub-clades.

2. MATERIALS AND METHODS

2.1 Taxon Sampling and selection of genomic regions

2.1.1 Taxon sampling

Half* (50%) of the genus *Englerophytum* and about a third* (31%) of the genus *Synsepalum* were sampled for this molecular study. Where possible, more than one individual per species was sampled. Effort was made to sample the genera from throughout their geographical range. Table A2.1 provides information on all the sequenced accessions.

*In the absence of a proper revision of these genera, percentage estimates are based the total number of species names from the The Plant List (2010). Lists of all the names within each genus are shown in Tables A1.1 & A1.2.

2.1.2 Selection of genomic regions

After referring to past literature on Sapotaceae phylogeny, (Smedmark *et al.*, 2006; Smedmark & Anderberg, 2007; Swenson *et al.*, 2008; Armstrong, 2010), it was decided to use two DNA regions for this study. The first was the nuclear ribosomal internal transcribed spacer (ITS). This is a relatively fast evolving region (Baldwin & Markos, 1998) and thus can be very useful at resolving relationships between closely related species (Feliner & Rossello, 2007). The second chosen region was the plastid trnH-psbA region. This region is a slower evolving region and can be useful at resolving relationships in the backbone of the phylogenetic tree (Shaw *et al.*, 2005).

2.2 Laboratory Protocols

2.2.1 DNA Extraction

A sample of about 20mg of leaf material, from silica gel-dried material or herbarium sheets, was placed in a 2ml Eppendorf tube together with an angular tungsten grinding bead. The leaf sample was then ground in a Qiagen Mixer Mill at 20Hz for 1-2minutes and the process was repeated until the sample was ground into a fine powder. This procedure was repeated for all leaf samples.

After consulting previous molecular work on the Sapotaceae (Smedmark *et al.*, 2006; Smedmark & Anderberg, 2007; Armstrong, 2010), the Qiagen Plant DNeasy DNA extraction protocol was chosen as the method for DNA extraction of *Englerophytum* and *Synsepalum.* Therefore, all the DNA samples were extracted using Qiagen's Plant DNeasy Mini Kit following the manufacturer's instructions.

2.2.2 Polymerase Chain Reaction (PCR)

Target DNA regions, that is, the nuclear ITS region and the plastid trnH-psbA region, were amplified using the polymerase chain reaction. PCR reactions of 25µl were set-up for each region using the recipes in Table 2.1. The recipes were taken from Armstrong (2010) with minor modifications made to the volume of BioTaq DNA polymerase used.

Table 2.1 – 25µl PCR reaction recipes for the ITS and trnH-psbA regions

Reagent	ITS recipe	trnH-psbA recipe
ddH_20	5.875 µL	15.25 µL
NH_4 (10x reaction buffer)	2.5 µL	2.5 µL
dNTP (2mM)	2.5 µL	2.5 µL
$MgCl_2$ (25mM)	1.25 µL	1.25 µL
Forward Primer (10 µM)	0.75 µL	0.75 µL
Reverse Primer (10 µM)	0.75 µL	0.75 µL
BioTaq DNA polymerase (Bioline UK) 5unit/ µL	0.125 µL	0.2 µL
Betaine (5mM)	10 µL	-
Bovine Serum Albumin (BSA 0.4%)	0.25 µL	0.8 µL
DNA template	1 µL	1 µL

The DNA was amplified on a Tetrad2 BioRadDNA Engine using the primers ITS5p and ITS8p for the ITS region and the primers trnH and psbA for the chloroplast region. Details of the primers are given in Table A.2.2.The PCR programs used for this amplification are shown in Tables 2.2 and 2.3 below:

Table 2.2 - PCR Program for the ITS region

1. Incubate at 95ºC for 5 minutes (denaturation)
2. Incubate at 95ºC for 30 seconds (denaturation)
3. Incubate at 50ºC for 30 seconds (annealing)
4. Incubate at 72ºC for 1 minute and 30 seconds (extension)
5. Cycle to step 2 for 34 more times
6. Incubate at 72ºC for 8 minutes (extension)
7. Incubate at 10ºC "forever" (finished)

Table 2.3 – PCR Program for the trnH-psbA region

1. Incubate at 80^{o}C for 5 minutes (denaturation)
2. Incubate at 95^{o}C for 1 minute (denaturation)
3. Incubate at 50^{o}C for 1 minute (annealing)
4. Ramp to 65^{o}C at 0.3^{o}C per second
5. Incubate at 65^{o}C for 4 miutes (extension)
6. Cycle to step 2 for 29 more times
7. Incubate at 65^{o}C for 5 minutes (extension)
8. Incubate at 10^{o}C "forever" (finished)

2.2.3 Gel Electrophoresis

Gel electrophoresis was used in order to check whether the target sequences were amplified successfully. A 1% agarose gel was made by adding 1g of agarose powder to 100mL of 1xTri-Borate-EDTA (TBE) buffer. The mixture was heated in a microwave until it became clear and later 5 µL of Sybrsafe DNA gel stain were added to it. Whilst still in the liquid state, the gel solution was poured in a tray fitted with combs, covered from light (to prevent degradation of Sybrsafe) and allowed to solidify.

Once the gel solidified, it was transferred into a tray containing 1xTBE buffer solution. A mixture containing 5 µL of PCR product and 3 µL of gel loading dye was loaded into the wells together with a 1kb ladder and a negative sample for reference. Electricity was then allowed to pass through the gel (80V, 400A) for 40 minutes so as to allow the negatively charged DNA in the wells to migrate to the positive electrode. The resulting gel was

visualized in a UV lightbox and successful PCR runs could be noted by the appearance of a fluorescent band on the gel.

2.2.4 Nested PCR reaction for ITS

In some cases, the ITS region did not amplify well enough to show a band on the agarose gel. Whenever DNA yield from the first PCR was low, a nested PCR was employed. This type of PCR makes use of two primers, ITS 1 and ITS 4 (Table A2.2), which attach within the region bounded by ITS5p and 8p and help to amplify the target region further. The nested PCR protocol was often required for material from herbarium specimens whose DNA quality was poorer than material stored in silica gel and did not give any bands in the first PCR.

The recipe and the program used for the nested PCR were the same as those used for the first PCR (Table 2.1 & 2.2) with two minor differences: (i) the primers used for the nested PCR were ITS 1 and 4 instead of ITS 5p and 8p (ii) the number of cycles in step 5 were reduced from 34 to 29 to reduce the risk of amplifying other non-target DNA fragments during the highly sensitive nested reaction.

The chloroplast region was much less problematic to amplify than the ITS region and no nested protocol was required. Samples which failed to show a band in the first run were re-tried using the same chloroplast recipe and program (Table 2.1 & 2.3) but with a larger volume of template (2 µL) and an extended number of cycles (39)

2.2.5 Purification of the PCR Product

All the PCR products from both the ITS and chloroplast regions were purified using the ExoSAP-IT protocol (GE Healthcare). To each 5µL sample of PCR product, 2µL of ExoSAP-IT enzyme solution was added. The resulting 7µL solutions were then centrifuged and subjected to the program below (Table 2.4).

Table 2.4 - ExoSAP-IT PCR Purification Program (7 µl reaction)

1. Incubate at 37º C for 15 minutes
2. Incubate at 80º C for 15 minutes
3. Incubate at 10º C “forever”

2.2.6 Sequencing

An aliquot of 1μL of purified PCR product was then used to prepare the sequencing mix according to the recipe below (Table 2.5).

Table 2.5 - Sequencing PCR Recipe for 10μl reaction

Reagent	Volume
Water	6.18 μL
ABI Sequencing Buffer	2 μL
Primer	0.32 μL
Big Dye	0.5 μL
Template	1 μL

In order to obtain a bi-directional read, two 10μL reaction mixtures were prepared for each sample, one containing the forward primer and one containing the backward primer. Depending on the nature of the products to be sequenced, different primers were used. The primers ITS5p & 8p were used for non-nested products, ITS 1 & 4 were used for nested products whilst the primers psbA and trnH were used for sequencing chloroplast DNA. The sequencing mixtures were then put into the Tetrad2 BioRadDNA Engine and subjected to the program below:

Table 2.6 - Sequencing PCR Program

1. Incubate at 95ºC for 30 seconds (denaturation)
2. Incubate at 50ºC for 20 seconds (annealing)
3. Incubate at 60ºC for 4 minutes (extension)
4. Cycle to step 1 for 24 more times
5. Incubate at 4ºC “forever” (finished)

Once the sequencing reaction was finished, the samples were sent for sequencing at the University of Edinburgh’s GenePool facility using ABI 3730 sequencer.

2.3 Processing of DNA Sequences

2.3.1 Sequence editing

Raw sequences were then imported into Sequencher ® ver. 5.1 (Gene Codes Corporation, 2013) and the primer sequences were trimmed off their ends. The trimmed forward and reverse sequences of each sample were then assembled into contiguous sequences (contigs) in preparation for alignment and any conflicts between the sequences were resolved to form a consensus sequence. Sequences were subjected to a nucleotide BLAST on the GenBank website (Benson *et al.*, 2011) to check for contaminants.

2.3.2 Alignment

Alignment was carried out using Bioedit ver. 7.2.0 (Hall, 1999). Initially an automatic alignment was carried out using the Clustal W Multiple alignment application (Thompson *et al.*, 1994). The alignment was then refined manually.

2.3.3 Gap Coding

Gaps were coded manually using binary characters according to the simple gap coding method (Simmons & Ochoterena, 2000). Gaps whose size was not equal to recurring gap patterns were assigned a question mark (?) and gaps in areas of multiple repeats were not coded. Once all gaps were coded, the aligned matrix was then converted to a nexus file using ProSeq3.5 (Filatov, 2009) in preparation for analysis.

2.4 Phylogenetic Analysis

2.4.1. Outgroup Selection

Eight outgroups were selected for this analysis by making reference to papers from Swenson & Anderberg (2005), Swenson *et al.* (2008) and Bartish *et al.* (2011). The outgroups were chosen in a way that could provide a suitable subfamilial framework within which the *Englerophytum –Synsepalum* clade could be analysed

The genus *Eberhardtia*, which is sister to the sub-family Chrysophylloideae, was also included in the analysis so as to provide further structure to the tree. A list of the selected outgroups and their GenBank accession numbers is given in Table A2.3.

2.4.2 Excluded and Misidentified Taxa

Four specimens, out of a total of 44, were excluded from the final analysis (Table A2.1). Three of the four excluded samples did not amplify well for neither the ITS nor the trnH-psbA region and thus could not be incorporated. The fourth excluded sequence ('Harris 8702') amplified well for both regions however, its poor alignment with the rest of the samples and its repeated appearance as sister to the *Englerophytum-Synsepalum* clade, roused suspects of the possibility of the sample being misidentified. A subsidiary analysis on this sample (Appendix 2, Section A2.5) proved that this sample belonged to the Sapotoideae and was closely related to *Neolemmoniera*. Hence it was excluded from the analysis.

This was not the only case of misidentification in the samples. However, since *Englerophytum* and *Synsepalum* have not yet been revised, some misidentifications in the sample material were expected. Therefore, prior to analysis, extra care was taken to ensure that the material sequenced was correctly identified. Whenever possible the material was examined and relabeled before being incorporated in the analysis. Only 3 samples were found to be misidentified. A list of these samples, their new determinations and reasons for making the changes are listed in Table A2.4.

2.4.3 Analysis of Datasets

Both ITS and trnH-psbA regions were initially analysed separately using Maximum Parsimony and Bayesian Inference Methods (explained below). After analyzing the trees by eye, and observing that no well supported incongruence was present between the trees, the data sets were combined into a single matrix and analysis was carried out on the combined data. Since the analyses were computationally intensive, they were carried out externally using the Oslo Bioportal. (Kumar *et al.*, 2009)

2.4.4 Maximum Parsimony

Heuristic Parsimony searches were implemented on the ITS, trnH-psbA and combined matrices, using PAUP* ver. 4.0 (Swofford, 2003). The heuristic search was made using TBR (Tree-Bisection Reconnection) branch swapping with 100,000 random-addition replicates and saving up to 1,000,000 trees. All character states were treated as unordered and equally weighed. Gaps were treated as missing data. The trees obtained were then saved with their branch lengths and Strict, Semi-Strict, Majority Rule and Adams consensus trees were

generated. A parsimony bootstrap search, with 10, 000 replicates, was also performed in order to obtain branch support bootstrap values.

2.4.5 Bayesian Inference

Bayesian analysis was carried out using Mr. Bayes ver.3.2.1 (Ronquist & Huelsenbeck, 2003). Prior to running the Bayesian analysis, jModel Test ver. 2.1.4 (Darriba *et al.*, 2012) was used together with the Bayesian Information Criterion (BIC) in order to select the optimum model of evolution for both the ITS and chloroplast sequences.

Once the models were chosen, the Markov Chain Monte Carlo (MCMC) algorithm was run for 20,000,000 generations with 1 cold and 3 heated chains, starting from a random tree and sampling one out of every 1000 generations. Tracer v.1.5 (Rambaut & Drummond, 2009) was used to check for convergence and estimate burn-in. Trees falling within the burn-in region were discarded and the remainder were used to construct a Bayesian consensus tree.

2.5 Morphology

2.5.1 Selection of Morphological Characters

Characters were selected by making reference to previous taxonomic treatments of *Englerophytum* and *Synsepalum* (Meeuse, 1960; Aubréville, 1961, 1964, 1965; Baehni, 1965; Pennington, 1991; Swenson & Anderberg, 2005) and noting the characters that were considered important to define the genera. Herbarium specimens were also examined and any additional characters which could potentially have taxonomic significance were identified. A list of the selected characters and their respective character states is presented in Table A2.6.

2.5.2 Analysis of Morphological Characters

Specimens from E, FHO and K were examined for morphological characters using a Leica MZ75 standard binocular dissecting microscope. Some online herbaria (HBG, BM, P, LISC) were also consulted in order to facilitate scoring of vegetative characters of species which were underrepresented in the study material. A BRAHMS-generated report of all the analyzed specimens is provided in Appendix A2.7.

Whenever floral material was available, this was boiled and dissected on card for analysis. In cases where floral material was lacking, floral characters were scored from literature (Lecomte, 1928; Meeuse, 1960; Aubréville, 1961, 1964, 1965; Baehni, 1965; Pennington, 1991; Swenson & Anderberg, 2005). The character matrix from the morphological analysis is shown in Table A2.8.

2.5.3 Mapping Morphological Characters on the Phylogenetic Tree

Once the character matrix was compiled, it was transferred into Mesquite ver. 2.75 (Maddison & Maddison, 2011). Morphlogical characters were mapped over one of the most parsimonious trees using the "Trace Character History" function in Mesquite and selecting the Parsimony Ancestral States option. The resulting trees were then analysed for morphological similarities and differences.

2.6 Databasing

Information from all the analysed specimens was inputted into BRAHMS ver. 7.3 (Brahms ©, 2013) in order to (i) facilitate organisation and handling of data (ii) generate reports of species analysed (iii) Explore and learn about the various advantages of databasing.

Data was inputted in the database by two main methods. The first method was that of manual data input. This method was used when no prior databased information was present about the specimens.

Whenever databased information was already present, this was imported into BRAHMS via its Rapid Data Entry option. All information of material from E, was first entered manually into the main garden database, BG-BASE (Walter & O'Neal, 2013), and then downloaded and imported into BRAHMS.

3. RESULTS

3.1 The Datasets

3.1.1 The ITS Region

DNA from the ITS region was more difficult to amplify than that from the chloroplast region and several samples, especially those from herbarium specimens, required a nested protocol. Despite this, the sequencing reaction produced very clean reads for this region and few problems were encountered.

3.1.2 The Chloroplast Region

A persistent problem was encountered when sequencing the chloroplast region. In all samples, a rich poly A motif was present at the beginning the trnH-psbA sequence. This repeatedly caused slippage of the polymerase enzyme; an issue which was impossible to resolve in the limited time available for this project. Therefore, during sequencing, the forward reaction, which was encountering the poly A region early on in the sequencing process, was producing very poor low quality reads since the polymerase slippage at the beginning was disrupting the whole read. On the other hand, the backward reaction was encountering the poly A region at the end of the sequencing process and hence was not being disrupted, producing nice clean reads. As a consequence, the consensus sequences used in the chloroplast analysis mostly relied on information from the backward read and had to have a considerable amount of bases trimmed from the ends. However, it was still possible to procure a good DNA read of 450-550 characters from the clean backward reaction and this was also incorporated in the analysis.

3.1.3 Combining Datasets

No hard incongruence was present between the ITS and trnH-psbA datasets and a decision was made to combine the datasets. Since the chloroplast region possessed less characters than the ITS region, it had a lesser impact on the final topology of the combined tree than the ITS region. Despite not influencing the topology much, it has increased the support of some clades in the analysis. Section A3.1 in Appendix 3 shows the combined datamatrix.

3.2 Maximum Parsimony

3.2.1 Parsimony Statistics

The statistics for each of the parsimony searches carried out on the datasets are shown in Table 3.1 below:

Table 3.1 – Parsimony Statistics

	ITS	trnH-psbA	Combined
Total Aligned Length	952	760	1712
Excluded Characters	38	172	210
Included Characters	914	588	1502
Parsimony Informative Characters	259	35	139
Uninformative characters	104	31	290
Variable sites	363	66	429
Percentage Variability	38%	8.7%	25%
Constant Sites	551	522	1073
CI	0.6492	0.9000	0.6684
RI	0.8489	0.9130	0.8498

Approximately a third (38%) of the ITS region is variable with 71% of the variable sites (259) being parsimony informative. The chloroplast region only varies in 8.7% of its sites and only 53% of these variable sites are parsimony informative. ITS thus appears to be more variable than trnH-psbA.

The retention and consistency indices, which give an idea of the amount of homoplasy in the tree, show that the chloroplast region (CI = 0.9, RI = 0.91) has less homoplasy than the ITS region (CI = 0.65, RI = 0.85). The combined dataset has values in between these two extremes.

Trees from the parsimony analysis are shown in Figure 3.1-3.3.

3.2.2 Trees from the Parsimony Analysis

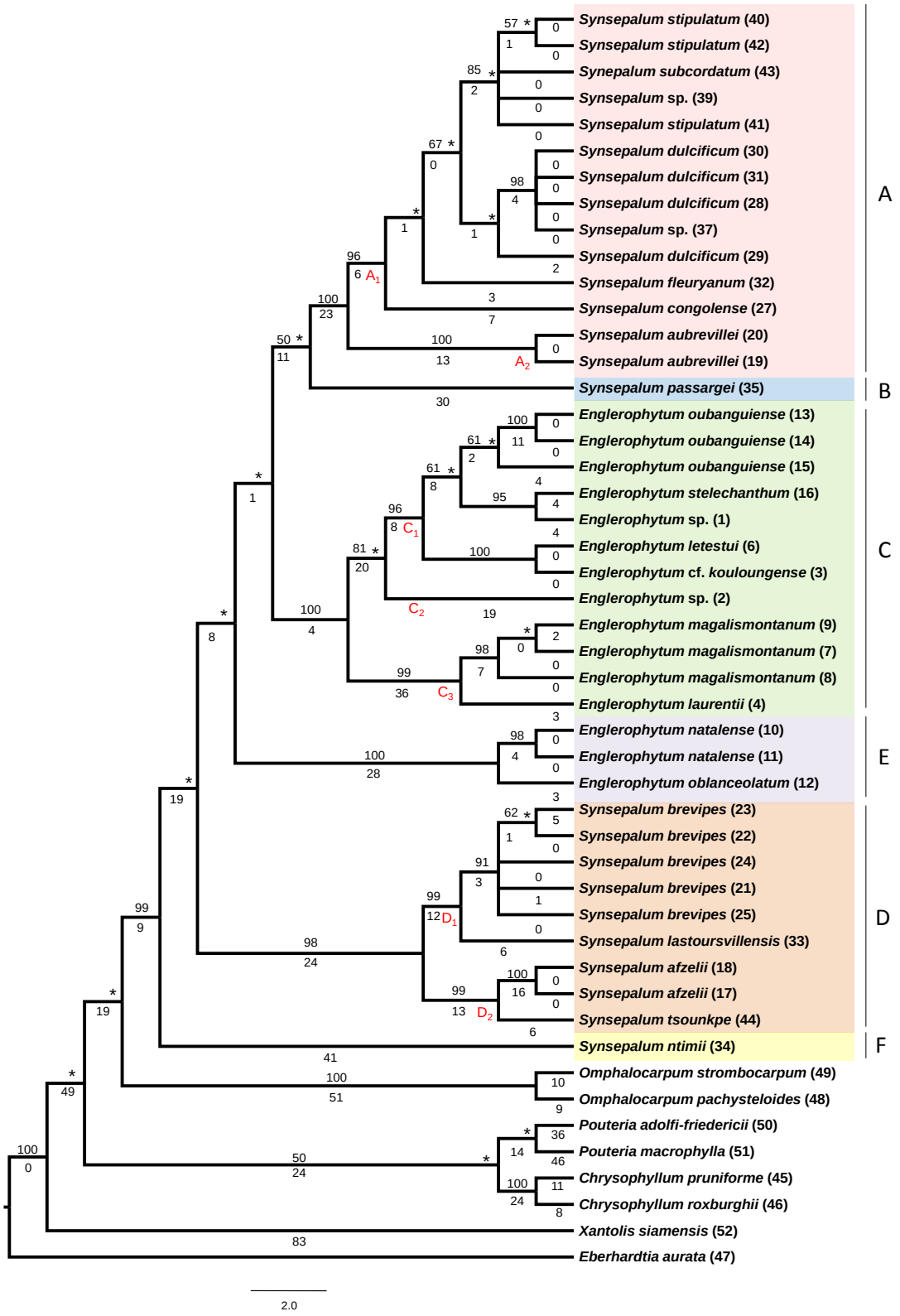


Figure 3.1 – 1 of 16,496 most parsimonious trees reconstructed from nuclear ITS data. Bootstrap support (bs) values are shown above the branches and branch length values are shown below. The symbol * indicates branches which collapse in the strict consensus. The tree has a consistency index of 0.5794 and a retention index of 0.8489. Tree length is 687steps.

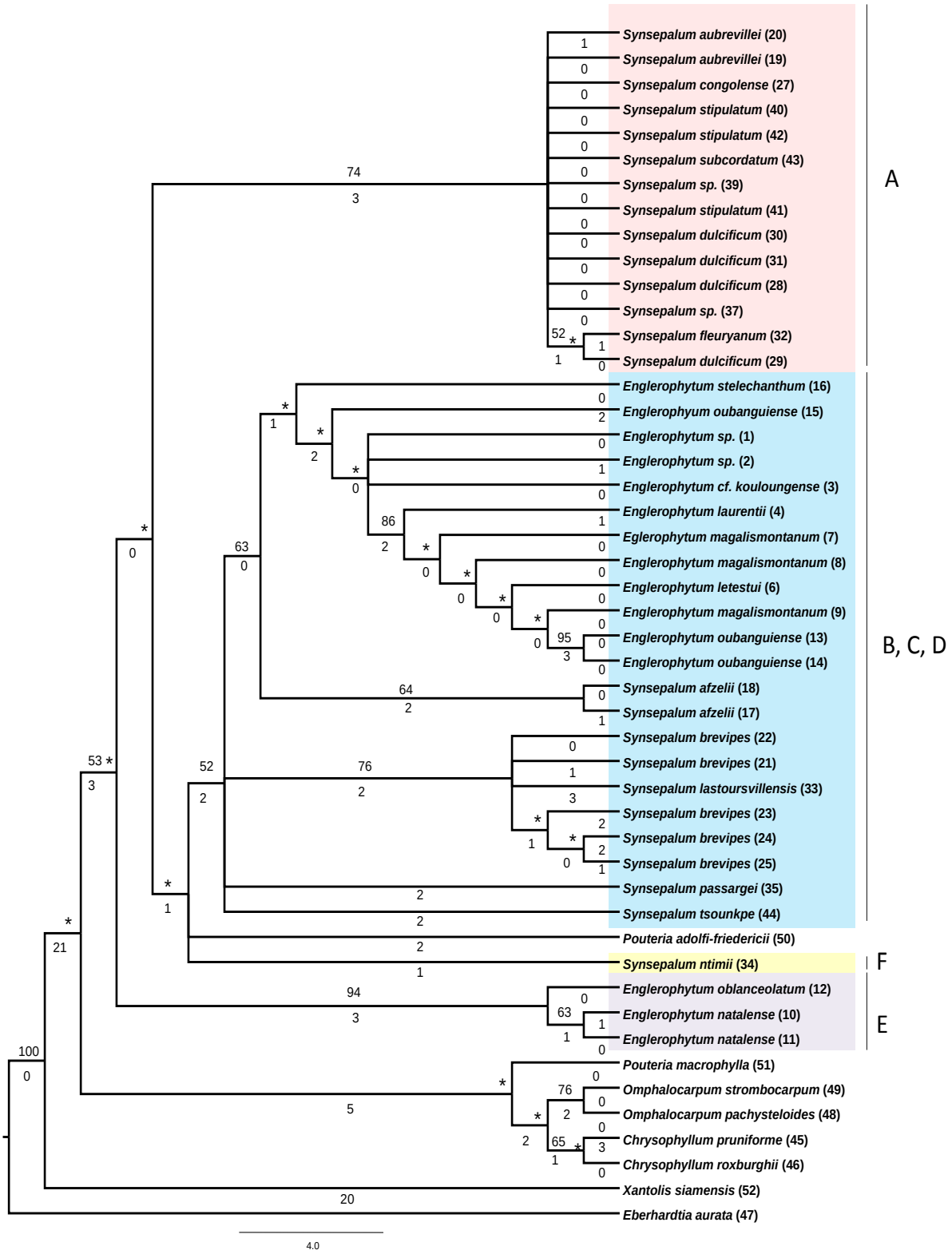


Figure 3.2 - One of 40,000 most parsimonious trees reconstructed from chloroplast data. Bootstrap support (bs) values are shown above the branches and branch length values are shown below. The symbol * indicates branches which collapse in the strict consensus. The tree has a consistency index of 0.9 and a retention index of 0.9310. Tree length is 80 steps.

Figure 3.3 - One of 29319 most parsimonious trees reconstructed from the combined dataset of the nuclear and chloroplast DNA. Bootstrap support (bs) values are shown above the branches and branch length values are shown below. The symbol * indicates branches which collapse in the strict consensus. The tree has a consistency index of 0.668 and a retention index of 0.850. Tree length is 775 steps.

3.3 Bayesian Inference

3.3.1 Partitioning of the Data

The data was split into three partitions representing the ITS region, the trnH-psbA region and binary gap data. The model used for each partition is shown in the table below:

Table 3.2 – Models used for each partition of the Bayesian analysis

Partition	Region	Model Used
1	ITS	SYM+G
2	trnH-psbA	F81+G
3	Gap data	F81

Initially the ITS region was going to be partitioned into three partitions representing ITS1, 5.8s and ITS2. However, when the analysis was run using three ITS partitions, no overall difference was observed between the three-partition tree and the tree with only one ITS partition, except for insignificant changes in the support values. Hence the simpler model was chosen and the ITS data was not partitioned.

3.3.2 Trees from the Bayesian Analysis

Trees from the Bayesian analysis are shown in Figures 3.4-3.6 overleaf

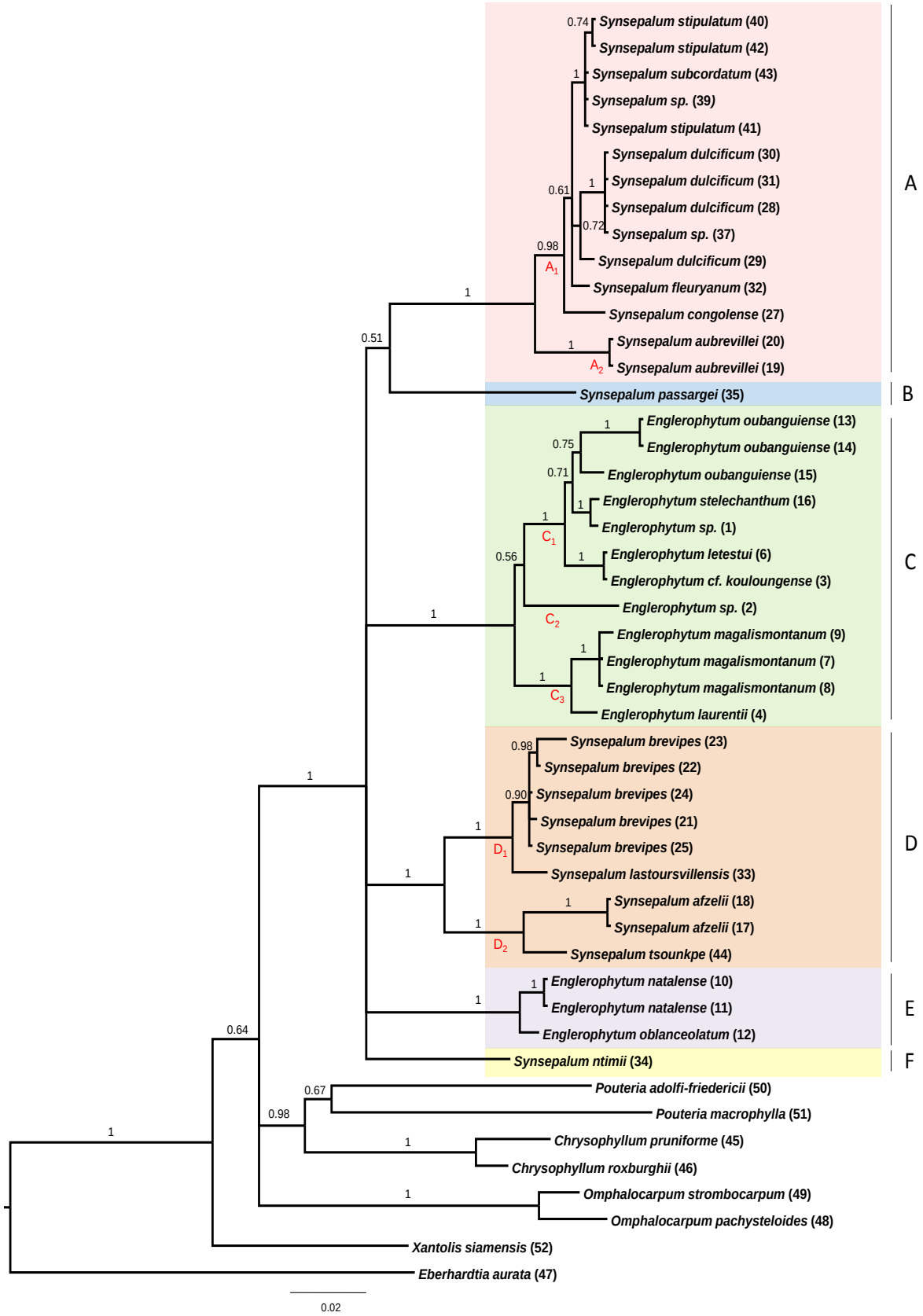


Figure 3.4 - A 50% majority rule Bayesian consensus tree of the ITS dataset showing the major sub-clades (Clades A-E) within the *Englerophytum-Synsepalum* clade and their posterior probabilities. Numbers in brackets are specimen numbers linking the specimens to information in Tables A2.0 and A.2.1

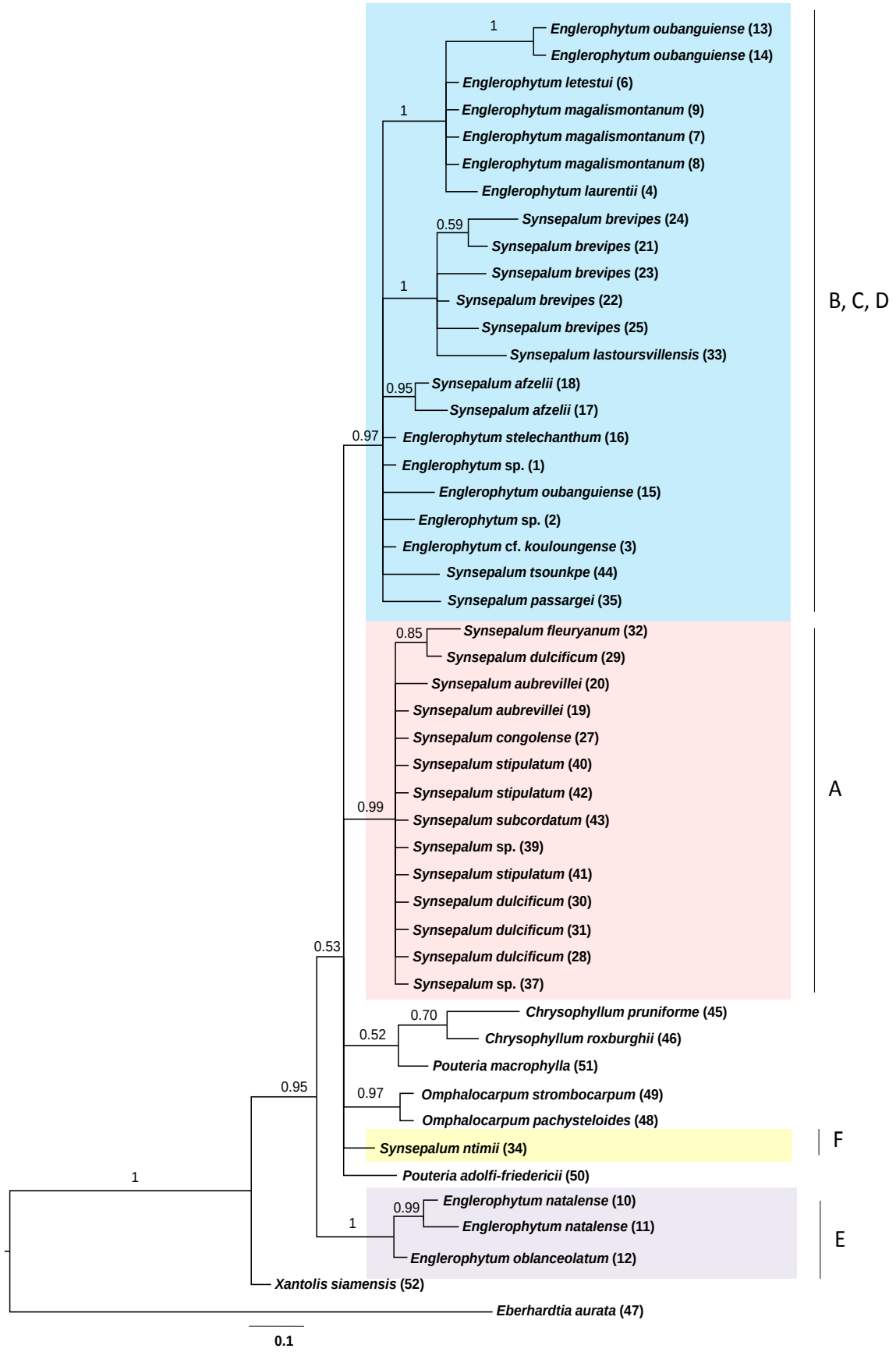


Figure 3.5 - A 50% majority rule Bayesian consensus tree of the chloroplast dataset showing the major sub-clades (Clades A-E) within the *Englerophytum-Synsepalum* clade and their posterior probabilities. Numbers in brackets are specimen numbers linking the specimens to information in Tables A2.0 and A.2.1

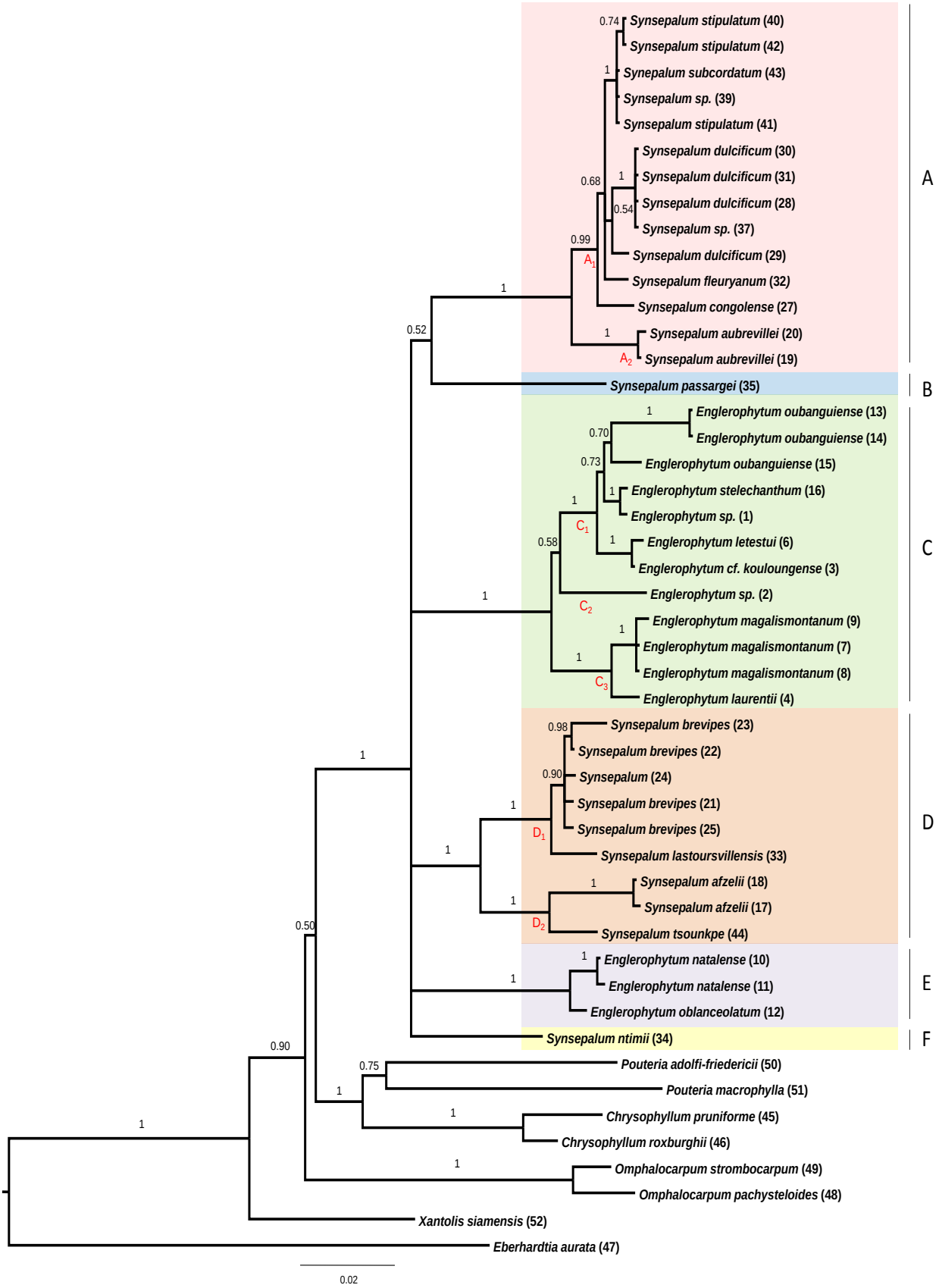


Figure 3.6 - A 50% majority rule Bayesian consensus tree of the combined dataset showing the major sub-clades (Clades A-E) within the *Englerophytum-Synsepalum* clade and their posterior probabilities. Numbers in brackets are specimen numbers linking the specimens to information in Tables A2.0 and A.2.1

3.4 Phylogenies

3.4.1 Monophyly of Chrysophylloideae and the *Englerophytum-Synsepalum* (ES) Clade

All the trees obtained from Bayesian inference and Parsimony analysis show maximum support for the presence of a monophyletic Chrysophylloideae Clade inclusive of all the ingroup and outgroup taxa except *Eberhardtia*, which is sister to the sub family.

The monophyly of the ES clade is strongly supported in the ITS and combined datasets (prob = 1, bs =99) but the resolution in the chloroplast dataset is too low to provide strong support. In the chloroplast tree, the ES clade collapses into a Chrysophylloideae polytomy in the strict consensus.

Chloroplast data, using more markers, has been shown (Swenson & Anderberg, 2005; Swenson *et al.*, 2008; Bartish *et al.*, 2011) to provide well supported monophyletic ES clades (prob=1,bs=94). Thus although, ES monophyly cannot be explicitly determined in our chloroplast dataset, it is highly unlikely that, if further analysis was made, the ES clade would prove to be polyphyletic. Better quality trnH-psbA reads and more chloroplast data could help provide better resolution and more support to the ES clade

3.4.2 Major lineages in the *Englerophytum-Synsepalum* Clade

A total of six lineages could be identified within the ES clade from the Parsimony and Bayesian analysis of the ITS and combined datasets. These lineages are labelled A-E in the Figures 3.1-3.6 and are summarised in the Table 3.3.

Four of these lineages exclusively consist of species currently belonging to the genus *Synsepalum* whilst the two other lineages exclusively consist of species currently belonging to *Englerophytum.* None of these lineages contained a mixture of species from both *Englerophytum* and *Synsepalum.* All lineages occur together in a six-way polytomy at the base of the strongly supported ES clade.

Table 3.3 – Major lineages within the *Englerophytum-Synsepalum* clade

Clade	Constituent Taxa	Probability*	Bootstrap*	Branch* length
A	*S. dulcificum, S. congolense, S. stipulatum, S. fleuryanum, S. aubrevillei, S. subcordatum*	1.00	100	25
B	*S. passargei*	N/A	N/A	33
C	*E. oubanguiense, E. stelechanthum, E. magalismontanum, E. letestui,, E. laurentii*	1	100	7
D	*S. brevipes, S. afzelii, S. lastoursvillensis, S. tsounkpe*	1	99	19
E	*E. natalense, E. oblanceolatum*	1	100	39
F	*S. ntimii*	N/A	N/A	42

* Posterior probability, bootstrap and branch length values were taken from the combined trees

Some of these lineages were also present in the chloroplast trees. Clade A is particularly well resolved in the chloroplast data, with a strong posterior probability (0.99) in the Bayesian analysis but a somewhat weaker bootstrap support in the parsimony tree (bs=74). Clade E (bs=94, prob=1) and lineage F are also easily distinguishable from other clades in the chloroplast tree.

On the other hand, lineages B, C, D are poorly resolved and generally appear in a single heterogeneous clade which is very weakly supported in the parsimony tree (bs=52) but somewhat better supported in the Bayesian analysis (prob=0.97).

3.4.3 Sub-clades within the major lineages

Three of the six major lineages (i.e. Clades A, C, D) also possessed well defined sub-clades within them. A summary of the sub-clades is given in Table 3.4. These sub-clades were clearly visible in the ITS and combined trees and were indicated accordingly with red labels.

However, since the chloroplast tree was poorly resolved, these clades were not mapped onto it.

Table 3.4 – Major sub-clades within clades A, C and D

Major Clade	Sub-Clade	Constituent Taxa	Probability*	Bootstrap*	Branch Lengths*
A	A_1	*S. dulcificum, S. congolense, S. stipulatum, S. subcordatum, S. fleuryanum,*	0.99	97	6
	A_2	*S. aubrevillei*	1	100	13
C	C_1	*E. oubanguiense, E. stelechanthum, E. letestui*	1	95	8
	C_2	*Englerophytum sp.* (Jongkind 5084)	N/A	N/A	20
	C_3	*E. laurentii, E. magalismontanum*	1	100	38
D	D_1	*S. brevipes, S. lastoursvillensis*	1	100	14
	D_2	*S. afzelii, S. tsounkpe*	1	100	13

* Posterior probability, bootstrap and branch length values were taken from the combined trees

3.5 Tracing morphology on the tree

Morphology was mapped over one of the most parsimonious trees from the combined analysis using the Parsimony ancestral states option on Mesquite. A selection of trees showing the most significant morphological differences have been displayed in Figures 3.7 and 3.8.

Several morphological characters were shown to help distinguish major lineages from each other. These characters will be discussed further in Chapter 4.

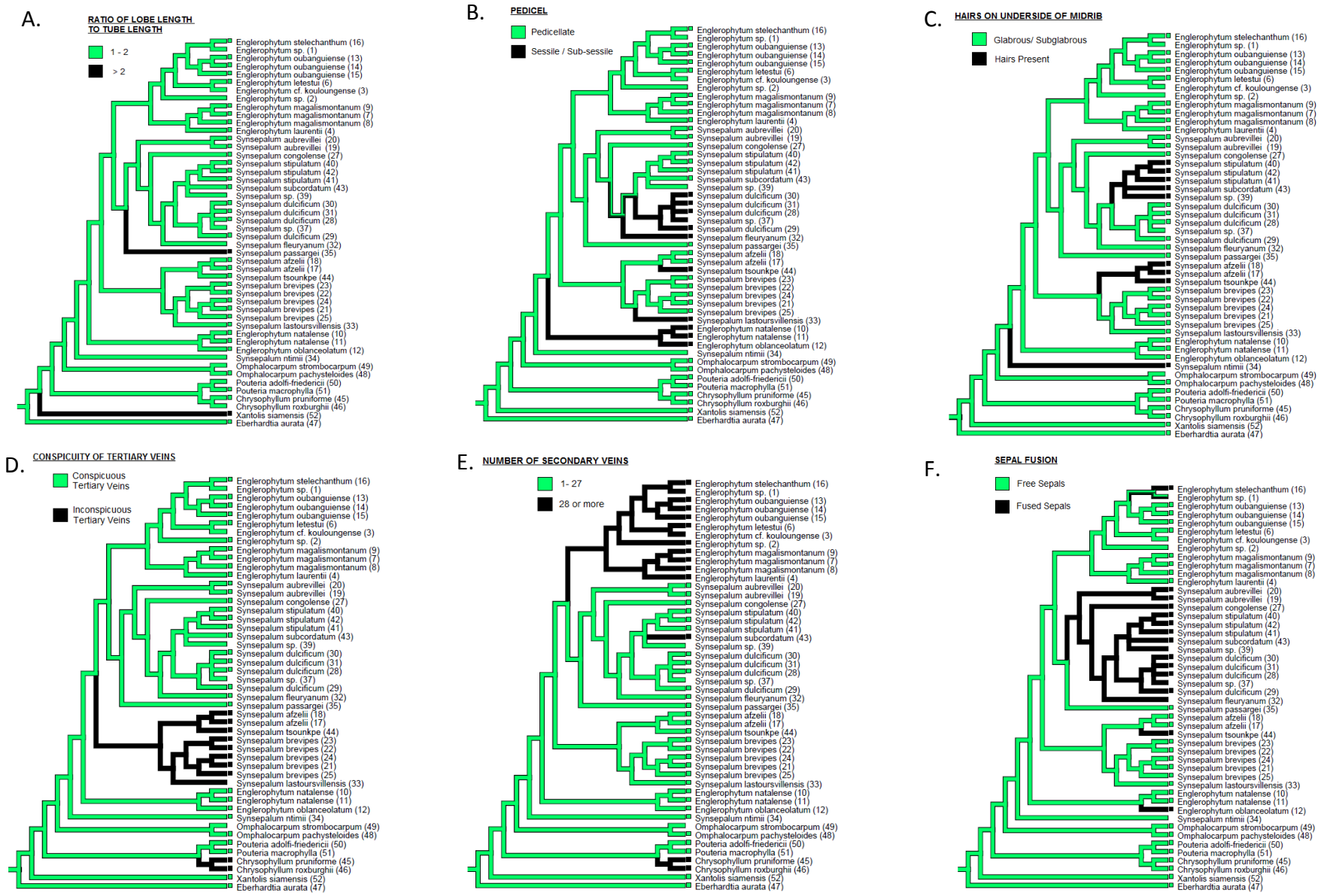


Figure 3.1 – A selection of morphological characters being mapped on one of the most parsimonious trees of the combined dataset. (A) Ratio of Lobe Length to Tube Length (B) Pedicel Length (C) Hairs on the underside of the midrib (D) Conspicuity of Tertiary Veins (E) Number of Secondary Veins (F) Sepal Fusion

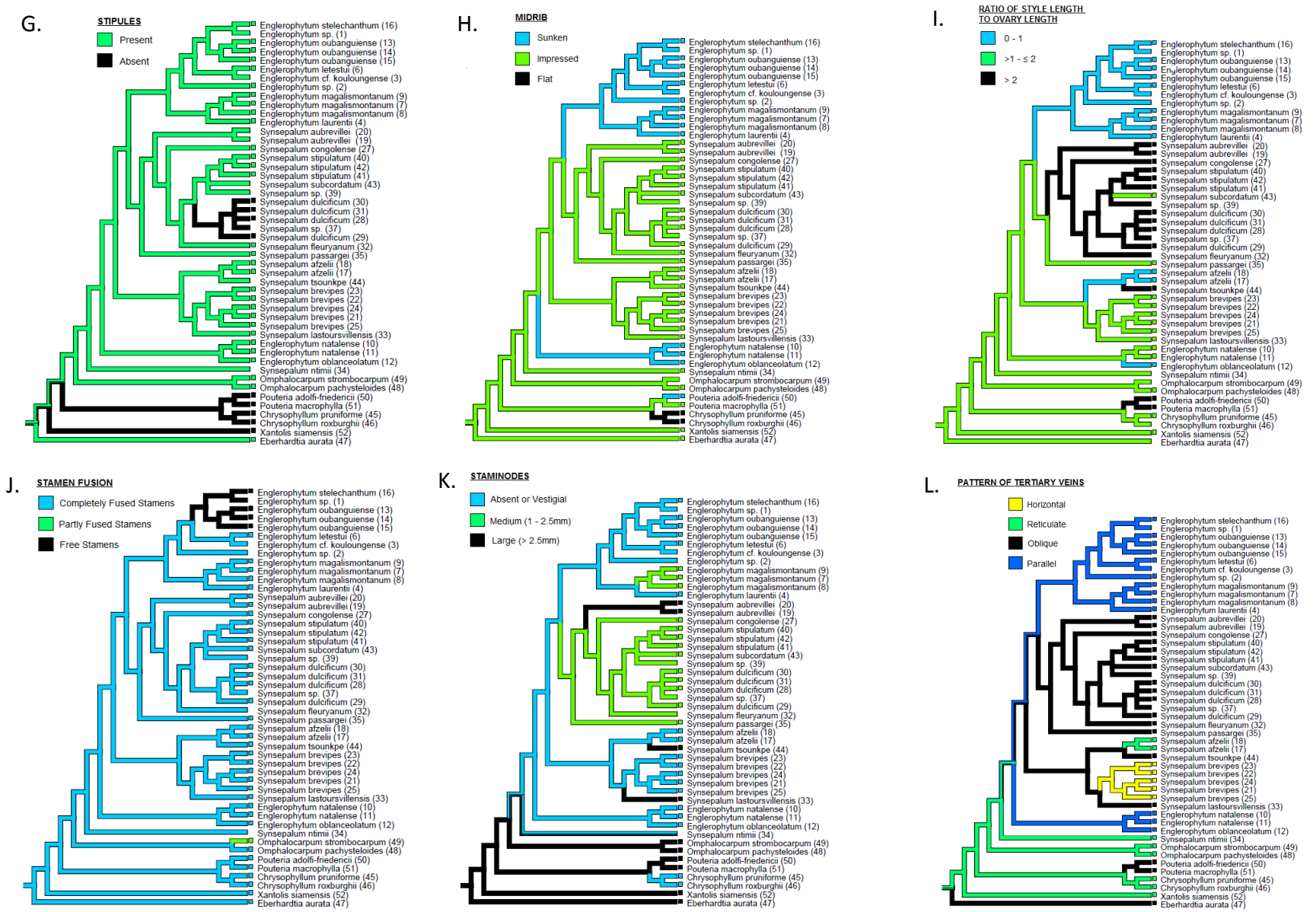


Figure 3.2 – A selection of morphological characters being mapped on one of the most parsimonious trees of the combined dataset. (G) Presence of Stipules (H) Shape of midrib (I) Ratio of Style Length to Ovary Length (J) Stamen Fusion (K) Staminodes (L) Pattern of tertiary Veins

4. DISCUSSION

4.1 The major lineages in the ES clade

One of the aims of this project was to determine whether the current morphology-based circumscription of *Synsepalum* and *Englerophytum* is consistent with the data from the molecular analysis.

In other words, the study was seeking to determine whether *Englerophytum* and *Synsepalum*, as currently delimited, formed two well-supported, morphology-compliant sub-clades within the ES clade, one consisting exclusively of *Synsepalum* and the other consisting exclusively of *Englerophytum*.

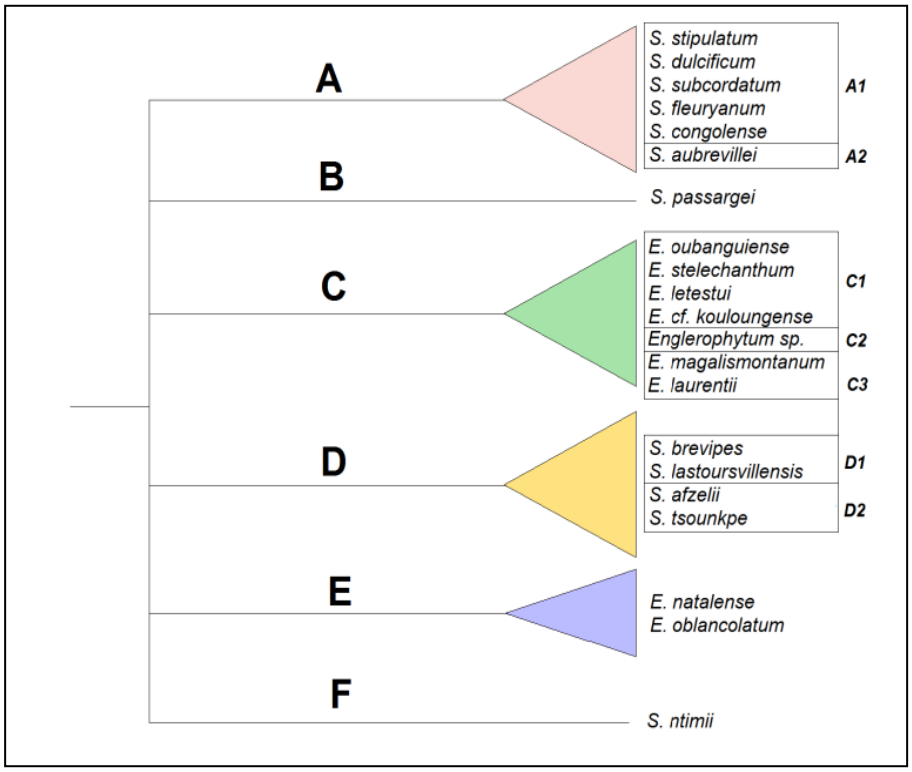


Figure 4.1 - Summary tree of the main lineages in the *Englerophytum-Synsepalum* Clade

Instead of the expected two clades, the tree obtained showed six lineages within the ES clade (Figure 4.1), four lineages containing species exclusively from *Synsepalum* and two lineages containing species exclusively from *Englerophytum.* This result can neither confirm nor reject the possibility that the current circumscription is correct since the polytomy at the base of the ES clade does not give any information about the inter-relationships between the major lineages.

However, despite not having two clearly defined lineages, the 6 lineages emerging from the present analysis still provide useful information on the structure of the ES clade. Each clade

will be discussed individually in the following sections and molecular and morphological evidence supporting each clade will be presented.

4.2 Clade A

Clade A consists of six species of *Synsepalum* which include *S. stipulatum, S. dulcificum, S. fleuryanum, S. subcordatum, S. congolense* and *S. aubrevillei*. These species range from small leaved short-petioled species such as *S. dulcificum,* to large leaved species such as *S. aubrevillei* and *S. subcordatum*. Although variable, this clade has several shared characters which distinguish it from other clades.

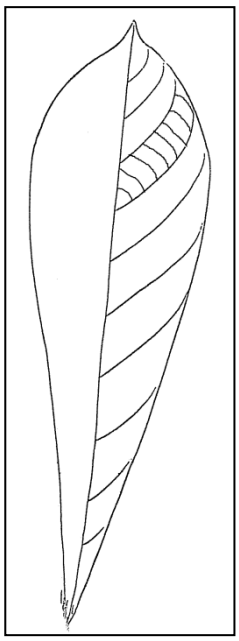


Figure 4.2 - Diagram showing Transverse-oblique tertiary venation

Species placed within Clade A all possess transverse-oblique tertiary venation (Figure 3.8L) where the tertiary veins form an oblique ladder-like pattern between two successive secondary veins (Figure 4.2). This character is not exclusive to clade A since it occurs in some other species (e.g. *S. lastoursvillensis* and *S. tsounkpe* in Clade D). However, venation pattern can be useful in combination with other characters in providing a unique morphological definition for Clade A.

Another shared character of members of this clade is the relatively long style length. All members of this clade have styles longer than 2.5mm. Some species such as *S. dulcificum* (7mm) and *S. aubrevillei* (7mm) have exceptionally long styles which are not found elsewhere in the ES clade. Consequently, since the ovary size does not change much in the ES clade (approx. 2mm), the style length: ovary length ratio of this clade always exceeds 2 (Fig. 3.8 I), with the exception of *S. subcordatum* whose style is the shortest (2.8mm).

Clade A also shares similar size and structure of staminodes. Representatives from Clade A always have prominent (>1mm) antisepalous staminodes with denticulate margins. Whenever staminodes were observed in other clades they were either small and vestigial (e.g. *S. afzelii, S. brevipes* in clade D), or large but lacking denticulate margins (E.g. *E. magalismontanum, S. lastoursvillensis*). Only *S. tsounkpe* (Clade D), had staminodes that were similar to those of representatives from Clade A.

The most significant shared character in this clade is the presence of fused sepals. All members of this clade have sepals which are fused for at least a third of their length. The fused sepals usually form a tight cup around the petal tube and are very difficult to tease apart without making deep incisions. This character is nearly exclusive to Clade A.

Besides Clade A, sepal fusion was only present in three other species in the analysis. These are *S. tsounkpe, E. stelechanthum* and *E. oblanceolatum*. Unfortunately, none of the floral material of the abovementioned species could be analysed first hand and literature data (Moore, 1907; Krause, 1914; Aubréville, 1959, 1961; Liben, 1989) was used to get the information for sepal fusion. The extent of sepal fusion for *E. oblanceolatum* and *E. stelecanthum* is difficult to determine from literature data since descriptions are not clear about this. In the protologue of *E. oblanceoatum* (Moore, 1907) the calyx is described as "connate below" (inferne connatis) whilst in that of *E. stelechanthum* (Krause, 1914) the calyx is described as "slightly connate at base" (basi breviter connata). Thus it is difficult to say whether the extent of fusion is slight enough to be easily distinguishable from the type of sepal fusion in Clade A.

On the other hand, the extent of fusion of sepals of *S. tsounkpe* is clearly stated in Aubréville's "Notulae Systematicae" (1961). *Synespalum tsounkpe* has sepals fused for about half their length; a similar type of fusion to that of the flowers in Clade A. However, despite its morphological similarities, *S. tsounkpe* seems to be phylogenetically closer to representatives of Clade D based on molecular data.

It is important to note that sepal fusion and some other synapomorphic characters mentioned here, were the basis upon which Aubréville (1961) made his generic circumscription of *Synsepalum*. Consequently, this clade is nearly identical to Aubréville's circumscription of *Synsepalum.* In fact, it contains species exclusively from Aubréville's *Synsepalum*. The only difference between this clade and Aubréville's *Synsepalum* is the placement of *S. tsounkpe.*

Morphologically this species fits very well into Clade A and as can be seen from Figures. 3.7 and 3.8, it is the only species in the phylogeny that possesses all the characters that define Clade A ie. a long style, fused sepals, dentate staminodes and an oblique tertiary venation pattern. Thus it is not surprising that Aubréville classified this species in *Synsepalum.*

However, in this analysis, this species appears in Clade D sister to *S. afzelii* with strong support.

In the light of this morphological evidence, there seems to be the need to revisit the placement of *S. tsounkpe* in Clade D. Unlike other species (e.g . *S. brevipes, S. dulcificum*) which were represented by more than one DNA sample in the analysis, *S. tsounkpe* was only represented by one sample. More samples of this species would be needed to confirm its placement.

In summary, Clade A in this analysis consists of six species (*S. stipulatum, S. dulcificum, S. fleuryanum, S. subcordatum, S. congolense* and *S. aubrevillei*) which share four consistent morphological characters i.e. an oblique tertiary venation pattern, a long style, fused sepals and dentate staminodes. This clade shares strong similarities with Aubréville's delimitation of *Synsepalum.*

4.3 Clade B

Clade B is only represented by one species, *S. passargei.* This species has unique floral characteristics not seen elsewhere in any of the other species analysed. A picture of the flower of *S. passargei* is shown in Fig. 4.3.

S. passargei is characterized by an extremely small petal tube (<1mm). The tube is so small that in some literature such as Aubréville (1961) the tube is overlooked and petals are termed "free". Due to the presence of a small tube, the species has a large "lobe length : tube length ratio" (lobes approx 15 times longer than tube – Fig 3.7A). Additionally, the petals of *S. passargei* are also unique in that they become strongly reflexed at maturity. They usually are so reflexed that the apices of the corolla lobes nearly touch the pedicel when fully reflexed.

The most prominent floral character in the flower of *S. passargei* is its ovary. *S. passargei* has a large ovary relative to the whole flower. The style is rather small and thus *S. passargei* has a "style length:ovary length" ratio below one (Fig. 3.8 I), which is rather unusual in the ES clade except for the *Englerophtyum* clade (Clade C).

The androecium of *S. passargei* also has unique characters. Unlike the petals, which are reflexed, the alternipetalous staminodes and antepetalous stamens are erect and immediately surround the massive ovary. Staminodes are linear, usually entire and approximately the same length as the petals. The stamens of this lineage have the smallest anthers of all the species included in the analysis. They measure below 1mm.

In this study the anther dehiscence of *S. passargei* was observed to be latrorse. This observation agrees with depictions of the *S. passargei* in Flora of East Tropical Africa (Hemsley, 1968) but disagrees with other literature (Kupicha, 1983; Swenson & Anderberg, 2005), where anther dehiscence is quoted as extrorse. This property is therefore still debatable, and further material of the species, preferably fresh, should be analysed to make more accurate conclusions.

Although Lineage B was only represented by a single species in this analysis, there are several other morphologically similar species to *S. passargei* which probably belong to this lineage and deserve further DNA work to verify this. These species include S. *revolutum, S.muelleri* and *S. brenanii*. The flower of one of these species, *S. revolutum*, is shown alongside the flower of *S. passargei* in Fig. 4.3 to show the extent of floral similarity existing between some of these species.

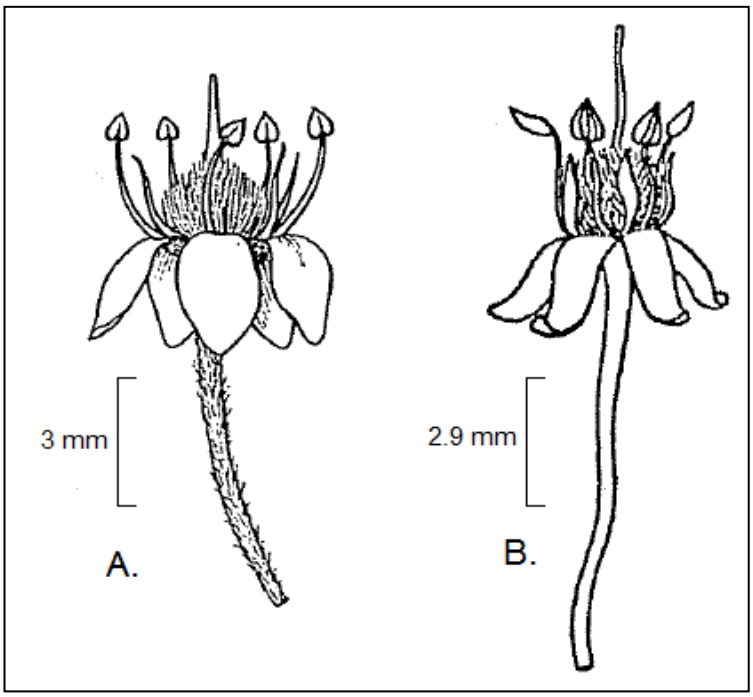


Figure 4.3 - Flowers of (A) *Synsepalum passargei* (Aubréville, 1964) and (B) *Synsepalum revolutum* (Aubréville, 1959). Both flowers have reflexed petals, small anthers, erect stamens and staminodes and a large ovary

These morphologically similar species have been previously grouped together (Aubréville & Pellegrin, 1934; Meeuse, 1960; Aubréville, 1964; Kupicha, 1983) with *S.passargei* into a single genus called *Vincentella* (recently synonymized into *Synsepalum* by Pennington (1991)). This genus was defined by the characteristics in Table 4.1.

Most of the characteristics of *Vincentella* apply to *S. passargei* in Lineage B. It would be very interesting to see whether species, previously classified under *Vincentella,* also fall into Lineage B. Two other species, *Vincentella ogouensis* and *Vincentella ovatostipulata*, which were also classified under *Vincentella by* Aubréville (1965) also deserve further study to check whether they belong to this lineage.

Table 4.1 – Characteristics used to define *Vincentella* (from Meeuse 1960)

a. Slender pedicels	b. Ovary large, ovoid and villous
c. Sepals small, free nearly to the base, later patent or reflexed	d. Staminodes alternipetalous, narrowly linear, erect and as long as corolla lobes
e. Corolla tube very short; the lobes many times longer	f. Filaments erect, several times longer than the oblong-sagittate, minutely apiculate anthers
g. Corolla lobes narrow and strongly reflexed	

In summary, lineage B consists of a single species *S. passargei* and is characterized by (i) a reflexed corolla with a very short corolla tube (ii) a large ovary (iii) erect stamens with very small anthers (possibly latrorse) and (iv) erect, entire alternipetalous staminodes. Characteristcs of this genus are similar to those of a former genus named *Vincentella*.

4.4 Clades C and E

These two lineages consist solely of species currently belonging to the genus *Englerophytum*. They will be treated together in the discussion because they share several common morphological characters especially in their vegetative parts. It is important to note that although these two clades are collectively easily distinguishable from the other four clades, they are very difficult to distinguish from each other.

4.4.1 Vegetative Similarities between Clades C and E

One of the reasons that Pennington (1991) grouped all representatives from Clade C and E into *Englerophytum,* was because of their characteristic leaf facies. A summary of the main leaf characters of these clades is shown in Figure 4.4. Their leaves differ from those of *Sysepalum* (Clades A, B, D, F) in 3 main aspects (i) structure of petiole and main vein, (ii) pattern of venation and (iii) leaf indumentum.

Leaves of *Englerophytum* have a sunken midrib which forms a channel along the lamina. This consequently affects the structure of the petiole as well as the apex of the leaf. The channeled petiole folds on itself forming a closed hollow groove whilst at the apex, the sunken main vein always exceeds the tip by a slight fraction forming a small mucronate tip. This structure strongly contrasts with that of *Synsepalum* whose leaf has a petiole with an open groove, an impressed midrib and lacks a mucronate tip.

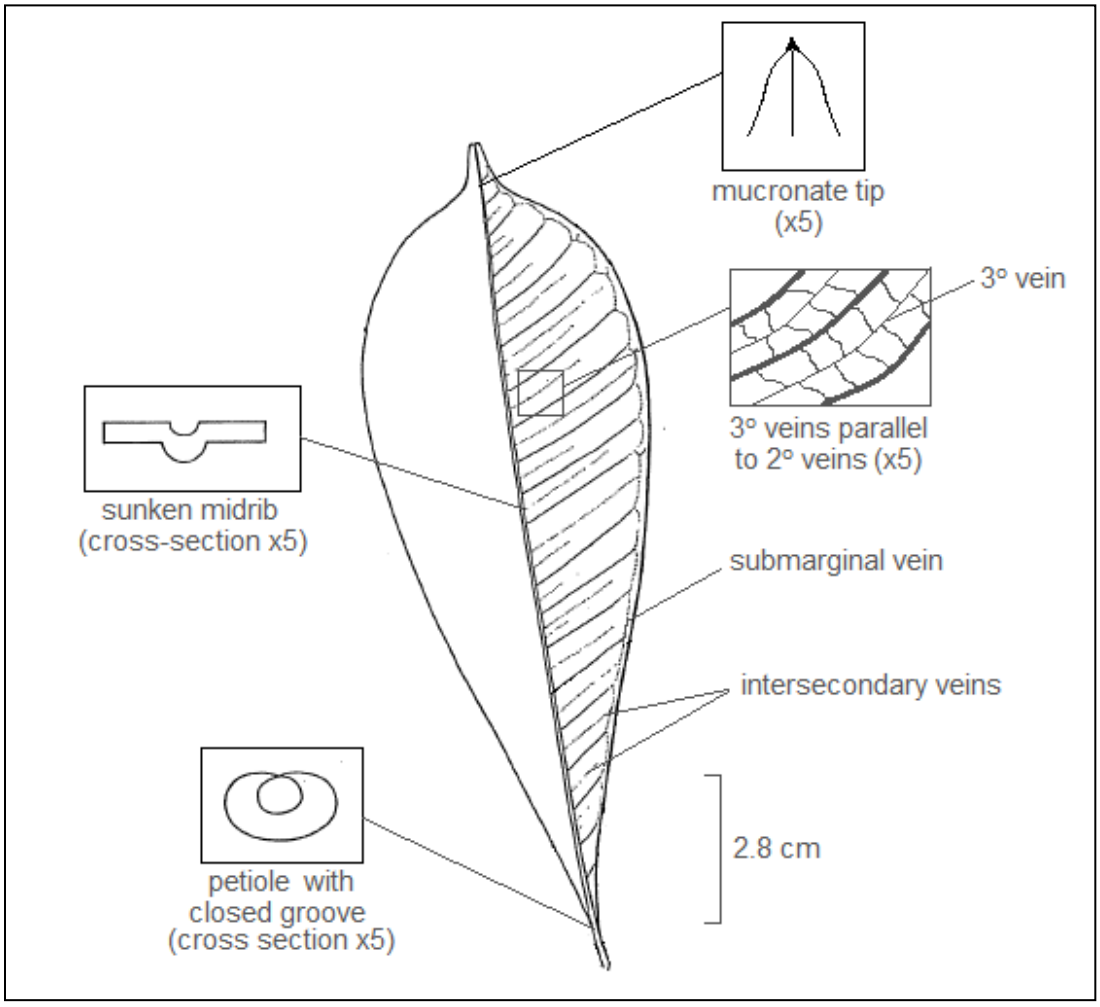


Figure 4.4 - Main features of the leaf of representatives from Clades C and E [assembled from pictures in Aubréville (1961) and Ng (1972)]

The pattern of venation of *Englerophytum* (Clades C and E) is also very characteristic. It consists of a brochiododromous (looping) pattern of venation with very closely parallel secondary veins. Due to the close proximity of the looping veins to the leaf margin, the loops tend to form a submarginal vein. Additionally, all species in Clades C and E have parallel

intersecondary veins. These are veins situated between the secondaries which are initially of equal thickness to the secondaries but become thinner as they approach the leaf margin. The intersecondaries and secondaries are also parallel to the tertiary veins, a pattern which is unique in the ES clade.

The final character shared by these clades is the velvety brown indumentum on the leaf undersurface. This character is not exclusive to Clades C and E and is also present in other species such as *S.brevipes* and *S. aubrevillei*. However, the coloration of the indumentum in *Englerophytum* is always orange brown whilst the indumentum in non-*Englerophytum* species is usually cream coloured.

4.4.2 Floral Similarities between Clades C and E

Unlike vegetative characters, whose characteristics clearly isolate Clades C and E from the rest of the ES clade, floral characters are more ubiquitous within the ES clade. All sub-clades in the ES clade usually have 5-merous flowers with an irregular presence of staminodes and stamens attached at the top of the corolla throat. Thus it is more difficult to isolate clades C and E from others using floral characters.

A unique feature that was used by Pennington (1991) to separate *Englerophytum* (Clades C and E) from *Synsepalum* was the tendency of fusion of filaments into a cone-like structure enclosing the pistil. However, during this study it became evident that this character is more suitable to define a specific sub-group within Clade C rather than to define both clades of *Englerophytum*. As shown in Fig. 3.8 J, this character seems to have evolved within Clade C and is only found in *E. stelechanthum* and *E. oubanguiense*. Nevertheless filament fusion is still a very useful character for identification of some species since it is a rare occurrence in the ES clade. Other species from *Englerophytum* (e.g. *E. congolense* and *E. iturense*), which were not included in this study, also show filament fusion and further work should also focus on sampling these species to find out whether this character evolved once or more than once in the ES clade.

4.4.3 Vegetative Differences between Clades C and E

Not much difference exists between vegetative characters of Clades C and E, however there are two characters which although not perfectly consistent throughout, might still prove

helpful in some cases. These include (i) the number of secondary veins and (ii) leaf shape. Clade E usually has less than 28 veins whilst Clade C usually has 28 or more (Fig 3.7 E). Leaf shape also shows some variation. Leaf shape in Clade C is usually oblanceolate, whilst that in Clade E is usually obovate, however some overlap is present.

4.4.4 Floral Differences between Clades C and E

Floral distinguishing characters are also lacking amongst these clades. The most consistent characters that can help make the distinction include (i) pedicel length (Fig. 3.7B) and (ii) the ratio of style length to ovary length (Fig 3.8 I).

Representatives of Clade E usually possess a very short pedicel appearing sessile or sub-sessile. On the other hand representatives from Clade C have longer pedicels which raise the flower above the flowering branch. Besides this difference, the ratio of style length to ovary length also varies. Members of clade C usually have styles which are shorter than the ovary, but representatives of Clade E have styles which are longer.

In literature (Aubréville & Pellegrin, 1958) *E. oblanceolatum* and *E. natalense* (Clade E) were also distinguished by having petal lobes being shorter than the petal tube. However, in the floral material analysed this character was not immediately evident and more material is required to confirm this.

4.4.5 Classification

It is interesting to note that *E. natalense* and *E. oblanceolatum* (Clade E) have been grouped together in the past through the use of morphological characters. Aubréville and Pellegrin (1958) erected the genus *Boivinella* (later changed to *Neoboivinella* in 1959 due to being a later homonym of a genus of grasses), which contained both species from Clade E together with *E. magalismontanum,* which was later removed from the genus. This genus was distinguished by having (i) lobes smaller than the petal tube (ii) shortly petiolate leaves (iii) short filaments (iv) absent or vestigial staminodia and (v) a wide ventrifixed hilium. However, in a later publication, (Aubréville, 1961), Aubréville changed the circumscription of this genus leaving only *E. natalense* (then *Neoboivinella natalense*) in the genus. *Neoboivinella* was synonymised with *Englerophytum* by Pennington (1991)

E. oblanceolatum and *E. natalense* (Clade E) were also grouped together by Heine & Hemsley (1960) into a genus erected by De Wildeman (1919) called *Bequaertiodendron*. Heine and Hemsley distinguished these species on the basis of seed characters; namely the absence of endosperm and the presence of thick and fleshy plano-convex cotyledons. *Bequaertiodendron* is now also a synonym of *Englerophytum.*

In summary Clades C and E can be collectively distinguished from the rest of the ES clade by their distinctive leaf facies (Fig. 4.3). Distinction between clades C and E is less obvious and can be helped by the following characters: (i) leaf shape (ii) number of secondary veins (iii) pedicel length and (iv) ratio of style length to ovary length.

4.5 Clade D

Clade D is composed here of four species which include *Synepalum lastoursvillenis, Synsepalum afzelii, Synsepalum brevipes* and *Synsepalum tsounkpe*. With the exception of *S. tsounkpe* which has a number of morphological similarities with Clade A, members of this clade share a number of morphological characters.

Members of this clade were observed to have inconspicuous tertiary venation. The tertiary veins of these species were so fine, that in order to determine their pattern of tertiary venation, the leaves had to be held against the light. This character contrasts with leaves from Clade A whose tertiaries were prominent and easily identifiable.

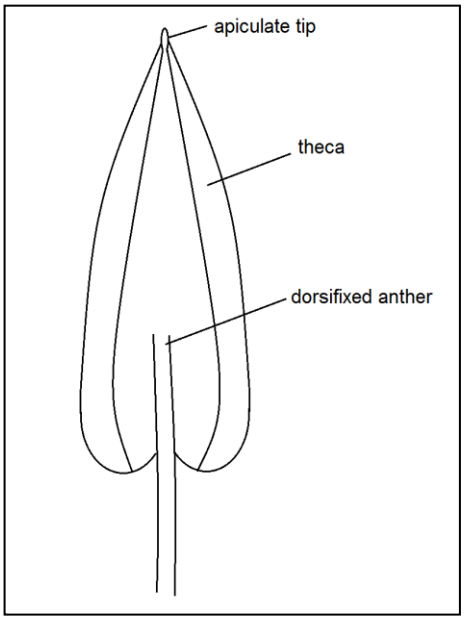


Figure 4.5 - Structure of narrowly sagittate anthers in Clade D

Members of this clade also have free sepals. This character is not exclusive to Clade D, however representatives from Clade D are most morphologically similar to members of Clade A and the fusion of the calyx is an important character that can help distinguish this clade from Clade A.

A third character that is shared by representatives of this clade is the presence of staminodes with entire margins. Although within this clade *S. afzelii* and *S. brevipes* do not

always possess staminodes, whenever they are present, these are rudimentary and have entire margins. *S. lastoursvillensis* on the other hand has the largest staminodes of the clade, which are also entire. *S. tsounkpe* is an exception and has dentate staminodes.
The shape of the anthers of *S. afzelii* and *S. brevipes*, whose flower material was available, showed a similar narrowly sagittate anther shape, shown in Figure 4.5. However, only line drawings of the other species in the clade were available and it was unclear whether they shared the same anther structure. Thus further analysis is required to check whether all representatives of this clade possess narrowly sagittate anthers.

In summary representatives of Clade D possess the following synapomorphies (i) inconspicuous tertiary venation (ii) free sepals (iii) entire staminodes and possibly (iv) narrowly sagittate anthers.

4.6 Clade F

In this analysis, Clade F was only composed of a single species, *Synsepalum ntimii*. Hawthorne (pers. comm. 2013) has suggested that this represents an undescribed species. A small description of the species can be found in "Woody Plants of Western African Forests" (Hawthorne & Jongkind, 2006). William Hawthorne has kindly provided a specimen of *S. ntimii* for inclusion in this project.

The leaf of *S. ntimii* was immediately distinguishable from the rest of the ES clade. *Synsepalum ntimii* is the only other species, besides ones from Clades C and E, to have brochiodromous venation (the rest have eucamptodromous venation). However, the venation patterns of *S. ntimii* are still distinguishable from the brochiododromous venation of Clades C and E.

Unlike clades C and E, *S. ntimii* has an impressed midrib and tertiary veins which are reticulate rather than parallel to the secondaries. Additionally, *S. ntimii* also has fewer secondary veins than species in Clades E and C. *Synsepalum ntimii* has an average number of 15 secondary veins, whilst Clades C and E always have more than 15.

Another distinguishing feature of *S. ntimii* is the leaf underside. The lamina of *S. ntimii* is glabrous and slightly shiny. This clearly distinguishes it from *Englerophytum* in Clades C and

E which have a velvety indumentum beneath. The lack of hairs on the midrib is also very distinctive of this lineage since usually the midrib of individuals in the ES clade possesses hairs (Fig. 3.7C).

It is important to note that this species was discovered before Pennington's comprehensive account of the genera of the Sapotaceae, when representatives of *Synsepalum* were still spread amongst a number of different genera. At that time, Hawthorne was unsure as to which of Aubréville's genera would best accommodate this morphologically distinct species (pers. comm. 2013). There did not seem to be an appropriate genus that could host the species.

In fact the morphological uniqueness of *S. ntimii* is also reflected in its DNA. Out of all the examined species, *S. ntimii* had the longest branch length in the parsimony tree (Fig. 3.3). Hawthorne only made the decision to put *S. ntimii* into *Synsepalum* after speaking with Pennington who advised him to recognise a very broad generic circumscription of *Synsepalum.*

Further study may reveal that the distinct lineages within the ES clade deserve generic status, in which case, *S. ntimii*, would represent an undescribed genus from tropical West Africa.

In summary, *S ntimii* can be distinguished from the other clades by (i) its brochiododromous venation with reticulate secondary venation (ii) the number of secondary veins (iii) and its glaborous, slightly shiny leaf underside.

4.7 Major patterns in the ES phylogeny

A summary table of all the clades and their diagnostic characteristics is given in Table 4.2. As one can observe, several lineages in this analysis possess a number of shared characters characters which enable distinction to be made amongst lineages.

It is still early to make any taxonomic decisions on the ES clade since more evidence needs to be collected. Future efforts should be focused on resolving the polytomy in the ES clade. This may provide further insight as to whether the current circumscription of the genera is correct or whether the taxonomy needs to be revised. Nevertheless, the six well-supported lineages in

Table 4.2 – Summary table of shared characters of each lineage			
Clade	**Constituent Taxa**	**Shared Characters**	
A	*S. dulcificum, S. congolense, S. stipulatum, S. fleuryanum, S. aubrevillei, S. subcordatum*	(i) Oblique tertiary venation (ii) Fused calyx (iii) Long style (iv) Denticulate staminodes	
B	*S. passargei*	(i) Large ovary (ii) Strongly reflexed petal lobes (iii) Very short corolla tube (iv) Anthers < 1mm (v) erect linear staminodes	
C	*E. oubanguiense, E. stelechanthum, E. magalismontanum, E. letestui,, E. laurentii*	(i) brochiododromous venation with parallel tertiary veins (ii) sunken midrib (iii) petiole with closed groove (iv) mucronate leaf tip (v) velvety brown leaf indumentum (vi) submarginal vein (vii) intersecondary veins	(i) More than 28 secondary veins (ii) style shorter than ovary (iii) pedicellate flowers (iv) usually oblanceolate leaves
E	*E. natalense, E. oblanceolatum*		(i) 17-27 secondary veins (ii) style longer than ovary (iii) sessile/sub-sessile (iv) usually obovate leaves
D	*S. brevipes, S. afzelii, S. lastoursvillensis, S. tsounkpe*	(i) inconspicuous tertiary venation (ii) calyx free (iii) entire staminodes [(iv) narrowly sagittate anthers]	
F	*S. ntimii*	(i) brochiododromous venation with reticulate tertiary venation(ii) 15 or less secondary veins (iii) glaborous leaf underside	

the ES clade are still rather distinct entities and if further study of this clade reveals further synapomorphies, considerations should be made to resurrect some genera (e.g. *Vicentella, Neoboivinella*) which in the past had very similar circumscriptions to some clades in the phylogeny.

5. CONCLUSIONS & FURTHER WORK

5.1 Conclusions

This study has provided further evidence of the monophyly of the Chrysophylloideae and the *Englerophytum-Synsepalum* clade and helped to better resolve the relationships within the ES clade. The clade is composed of six lineages, many of which are distinguishable by means of vegetative and floral characters. Calyx fusion, leaf venation patterns, style length and staminode structure were amongst the most useful characters for distinguishing one lineage from another. Leaf size and shape, sepal length and ovary length were amongst the least useful to make distinctions.

The ITS region proved to be much more informative than the trnH-psbA region and consequently the phylogeny obtained from the ITS region was better resolved than that obtained from trnH-psbA. The results, neither contradict nor corroborate the present generic circumscription of *Englerophytum* and *Synsepalum* (Pennington, 1991), since the deep-seated polytomy at the base of the ES clade does not provide any information on the relationship between any of the main *Englerophytum* and *Synsepalum* clades. However, there seems to be close similarities between some sub-clades in the ES clade and formerly recognised genera (e.g *Vincentella, Neoboiviella*). Further work on this clade might help provide the necessary evidence to make educated taxonomic decisions on representatives of the ES clade.

5.2 Further Work

In order to gain a better understanding of the *Englerophytum-Synsepalum* clade, further work should be focused on the following points:

a. Revision: There is an immediate need of a revision of the taxa within these two genera. This can help establish better species concepts and reduce misidentifications in herbarium specimens.
b. Wider species sample: Further DNA work should focus on including more species in the phylogeny so as to allow a better understanding of the size and constituent taxa of each of the sub-clades within the ES clade.

c. DNA Evidence: More DNA data, especially from chloroplast regions, should be gathered in order to attempt to resolve the deep-seated polytomy at the base of the ES clade.
d. Examination of more morphological characters: Characters which were not considered in this project (e.g. seed, fruit and wood characters) should be explored in order to help characterize fully the sub-clades within the ES clade.
e. Collections: More collections of species from this clade are required, especially specimens with well preserved fruit and flower material. This can allow a more thorough analysis of each species and reduce reliance on literature data.
f. Biogeography and evolution: A dated phylogeny should also be carried out in order to better understand the age, origin and evolution of the representatives in the ES clade.

6. REFERENCES

ANGIOSPERM PHYLOGENY GROUP. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of The Linnean Society* 161(2): 105–121.

ARMSTRONG, K. E. (2010). *Systematics and biogeography of the pantropical genus* Manilkara *(Sapotaceae)*. PhD thesis, Univerity of Edinburgh.

AUBRÉVILLE, A. (1959). *La Flore Forestière de la Côte d'Ivoire*. Nogent-sur-Marne, Seine: Centre Technique Forestier Tropical.

AUBRÉVILLE, A. (1961). Notes sur les Sapoteés de l'Afrique équatoriale. *Notulae Systematicae (Paris)* 16: 260–268.

AUBRÉVILLE, A. (1964a). Sapotacées: taxonomie et phytogéographie. *Adansonia* 1: 1–157.

AUBRÉVILLE, A. (1964b). *Flore Du Cameroun. Monograph 2. Sapotaceés*. Paris: Muséum national d'histoire naturelle, Laboratoire de phanérogamie.

AUBRÉVILLE, A. (1965). *Flore Du Gabon. Monograph 1. Sapotaceés*. Paris: Muséum national d'histoire naturelle, Laboratoire de phanérogamie.

AUBRÉVILLE, A. (1971). Essais de geophyletique des Sapotacees: 2. *Adansonia* 3: 425–436.

AUBRÉVILLE, A. & PELLEGRIN, F. (1934). De quelques Sapotacées de la Côte d'Ivoire. *Bulletin de la Société botanique de France* 81: 792.

AUBRÉVILLE, A. & PELLEGRIN, F. (1957). Sapotacées nouvelles d'Afrique tropicale. *Bulletin de la Société botanique de France* 104: 276–281.

AUBRÉVILLE, A. & PELLEGRIN, F. (1958). Réhabilitation de deux genres de Sapotacées. *Bulletin de la Société botanique de France* 105: 35.

AUBRÉVILLE, A. & PELLEGRIN, F. (1959). Rectification au sujet de Sapotacées africaines. *Bulletin de la Société botanique de France* 106: 22.

AYENSU, E. S. (1972). Morphology and Anatomy of *Synsepalum dulcificum* (Sapotaceae). *Botanical Journal of The Linnean Society* 65(2): 179–187.

BAEHNI, C. (1965). Mémoires sur les Sapotacées 3. Inventaire des genres. *Boissiera* 11: 1–262.

BAILLON, M. H. (1890). Observations sur les Sapotacées de la Nouvelle-Calédonie. *Bulletin Mensuel de la Société Linnéenne de Paris* 2: 881–888.

BALDWIN, B. G. & MARKOS, S. (1998). Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: Congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10(3): 449–463.

BARTISH, I. V., ANTONELLI, A., RICHARDSON, J. E. & SWENSON, U. (2011). Vicariance or long-distance dispersal: historical biogeography of the pantropical subfamily Chrysophylloideae (Sapotaceae). *Journal of Biogeography* 38(1): 177–190.

BARTISH, I. V., SWENSON, U., MUNZINGER, J. & ANDERBERG, A. A. (2005). Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): Molecular evidence for generic polyphyly and repeated dispersal. *American Journal of Botany* 92(4): 667–673.

BENSON, D. A., KARSCH-MIZRACHI, I., CLARK, K., LIPMAN, D. J., OSTELL, J. & SAYERS, E. W. (2011). GenBank. *Nucleic Acid Research* 40(Database Issue): D48–D53.

BOLZA, E. & KEATING, W. G. (1972). *African Timbers: the properties, uses and characteristics of 700 species*. Melbourne: Division of Buildings Research, CSIRO.

BRAHMS ©. (2013). *Botanical Research And Herbarium Management Software.* Version 7.3 Build 1. Oxford: Department of Plant Sciences, University of Oxford.

BURKILL, H. M. (2000). *The Useful Plants of West Tropical Africa. Volume 5. 2nd edition.* London: Royal Botanical Gardens, Kew.

CHEVALIER, A. (1943). A propos de la nomenclature de quelqes Sapotaceés africaines. *Revue de Botanique Appliquée et d'Agriculture Tropicale* 23: 282–294.

DANIELL, W. F. (1852). On the *Synsepalum dulcificum* De Cand.; or, miraculous berry of Western, Africa. *Pharmaceutical Journal and Transactions* 11: 445–448.

DARRIBA, D., TABOADA, G. L., DOALLO, R. & POSADA, D. (2012). jModelTest 2.1.4 : more models, new heuristics and parallel computing. *Nature Methods* 9(8): 772.

DE CANDOLLE, A. (1844). *Prodromus Systematis Naturalis Regni Vegetabilis*. Paris: Sumptibus Sociorum Treuttel et Würtz.

DE WILDEMAN, E. A. J. (1914). Neuen Arten aus Zentral-Afrika (Belgisch-Kongo). *Repertorium Specierum Novarum Regni Vegetabilis* 13: 377.

DE WILDEMAN, E. A. J. (1919). Sur quelques espéces congolaises de la famille des Sapotacées. *Revue de Zoologie et de Botanique Africaines* 7(Suppl): 21.

DUBARD, M. (1911). Descriptions de quelques espéces de Lucumées africaines d'apres les documents de L. Pierre. *Notulae Systematicae (Paris)* 2: 89–91.

DUBARD, M. (1912). Les Sapotacées du groupe des Sideroxylinées. *Annales de l'Institut Botanico- Géologique Colonial de Marseille, Série 2* 10: 1–90.

ENGLER, A. (1904). Sapotaceae africanae. *Monographieen Afrikanisher Pflanzen-Familien und Gattungen* 8: 32–33.

FELINER, G. N. & ROSSELLO, J. A. (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* 44(2): 911–919.

FILATOV, D. A. (2009). Processing and population genetic analysis of multigenic datasets with ProSeq3 software. *Bioinformatics* 25: 3189–3190.

GENE CODES CORPORATION. (2013). Sequencher ® - Sequence analysis software. Version 5.1. Michigan: Ann Arbor.

GOVAERTS, R., FRODIN, D. G. & PENNINGTON, T. D. (2001). *World Checklist and Bibliography of Sapotaceae*. London: Royal Botanic Gardens, Kew.

HALL, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

HARTOG, M. M. (1878). On the floral structure and affinities of Sapotaceae. *Journal of Botany* 16: 65–72.

HAWTHORNE, W. D. & JONGKIND, C. C. H. (2006). *Woody Plants of Western African Forests: A guide to the forest trees, shrubs and lianes from Senegal to Ghana*. London: Royal Botanic Gardens, Kew.

HEINE, H. & HEMSLEY, J. H. (1960). Notes on African Sapotaceae: II: The Genus *Bequaertiodendron* De Wild. *Kew Bulletin* 14(2): 304–309.

HEMSLEY, J. H. (1968). Sapotaceae. In: MILNE-REDHEAD, E., POLHILL, R. M. & JEFFRIC, C. (eds), *Flora of Tropical East Africa*, pp. 1–79. London: Royal Botanical Gardens, Kew.

KEAY, R. W. J. (1989). *Trees of Nigeria*. Oxford: Claredon Press.

KRAUSE, K. (1914). *Englerophytum*, eine neue afrikanische Gattung der Sapotaceen. *Botanischer Jahrbücher für Systematik* 50(Suppl): 343–348.

KUMAR, S., SKJÆVELAND, Å., ORR, R. J., ENGER, P., RUDEN, T., MEVIK, B.-H., BURKI, F., BOTNEN, A. & SHALCHIAN-TABRIZI., K. (2009). AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* 10: 357.

KUPICHA, F. K. (1983). Sapotaceae. In: LAUNERT, E., BRENAN, J. P. M., MENDES, E. J. & WILD, H. (eds), *Flora Zambesaica. Volume 7.* London: The Flora Zambesiaca Managing Committee.

LAM, H. J. (1939). On the system of the Sapotaceae, with some remarks on taxonomical methods. *Recueil des Travaux Botaniques Neerlandais* 36: 509–525.

LECOMTE, M. H. (1928). Deux Sapotacées nouvelles de Madagascar et d'Afrique. *Bulletin of Moss Flora of Arctic North America and Greenland* 34: 355–356.

LIBEN, L. (1989). La véritable identité des genres et espèces confondus sous le nom de "*Bequaertiodendron magalismontanum*" (Sond.) Heine & Hemsley (Sapotaceae) en Afrique centrale et occidentale. *Bulletin du Jardin botanique national de Belgique* 59: 151–169.

LOUPPE, D., OTENG-AMOAKO, A. A. & BRINK, M. (eds.). (2008). *Plant Resources of Tropical Africa 7(1). Timbers 1.* Netherlands: PROTA Foundation and Backhuys Publishers.

MADDISON, W. P. & MADDISON, D. R. (2011). Mesquite: a modular system for evolutionary analysis. Version 2.75. U.S.A: University of British Columbia/ Oregon State University.

MEEUSE, A. D. J. (1960). Notes on the Sapotaceae of Southern Africa. *Bothalia* 71(2): 317–379.

MÖLLER, M. & CRONK, Q. C. B. (1997). Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. *American Journal of Botany* 84(7): 956–965.

MOORE, S. (1907). Alabastra Diversa - Part XIV - New or little known African gamopetalae. *Journal of Botany, British and Foreign.* 45: 47.

NG, F. S. P. (1972). Sapotaceae. In: WHITMORE, T. C. (ed), *Tree Flora of Malaya - A Manual for Foresters. Volume 1.*, Kuala Lump., pp. 388–441. Longman.

PENNINGTON, T. D. (1991). *The Genera of the Sapotaceae*. London: Royal Botanical Gardens, Kew/ New York: New York Botanical Garden.

PIERRE, J. B. L. (1890). *Notes Botaniques Sapotacées*. Paris: Libraire des Sciences Naturelles.

RAMBAUT, A. & DRUMMOND, A. J. (2009). Tracer v.1.5. http://beast.bio.ed.ac.uk/Tracer (accessed August 2013).

RONQUIST, F. & HUELSENBECK, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

SCHÖNENBERGER, J., ANDERBERG, A. A. & SYSTMA, K. J. (2005). Molecular Phylogenetics and Patterns of Floral Evolution in the Ericales. *International Journal of Plant Sciences* 166(2): 265–288.

SHAW, J., LICKEY, E. B., BECK, J. T., FARMER, S. B., LIU, W., SIRIPUN, K. C., WINDER, C. T., SCHILLING, E. E. & SMALL, R. L. (2005). The Tortoise and the Hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92(1): 142–166.

SIMMONS, M. P. & OCHOTERENA, H. (2000). Gaps as characters in sequence-based pylogenetic analyses. *Systematic Biology* 49(2): 369–381.

SMEDMARK, J. E. E. & ANDERBERG, A. A. (2007). Boreotropical Migration Explains Hybridization between Geographically Distant Lineages in the Pantropical Clade Sideroxyleae (Sapotaceae). *American Journal of Botany* 94(9): 1491–1505.

SMEDMARK, J. E. E., SWENSON, U. & ANDERBERG, A. A. (2006). Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Molecular Phylogenetics and Evolution* 39(3): 706–21.

SNEATH, P. H. A. (1962). The construction of taxonomic groups. In: *Microbial Classification, 12th Symposium of the Society for General Microbiology*, p. 483. Cambridge: Cambridge University Press.

SWENSON, U. & ANDERBERG, A. A. (2005). Phylogeny, character evolution, and classification of Sapotaceae. *Cladistics* 21: 101–130.

SWENSON, U., RICHARDSON, J. E. & BARTISH, I. V. (2008). Multi-gene phylogeny of the pantropical subfamily Chrysophylloideae (Sapotaceae): evidence of generic polyphyly and extensive morphological homoplasy. *Cladistics* 24: 1006–1031.

SWOFFORD, D. L. (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods).* Version 4. Sunderland, Massachusetts: Sinauer Associates.

THE PLANT LIST. (2010). *Synsepalum* and *Englerophytum*. http://www.theplantlist.org/ (accessed August 2013).

THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through through dequence weighting, position specific gap panalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

WALTER, K. S. & O'NEAL, M. J. (2013). BG-BASE, Collections Management Software.

WATT, J. M. & BREYER-BRANDWIJK, M. G. (1962). *The Medicinal and Poisonous Plants of Southern and Eastern Africa*. Edinburgh: E. & S. Livingstone.

WERNHAM, H. F. (1917). New Tropical African Rubiaceae. *Journal of Botany* 55: 82.

WHITE, F. (1962). Geographic variation and speciation in Africa with particular reference to *Diospyros* . *Systematics Association Publication* 4: 71.

WHITE, T. J., BURNS, T., LEE, S. & TAYLOR, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: INNIS, M., GELFAND, D., SNINSKY, J. & WHITE, T. (eds), *PCR protocols: a guide to methods and applications*, pp. 315–322. San Diego: California Academic Press.

WSCP. (2013). World Checklist of the Sapotaceae. http://apps.kew.org/wcsp/ (accessed August 2013).

APPENDIX 1

Table A1.1 – Species of *Synsepalum* (adapted from the Plant List)

Species Name	Status	Species sequenced
Synsepalum afzelii (Engl.) T.D.Penn.	Accepted	✓
Synsepalum aubrevillei (Pellegr.) Aubrév. & Pellegr.	Accepted	✓
Synsepalum batesii (A.Chev.) Aubrév. & Pellegr.	Accepted	
Synsepalum bequaertii De Wild.	Accepted	
Synsepalum brenanii (Heine) T.D.Penn.	Accepted	
Synsepalum brevipes (Baker) T.D.Penn.	Accepted	✓
Synsepalum buluensis (Greves) ined.	Accepted	
Synsepalum carrieanum (Dubard) Pierre ex ined.	Accepted	
Synsepalum cerasiferum (Welw.) T.D.Penn.	Accepted	
Synsepalum congolense Lecomte	Accepted	✓
Synsepalum dulcificum (Schumach. & Thonn.) Daniell	Accepted	✓
Synsepalum fleuryanum A.Chev.	Accepted	✓
Synsepalum gabonense (Aubrév. & Pellegr.) T.D.Penn.	Accepted	
Synsepalum kassneri (Engl.) T.D.Penn.	Accepted	
Synsepalum lastoursvillensis (Aubrév. & Pellegr.) ined.	Accepted	✓
Synsepalum laurentii (De Wild.) D.J.Harris	Accepted	
Synsepalum le-testui Aubrév. & Pellegr.	Accepted	
Synsepalum letouzei Aubrév.	Accepted	
Synsepalum msolo (Engl.) T.D.Penn.	Accepted	
Synsepalum muelleri (Kupicha) T.D.Penn.	Accepted	
Synsepalum nyangense (Pellegr.) McPhersen & J.T.White	Accepted	
Synsepalum ogouense (Aubrév. & Pellegr.) ined.	Accepted	
Synsepalum ovatostipulatum (De Wild.) ined.	Accepted	
Synsepalum oyemense (Aubrév. & Pellegr.) ined.	Accepted	
Synsepalum passargei (Engl.) T.D.Penn.	Accepted	✓
Synsepalum pobeguinianum (Dubard) Aké Assi & L.Gaut.	Accepted	
Synsepalum revolutum (Baker) T.D.Penn.	Accepted	
Synsepalum seretii (De Wild.) T.D.Penn.	Accepted	
Synsepalum stipulatum (Radlk.) Engl.	Accepted	✓
Synsepalum subcordatum De Wild.	Accepted	✓
Synsepalum subverticillatum (E.A.Bruce) T.D.Penn.	Accepted	
Synsepalum tomentosum (Aubrév. & Pellegr.) ined.	Accepted	
Synsepalum tsounkpe Aubrév. & Pellegr.	Accepted	✓
Synsepalum ulugurense (Engl.) Engl.	Accepted	
Synsepalum zenkeri Aubrév. & Pellegr.	Accepted	

Table A1.2 – Species of *Englerophytum* (adapted from the Plant List)

Species Name	Status	Species sequenced
Englerophytum kouloungense Aubrév. & Pellegr.	Unresolved	
Englerophytum kennedyi Aubrév.	Unresolved	
Englerophytum congolense (De Wild.) Aubrév. & Pellegr.	Accepted	
Englerophytum iturense (Engl.) L.Gaut.	Accepted	
Englerophytum koulamoutouense (Aubrév. & Pellegr.) ined.	Accepted	
Englerophytum laurentii (De Wild.) ined.	Accepted	✓
Englerophytum letestui (Aubrév. & Pellegr.) ined.	Accepted	✓
Englerophytum longipedicellatum (De Wild.) ined.	Accepted	
Englerophytum magalismontanum (Sond.) T.D.Penn.	Accepted	✓
Englerophytum mayumbense (Greves) ined.	Accepted	
Englerophytum natalense (Sond.) T.D.Penn.	Accepted	✓
Englerophytum oblanceolatum (S.Moore) T.D.Penn.	Accepted	✓
Englerophytum oubanguiense (Aubrév. & Pellegr.) Aubrév. & Pellegr.	Accepted	✓
Englerophytum rwandense (Troupin) ined.	Accepted	
Englerophytum somiferanum Aubrév.	Accepted	
Englerophytum stelechanthum K.Krause	Accepted	✓

APPENDIX 2

Table A2.1 – Data for all the sequenced samples of *Englerophytum* and *Synsepalum*

No.	Species name	Accession Number	Collector & Collector Number	Source	Country	ITS	trnH-psbA
1	*Englerophytum sp.*	EDNA13-0000032	Wieringa 7702 (WAG)	S	Gabon	✓	✓
2	*Englerophytum sp.*	EDNA13-0033180	Jongkind 5084 (FHO)	H	Ivory Coast	✓	✓
3	*Englerophytum cf kouloungense*	EDNA13-0000033	Maas 10325 (WAG)	H	Gabon	✓	✓
4	*Englerophytum laurentii*	EDNA13-0033184	Harris 9685 (E)	H	Congo	✓	✓
5	*Englerophytum laurentii*	EDNA13-0000026	Van der Laan 231 (WAG)	H	Cameroon	×	×
6	*Englerophytum letestui*	EDNA13-0000025	Sosef 2025 (WAG)	S	Gabon	✓	✓
7	*Englerophytum magalismontanum*	EDNA13-0000017	Balkwill *et al.* 11986 (E)	H	South Africa	✓	✓
8	*Englerophytum magalismontanum*	EDNA13-0000029	Stronkhorst 1 (WAG)	H	Botswana	✓	✓
9	*Englerophytum magalisontanum*	EDNA13-0000019	Chapman 6922 (E)	H	Malawi	✓	✓
10	*Englerophytum natalense*	EDNA13-0000016	Gereau *et al.* 6120 (E)	H	Tanzania	✓	✓
11	*Englerophytum natalense*	EDNA13-0000018	Chapman 6479 (E)	H	Malawi	✓	✓
12	*Englerophytum oblanceolatum*	EDNA13-0000024	Van der Maesen 6154 (WAG)	H	Benin	✓	✓
13	*Englerophytum oubanguiense*	EDNA13-0000020	Harris 8166 (E)	H	Congo	✓	✓
14	*Englerophytum oubanguiense*	EDNA13-0000022	Harris 4924 (E)	H	Central Afr. Rep.	✓	✓
15	*Englerophytum oubanguiense*	EDNA13-0000030	Jongkind 11443 (WAG)	H	Guinea	✓	✓
16	*Englerophytum stelechanthum*	EDNA13-0000015	Waterman & Mckey 868 (E)	H	Cameroon	✓	✓
17	*Synsepalum afzelii*	EDNA13-0033170	Hawthorne, Gyakari 201a 121 (FHO)	H	Ghana	✓	✓
18	*Synsepalum afzelii*	EDNA13-0033175	Hawthorne,Gyakari 200b 212 (FHO)	H	Ghana	✓	✓
19	*Synsepalum aubrevillei*	EDNA13-0033172	Hawthorne, Gyakari 200b 32 (FHO)	H	Ghana	✓	✓
20	*Synsepalum aubrevillei*	EDNA13-0033176	Hawthorne,Gyakari 200b 166 (FHO)	H	Ghana	✓	✓
21	*Synsepalum brevipes*	EDNA13-0033055	Harris 8441 (E)	S	Gabon	✓	✓
22	*Synsepalum brevipes*	EDNA13-0033178	Hawthorne,Gyakari 200b 131 (FHO)	H	Ghana	✓	✓

Table A2.1 (continued) – Data for all the sequenced samples of *Englerophytum* and *Synsepalum*

No.	Species name	Accession Number	Collector & Collector Number	Source	Country	ITS	trnH-psbA
23	*Synsepalum brevipes*	EDNA13-0033179	Hawthorne *et al.* AM1219 (FHO)	H	Senegal	✓	✓
24	*Synsepalum brevipes*	EDNA13-0033182	Moutsambote 6093 (E)	H	Congo	✓	✓
25	*Synsepalum brevipes*	EDNA13-0033185	Harris 9712 (E)	H	Congo	✓	✓
26	*Synsepalum brevipes*	EDNA13-0033520	Sosef 2134 (WAG)	S	Gabon	×	×
27	*Synsepalum congolense*	EDNA13-0033054	Harris 8325 (E)	S	Gabon	✓	✓
28	*Synsepalum dulcificum*	EDNA13-0033053	Moutsambote 6060 (E)	S	Congo	✓	✓
29	*Synsepalum dulcificum*	EDNA13-0033174	Hawthorne, Gyakari 200b 138 (FHO)	H	Ghana	✓	✓
30	*Synsepalum dulcificum*	EDNA13-0033181	Moutsambote 6013 (E)	H	Congo	✓	✓
31	*Synsepalum dulcificum*	EDNA13-0033183	Kami 4327 (E)	H	Congo	✓	✓
32	*Synsepalum fleuryanum*	EDNA13-0033056	Harris 8456 (E)	S	Gabon	✓	✓
33	*Synsepalum latoursvillensis*	EDNA13-0000031	Bissiengou 771 (WAG)	S	Gabon	✓	✓
34	*Synsepalum ntimii*	EDNA13-0033171	Hawthorne, Gyakari 203a 24 (FHO)	H	Ghana	✓	✓
35	*Synsepalum passargei*	EDNA13-0033177	Reitsma 3820 (FHO)	H	Guinea	✓	✓
36	*Synsepalum revolutum*	EDNA13-0000023	Harris 5735 (E)	H	Central Afr. Rep.	×	×
37	*Synsepalum* sp.	EDNA13-0033058	Harris 9579 (E)	S	Congo	✓	✓
38	*Synsepalum* sp.	EDNA13-0033057	Harris 8702 (E)	S	Gabon	✓	✓
39	*Synsepalum* sp.	EDNA13-0000028	Sosef 2619 (WAG)	H	Gabon	✓	✓
40	*Synsepalum stipulatum*	EDNA13-0033052	Harris 9130 (E)	S	Congo	✓	✓
41	*Synsepalum stipulatum*	EDNA13-0033051	Harris 9014 (E)	S	Congo	✓	✓
42	*Synsepalum stipulatum*	EDNA13-0000027	Wieringa 5228 (WAG)	S	Gabon	✓	✓
43	*Synsepalum subcordatum*	EDNA13-0000021	Harris 7562 (E)	H	Central Afr. Rep.	✓	✓
44	*Synsepalum tsuonkpe*	EDNA13-0033173	Hawthorne, Gyakari H200 661 (FHO)	H	Ivory Coast	✓	✓

H = Sample from herbarium specimen; S = Silica gel dried sample; ✓ = region successfully sequenced; × = region not successfully sequenced
WAG = Nationaal Herbarium Nederland – Wageningen Branch; FHO = Daubney Herbarium, Oxford; E = Royal Botanical Gardens Edinburgh.

Table A2.2 – Details of primers used for amplifying target DNA regions

Primer	Direction	Primer sequence (5'-3')	Author
ITS 1	forward	GTAGGTGAACCTGCAGAAGGA	modified White *et al.* (1990)
ITS 4	reverse	TCCTCCGCTTATTGATATGC	White *et al.* (1990)
ITS 5p	forward	GGAAGGAGAAGTCGTAACAAG	Möller & Cronk (1997)
ITS 8p	reverse	CACGCTTCTCCAGACTACA	Möller & Cronk (1997)
trnH	forward	ACTGCCTTGATCCACTTGGC	Hamilton 1999
psbA	reverse	CGAAGCTCCATCTACAAATGG	Hamilton 1999

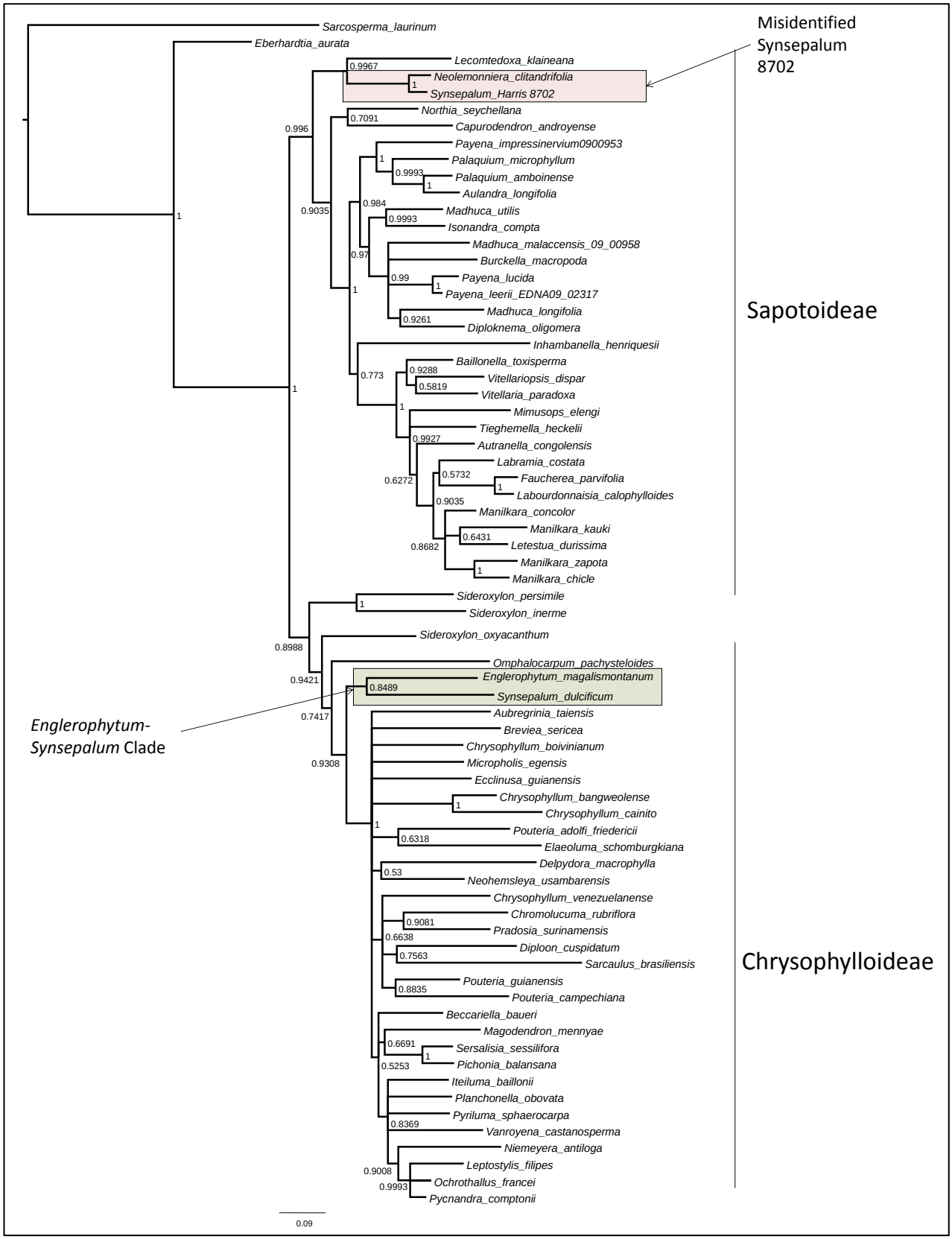
Table A2.3 - Details of the outgroup sequences used in the analysis

No.	Outgroup	Voucher	DNA Region	Accession Number	GI Number	Publication
45	*Chrysophyllum pruniforme*	Jongkind 3762 (WAG)	ITS	DQ246671	82698903	Swenson *et al.* 2008
			trnH-psbA	DQ344100	86774491	Swenson *et al.* 2008
46	*Chrysophyllum roxburghii*	Solo & Randrianasolo 33 (WAG)	ITS	DQ246672	82698904	Swenson *et al.* 2008
			trnH-psbA	DQ344101	86774492	Swenson *et al.* 2008
47	*Eberhardtia aurata*	Hao 534 (S)	ITS	EF558617	156106155	Swenson *et al.* 2008
			trnH-psbA	DQ344106	86774497	Swenson *et al.* 2008
48	*Omphalocarpum pachysteloides*	Jongkind 2351 (WAG)	ITS	AY552151	49344958	Bartish *et al.* 2005
			trnH-psbA	DQ344122	86774513	Swenson *et al.* 2008
49	*Omphalocarpum strombocarpum*	Frimodt-Moller, Joker & Ndangalasi TZ538 (C)	ITS	DQ246685	82698917	Swenson *et al.* 2008
			trnH-psbA	DQ344123	86774514	Swenson *et al.* 2008
50	*Pouteria adolfi-friedericii*	Friis et al. 3502 (UPS)	ITS	AY552115	49344920	Bartish *et al.* 2005
			trnH-psbA	DQ344127	86774518	Swenson *et al.* 2008
51	*Pouteria macrophylla*	Seidel & al. 5905 (K)	ITS	DQ246692	82698924	Swenson *et al.* 2008
			trnH-psbA	DQ344137	86774528	Swenson *et al.* 2008
52	*Xantolis siamensis*	Smitairi 1 (L)	ITS	AY552154	49344959	Bartish *et al.* 2005
			trnH-psbA	DQ344151	86774542	Swenson *et al.* 2008

Table A2.4 – Table showing the list of misidentified taxa and their new determination

Specimen	Old Identification	New identificaton	Reasons for making the change
Jongkind 5084	*Englerophytum oubanguiense*	*Englerophytum* sp.	When foliage of this specimen was compared to known material of *E. oubanguiense*, major differences could be noted in leaf shape & size, leaf indumentum and number of secondary veins. Hence, the possibility of the specimen being *E. oubanguiense* was excluded. In the absence of floral and fruit material, this specimen could not be identified to species however, it possessed characters consistent with the Pennington's current circumscription of *Englerophytum* (leaf striations, brochiododromous venation, parallel tertiary veins) and hence its determination was changed to *Englerophytum* sp.
Sosef 2619	*Englerophytum* sp.	*Synsepalum* sp.	Only 1 leaf was available from this specimen. However, this leaf showed all traits of *Synsepalum*, as presently circumscribed by Pennington (incl. eucamptodromous vention, non-parallel tertiary veins, absence of leaf striations) and was hence relabelled. In the absence of fertile material identification was restricted to the genus level.
Harris 8702	*Synsepalum* sp.	belongs to the Sapotoideae, probably a close relative to *Neolemmoniera*	In all the analyses, this specimen always appeared as sister to the rest of the *Englerophytum-Synsepalum* clade and its sequence aligned better with the outgroups than with representatives from *Synsepalum*. Thus in order to check whether this specimen belonged to the *Englerophytum-Synsepalum* clade, its sequence was incorporated in a matrix of representatives from throughout the Sapotaceae and a Bayesian analysis was run on this matrix (Figure A2.4). It appeared close to *Neolemmoneira* and was hence excluded from the analysis.

Section A 2.5 – Placement of 'Harris 8702' within the Sapotaceae



Note: ITS Data matrix of the Sapotaceae provided by James Richardson

Table A2.6 – Character states for all scored morphological characters

Character	State 1	State 2	State 3	State 4
1. Leaf Length	0-15cm	>15 – 30cm	>30cm	
2. Leaf Width	0-6cm	>6-12cm		
3. Ratio Leaf Width : Leaf Length	0.2-0.3	>0.3 – 0.4	>0.4 – 0.5	
4. Petiole Length	0 – 1cm	>1-2cm	>2-3cm	
5. Ratio Petiole Length: Leaf Length	0-0.1	>0.1		
6. Distance between secondary veins	0-0.5	>0.5-1cm	>1-1.5cm	>1.5cm
7. Ratio Distance Between Secondaries: Leaf length	0-0.05	>0.05		
8. Number of Secondary veins	0 -27	≥28		
9. Ratio Number of Secondary Veins: Leaf Length	0-1	>1-2cm	>2-3cm	>3
10. Leaf Venation	Brochiododromous	Eucamptodromous		
11. Petiole	With closed groove	With open groove	Flat and thickened at base	
12. Intramarginal Vein	Present	Absent		
13. Midrib	Sunken Above	Impressed Above	Flat Above	
14.Intersecondary veins	Present	Absent		
15. Conspicuity of Tertiary venation	Conspicuous	Inconspicuous		
16. Tertiary venation	Parallel to Secondaries	Reticulate	Horizontal	Oblique
17. Stipules	Present	Absent		
18. Hairs on Midrib underside	Present	Glabrous/Subglabrous		
19. Hairs on lamina underside	Velvety	Sparse	Glabrous/Subglabrous	
20. Pedicel	Pedicellate	Sessile/Subsessile		
21. Sepal Length	0-2.5mm	>2.5-5cm		
22. Sepal fusion	Free	Fused		

Table A2.6 (continued) – Character states for all scored morphological characters

Character	State 1	State 2	State 3	State 4
23.Petal Tube length	0-2mm	>2mm		
24. Petal Lobe length	0-2.5mm	>2.5mm		
25. Ratio Petal Lobe Length: Petal Tube Length	1-2	>2		
26. Stamen Fusion	Free	Partly Fused	Completely Fused	
27. Filament Length	0-1.5mm	>1.5mm		
28. Anther dehiscence	Extrorse	Latrorse		
29. Anther length	0-1mm	1-2mm	>2mm	
30. Ratio Anther length: Filament Length	0-0.4	0.4-0.8	0.8-12	>12
31. Staminodes	Absent/Vestigial	Medium (1-2.5mm)	Large (>2.5cm)	
32. Style Length	1-2mm	>2-3mm	>3mm	
33. Ovary Length	1-2mm	>2mm		
34. Ratio Style Length: Ovary Length	0-1	>1-2	>2	

Section A. 2.7 – List of specimens viewed for morphological analysis

Englerophytum laurentii (De Wild.) ined.
Gossweiler, J. 13785 (LISC);

Englerophytum letestui Aubrév. & Pellegr.
Le Testu, G. 8806 - P00417600 (P); Le Testu, G. 8806 - LISC002852 (LISC); Le Testu, G. 8806 - K000430628 (K);

Englerophytum magalismontanum (Sond.) T.D.Penn.
Gossweiler, J. 14151 (K); Chapman, J.D. 6536 (E); Chapman, J.D. & Chapman E.G. 6699 (E); Chapman, J.D. & Chapman E.G. 6878 (E); Chapman, J.D. & Chapman E.G. 6922 (E); Timberlake, J. 5282 (K); Timberlake, J. 5296 (K); Taylor, H.C. 1813 (E); Balkwill, K., McCallum, D.A. & Reddy, R.A. 11986 (E); Wilms, F. 1812 (E); Brenan, J.P.M. 7818 (K); Holmes, W.D. 1170 (K); Nyariri, P. 254 (E);

Englerophytum natalense (Sond.) T.D.Penn.
Balkwill, K. 1769 (E); Chapman, J.D. 6553 (E); Chapman, J.D. 6479 (E); Chapman, J.D. & Chapman E.G. 6935 (E); Rudatis, H. 904 (E); Sim, T.R. 2374 (E); Wood, J.M. 110 (E); Bourquin, O. s.n. (E); Hilliard, O.M.; Burtt, B.L. 8452 (E); Balkwill, K. 6853 (E); Balkwill, K.; Cadman, M.J. 2197 (E); Gereau, R.E., Kayombo, C.J. & Mwangoka, M.A. 6120 (E); Boivin, M. s.n. (E);

Englerophytum oblanceolatum (S.Moore) T.D.Penn.
Adam, J.G. 3477 (P); Adam, J.G. 23191 (P); Scott-Elliott, G.F. 4867 (K); Bangshawe, A.G. 1087 (BM);

Englerophytum oubanguiense (Aubrév. & Pellegr.) Aubrév. & Pellegr.
Waterman, P.G. & McKey, D. 887 (E); Harris, D.J. 5027 (E); Harris, D.J. 4623 (E); Harris, D.J. 4924 (E); Harris, D.J. 8166 (E); Jongkind, C.C.H. 5084 (FHO);

Englerophytum stelechanthum K. Krause
Waterman, P.G. & McKey, D. 868 (E); Mildbraed, J. 6113 - HBG510668 (HBG); Mildbraed,

J. 6119 - HBG510669 (HBG).

Synsepalum afzelii (Engl.) T.D.Penn.
Irvine, F.R. 2253 (E); Hawthorne, W.D. 200b 212 (FHO); Hawthorne, W.D. 201a 121 (FHO); Kennedy, J.D. 1021 (E);

Synsepalum aubrevillei (Pellegr.) Aubrév. & Pellegr.
Hawthorne, W.D. 200b 166 (FHO); Hawthorne, W.D. 200b 32 (FHO); Aubréville, A. 133 - P00417563 (P); Aubréville, A. 133 - P00417564 (P);

Synsepalum brevipes (Baker) T.D.Penn.
Goetze, W. 883 (E); Zenker, G.A. 2587 - E00330774 (E); Bates, G.L. 325 (E); Zenker, G.A. 2404 - E00330773 (E); Zenker, G.A. 2404 - E00330788 (E); Zenker, G.A. 2587 - E00330769 (E); Zenker, G.A. 3797 - E00330519 (E); Zenker, G.A. 3797 - E00330772 (E); Zenker, G.A. 3817 (E); Zenker, G.A. 4324 (E); Harris, D.J. 7683 (E); Harris, D.J. & Fay, M. 1452 (E); McDonald, K. 25 (E); Harris, D.J. 8441 (E); Hawthorne, W.D. 200b 131 (FHO); Chapman, J.D. & Chapman E.G. 7715 (E); Chapman, J.D. & Chapman E.G. 9423 (E); Buchanan, J. 151 (E); Chapman, J.D. & Chapman E.G. 7853 (E); Dalziel, J.M. 1273 (E); Dalziel, J.M. 1385 (E); Hawthorne, W.D. AM 1291 (FHO);

Synsepalum congolense Lecomte
Le Testu, G. 1769 - K000430659 (K); Harris, D.J. 8325 (E); Le Testu, G.M.P.C. 1769 (E);

Synsepalum dulcificum (Schumach. & Thonn.) Daniell
Dalziel, J.M. 5 (E); Irvine, F.R. 1582 (E); Hawthorne, W.D. 200b 138 (FHO);

Synsepalum fleuryanum A.Chev.
Chevalier, A. 26309 - P00417569 (P); Chevalier, A. 26309 - P00417570 (P); Harris, D.J. 8456 (E);
Synsepalum lastoursvillensis (Aubrév. & Pellegr.) Ewango
Le Testu, G. 8280 - P00417556 (P); Le Testu, G. 8280 - P00417557 (P); Le Testu, G. 8280 - P00417558 (P);

Synsepalum ntimii Hawthorne
Hawthorne, W.D. 203a 24 (FHO);

Synsepalum passargei (Engl.) T.D.Penn.
Jacques-Felix, H. 4280 (K); De Wild. 2577 (K); De WItte, L. 6093 (K); Reitsma, J.M. 3820 (FHO); Stolz, A.E. 89 (E); Milne-Redhead, E. 9137 (K);

Synsepalum stipulatum (Radlk.) Engl.
Dawe, M.T. 291 (K); De Wild. 1265 (K); Zenker, G.A. 3662 (E); Thomas, D.W. 428 (K); Harris, D.J. 4937 (E); Gentry, A.H. & Harris, D.J. 62781 (E); Goldsmith, M. 225 (E); Harris, D.J. 3992 (E); Harris, D.J. 4325 (E); Harris, D.J. 4400 (E); Harris, D.J. 5365 (E); Harris, D.J. 5532 (E); Harris, D.J. 7168 (E); Harris, D.J. & Fay, M. 48 (E); Harris, D.J. & Fay, M. 774 (E); Ndolo Ebika, S.T. 327 (E); Bokdam, J. 4493 (K); Wilks, C. 986 (E); Wieringa, J.J. 5228 (E); Wieringa, J.J. 5228 (WAG); Kennedy, J.D. 528 (E);

Synsepalum subcordatum De Wild.
Tisserant, R.P. 890 (K); Harris, D.J. 3562 (E); Harris, D.J. 8939 (E); Harris, D.J. 9321 (E); Ndolo Ebika, S.T. 349 (E); Lisowski, S. 15358 (K); Louis, J. 1429 (K); Louis, J. 3270 (K);

Synsepalum tsuonkpe Aubrév. & Pellegr.
Hawthorne, W.D. H200 661 (FHO);

Table A2.8 – Score table showing the scores of each specimen for all 35 morphological characters investigated

	0 1	**0 2**	**0 3**	**0 4**	**0 5**	**0 6**	**0 7**	**0 8**	**0 9**	**1 0**	**1 1**	**1 2**	**1 3**	**1 4**	**1 5**	**1 6**	**1 7**	**1 8**	**1 9**	**2 0**	**2 1**	**2 2**	**2 3**	**2 4**	**2 5**	**2 6**	**2 7**	**2 8**	**2 9**	**3 0**	**3 1**	**3 2**	**3 3**	**3 4**	**3 5**
Englerophytum sp. (1)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Englerophytum sp. (2)	2	1	1	2	1	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
E. cf. kouloungense (3)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
E. laurentii (4)	1	1	3	2	1	1	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	2	3	1	1	2	1
E. letestui (6)	2	1	1	3	2	2	1	2	4	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	1	1	2	?	3	3	1	2	2	1
E. magalismontanum (7)	1	1	1	2	2	1	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	2	2	2	2	2	1	2	1
E. magalismontanum (8)	1	1	1	2	2	1	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	2	2	2	2	2	1	2	1
E. magalismontanum (9)	1	1	1	2	2	1	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	2	2	2	2	2	1	2	1
E. natalense (10)	1	1	1	2	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2	1	2	1	1	1	2	2	2	2	1	2	1	2
E.natalense (11)	1	1	1	2	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2	1	2	1	1	1	2	2	2	2	1	2	1	2
E. oblanceolatum (12)	1	1	2	2	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	2	?	2	2	1	2	2	1
E. oubanguiense (13)	2	2	1	2	1	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	1	3	2	1	3	2	1	2	2	1
E. oubanguiense (14)	2	2	1	2	1	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	1	3	2	1	3	2	1	2	2	1
E.oubanguiense (15)	2	2	1	2	1	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	1	3	2	1	3	2	1	2	2	1
E.stelechanthum (16)	2	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	3	2	1	3	3	1	2	2	1
S. afzelii (17)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	2	2	1	2	3	2	1	1	1	1	1	1	1	1	1	2	3	1	1	2	1
S.afzelii (18)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	2	2	1	2	3	2	1	1	1	1	1	1	1	1	1	2	3	1	1	2	1
S. aubrevillei (19)	2	2	3	1	1	4	2	1	1	2	2	2	2	2	1	4	?	1	1	2	1	2	2	2	2	1	1	1	1	3	4	3	3	1	3
S.aubrevillei (20)	2	2	3	1	1	4	2	1	1	2	2	2	2	2	1	4	?	1	1	2	1	2	2	2	2	1	1	1	1	3	4	3	3	1	3
S. brevipes (21)	2	1	2	1	1	4	2	1	1	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	2	1	1	2	1	2	3	1	2	1	2
S.brevipes (22)	2	1	2	1	1	4	2	1	1	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	2	1	1	2	1	2	3	1	2	1	2
S. brevipes (23)	2	1	2	1	1	4	2	1	1	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	2	1	1	2	1	2	3	1	2	1	2
S. brevipes (24)	2	1	2	1	1	4	2	1	1	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	2	1	1	2	1	2	3	1	2	1	2
S.brevipes (25)	2	1	2	1	1	4	2	1	1	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	2	1	1	2	1	2	3	1	2	1	2

Table A2.8 (continued) – Score table showing the scores of each specimen for all 35 morphological characters investigated

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
S. congolense (27)	1	1	2	2	2	3	2	1	2	2	3	2	2	2	1	4	1	1	2	2	1	2	2	2	2	1	1	2	1	2	2	2	3	1	3
S. dulcificum (28)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	1	4	2	1	2	2	2	2	2	2	2	1	1	2	1	2	2	2	3	1	3
S. dulcificum (29)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	1	4	2	1	2	2	2	2	2	2	2	1	1	2	1	2	2	2	3	1	3
S. dulcificum (30)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	1	4	2	1	2	2	2	2	2	2	2	1	1	2	1	2	2	2	3	1	3
S.dulcificum (31)	1	1	2	1	1	2	2	1	2	2	2	2	2	2	1	4	2	1	2	2	2	2	2	2	2	1	1	2	1	2	2	2	3	1	3
S. fleuryanum (32)	1	1	1	1	1	2	2	1	2	2	2	2	2	2	1	4	1	1	1	2	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?
S. lastoursvillensis (33)	3	2	2	3	1	4	1	1	1	2	2	2	2	2	?	4	1	1	?	2	2	2	1	2	2	1	1	2	2	3	2	3	?	2	2
S. ntimii (34)	1	1	3	1	1	2	2	1	2	1	2	2	2	1	1	2	?	2	3	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
S.passargei (35)	1	1	2	2	2	2	2	1	2	2	2	2	2	2	1	4	1	1	2	2	1	1	1	1	2	2	1	2	2	1	1	2	1	2	2
Synsepalum sp. (37)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Synsepalum sp. (39)	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
S. stipulatum (40)	1	1	2	3	2	2	2	1	2	2	3	2	2	2	1	4	1	2	3	2	1	2	2	2	2	1	1	2	1	1	2	2	3	1	3
S. stipulatum (41)	1	1	2	3	2	2	2	1	2	2	3	2	2	2	1	4	1	2	3	2	1	2	2	2	2	1	1	2	1	1	2	2	3	1	3
S. stipulatum (42)	1	1	2	3	2	2	2	1	2	2	3	2	2	2	1	4	1	2	3	2	1	2	2	2	2	1	1	2	1	1	2	2	3	1	3
S.subcordatum (43)	3	2	1	1	1	4	2	2	1	2	2	2	2	2	1	4	?	1	3	2	1	2	2	1	2	1	1	1	1	2	3	2	3	1	2
S. tsounkpe (44)	2	2	2	2	1	4	2	1	1	2	2	2	2	2	2	4	?	2	3	2	2	2	2	2	2	1	1	2	?	1	1	3	3	1	3
C. pruniforme (45)	1	1	2	1	1	1	1	2	4	1	?	1	3	2	2	2	2	1	3	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2
C.roxburghii (46)	1	1	1	1	1	1	1	2	4	1	?	1	3	2	2	2	2	1	3	1	1	1	1	?	?	?	1	?	1	?	?	1	1	?	?
Eberhardtia aurata (47)	2	2	2	3	2	4	2	1	1	2	2	2	2	2	1	4	1	1	1	2	1	2	1	2	2	1	1	2	1	?	?	3	?	?	?
O.pachysteloides (48)	2	2	3	3	2	3	2	1	1	2	?	2	2	1	1	2	1	1	3	2	1	2	1	1	2	1	1	2	1	2	3	3	3	1	2
O.strombocarpum (49)	2	2	1	3	?	?	?	?	?	2	?	?	?	1	?	2	1	?	3	2	1	?	1	?	?	?	2	?	1	?	?	3	?	?	?
P. adolfi-friedericii (50)	2	2	2	2	1	3	2	1	1	2	?	2	1	2	1	4	2	1	1	2	1	2	1	?	?	?	1	?	2	2	?	3	3	1	3
P.macrophylla (51)	1	2	3	2	2	3	2	1	2	2	2	2	2	2	1	4	2	1	2	2	1	2	1	2	2	1	1	2	1	2	2	3	3	1	3
Xantolis siamensis (52)	1	1	3	1	2	1	2	1	3	1	?	2	2	1	1	2	2	1	3	2	1	2	1	1	2	2	1	2	1	2	3	3	3	1	2