PHYLOGENETIC RELATIONSHIPS OF THE GENUS *ANDREAEA* HEDW. (ANDREAEACEAE, BRYOPHYTA) AS INFERRED FROM *RPS4* AND *TRNL-F* SEQUENCES AND MORPHOLOGY

David Kananga Chuba

A THESIS SUBMITTED TO THE DEPARTMENT OF BOTANY, FACULTY OF SCIENCE, UNIVERSITY OF CAPE TOWN, IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN SYSTEMATICS AND BIODIVERSITY SCIENCE.

MARCH 2001

Supervisor: Dr. T. A. J. Hedderson

ABSTRACT

The moss genus *Andreaea* possesses some sporophyte features that resemble liverworts or mosses and some gametophyte features resembling only the mosses, whereas other features are unique. This thesis presents the first phylogenetic study of the genus, based on both morphological and molecular evidence. Gametophyte and sporophyte characters were utilised for cladistic analysis. Sequence data was also generated from two chloroplast gene loci, the *trnL-F* intergenic spacer and the coding region of the ribosomal protein S4 (*rps4*). Separate morphological and molecular analyses produced topologies incongruent in certain parts and congruent in others. However, their combined analysis was better supported and therefore offered a more reliable hypothesis. The inferred phylogeny supported the monophyly of the genus. However, the monophyly of most infra-generic groups was largely contradicted. The putative subgenus *Chasmocalyx* is monotypic with *A. nivalis* as the sole species, whereas *A. australis* and *A. nitida* were resolved within the more broadly circumscribed section *Andreaea* of subgenus *Andreaea*. The section *Nerviae* of subgenus *Andreaea* is more narrowly circumscribed. *A. blyttii* (presumed member of section *Nerviae*) forms a basal lineage separate from all other species and apparently should constitute another monotypic subgenus. *A. wilsonii* (traditional Subgenus *Acroschisma*) is embedded within section *Andreaea* of Subgenus *Andreaea*. *A. subulata*, a presumed member of the section *Nerviae* (Subgenus *Andreaea*) is included in the section *Andreaea*. Character state optimisation has shown that falcate leaves, possession of a leaf costae and medium sized spores are some of the pleisiotypic features within *Andreaea*. However, a number of phylogenetic questions, regarding infra-generic relationships of the genus still remain unanswered. Directions for further future work have been suggested.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. T. A. J. Hedderson for his patience and guidance in my work. I would also like to thank Tracey Nowell, who guided me through the ropes of the molecular techniques that were utilised in this study. I would also like to thank the electron microscopy unit particularly Ms. Miranda Waldron, for making available their microscope facilities, which were utilised for taking the pictures presented in this Thesis. Further appreciation goes to Mr. George Gibbings' for his help with the running of all the sequence product samples through the sequencer at University of Reading. I would like to thank Prof. H. P. Linder for making available his laboratory facilities, especially for the morphological work and also part of the analysis. The Microbiology Department (now Department of Molecular and Cellular Biology) is thanked for permitting the use of their laboratory facilities for certain aspects of the study.

I thank the Southern African Botanical Diversity Network (SABONET) for financial assistance that enabled me to pursue the Master of Science degree programme under which the current study falls. SABONET is a GEF project implemented by UNDP and the regional Centre for Southern Africa, Gaborone, Botswana, US Agency for International Development, Plot no. 14818, Lebatlane Road, Gabarone West, extension 6, Gabarone, under the terms and grant No. 690-0283-A-00-5950.

I also thank all those who encouraged me throughout the period of my Masters. Lastly I thank my Wife and my Son for moral support and for patiently enduring my long periods of absence from home during the course of the MSc programme.

CONTENTS

Arrangement of the Thesis

This thesis is based on a six-months study, involving work towards the determination of phylogenetic relationships of the moss genus *Andreaea*, and thus attempting to answer the questions outlined above. Chapter 1 is an introductory chapter, presenting a brief introduction to the genus, an overview of the taxonomic history, a description of some morphological features in a phylogenetic context, and a discussion of putative relationships of the genus to the major moss groups. Chapter 2 is the first analytical chapter, investigating morphological features of the genus *Andreaea* and analysing them cladistically using the maximum parsimony approach. Chapter 3 is the second analytical chapter describing the molecular methods utilised and presenting the cladistic analysis of the separate *trnL* dataset, and the various combined *trnL/rps4* datasets. In chapter 4, the analyses of various other combinations of morphological and molecular datasets are presented. Finally, in chapter 5, the general discussion of results and conclusions, and implications for classification and character evolution tendencies are given. Possible infra-generic groups are also suggested.

CHAPTER 1

GENERAL INTRODUCTION

The moss genus *Andreaea* was established by Hedwig in 1801. It is cosmopolitan, growing in montane to alpine regions (Scott, et al., 1976; Murray, 1988b). Species are especially abundant in the cooler climates of the temperate, oceanic and sub-polar regions of both Hemispheres, and fewer species are found only on top of high mountains in the tropics or subtropics (Cavers, 1911, Schofield, 1985). Most species of *Andreaea* grow as blackish or reddish patches (Magill, 1981) on rocks that are rich in silica and thus acidic (Braithwaite, 1887; Schofield, 1985). They usually grow in exposed places though some species (e.g. *A. nivalis* Hook.) grow on wet stones in streams (Cavers, 1911).

Composition

Estimates of the number of species in the genus are varied, e.g. 100 (Cavers, 1911; Magill, 1981; Schofield, 1985) and 50 (Braithwaite, 1887; Murray, 1988b). This has probably been due to confusion arising from the large number of nomenclatural changes that have taken place in the group coupled with the large number of species described since the establishment of the genus. Another source of differences in numbers of species is the diverse opinions about characters on which the specific classification has to be based e.g. mostly gametophytic characters (Sainsbury, 1955) or mostly sporophytic characters (Murray, 1988a).

The genus *Andreaea* alone constitutes the family Andreaeaceae (Murray, 1988a), whereas the Andreaeaceae and Andreaeobryaceae (consisting of the monotypic genus *Andreaeobryum* Steere & B. Murr.) constitute the orders Andreaeales and Andreaeobryales respectively, the two orders together constituting the subclass Andreaeidae (Schofield, 1985) as well as the class Andreaeopsida (Murray, 1988a).

Primary systematic features

Features that primarily distinguish the genus *Andreaea* include the valvate capsule (which with that of *Andreaeobryum*, is unique among mosses), the presence of a pseudopodium (a leafless prolongation of the gametophyte on which a capsule is elevated during maturation), the dome-like sporogenous sac that overarches the columella (similar to features occuring in the genus *Sphagnum* L. and hornworts), the well developed protonema and biseriate rhizoids, and the spirally arranged, sometimes costate leaves (Scott, et al., 1976, Murray, 1988a). Other systematic features include a perichaetium (i.e. archegonium and specialised protecting leaves) that is developed partly before fertilization, protonemal appendages that are dorsiventral and rare, and a small, mostly bi-stratose, apically persistent calyptra (Murray, 1988a).

Brief taxonomic history of *Andreaea*

The earliest known species of *Andreaea* were described even before the genus *Andreaea* was established. These were described by Dillenius (1741) under the genus *Lichenastrum* Dill. ex C. Stewart, in his "Historia Muscorum". Dillenius actually placed these species of *Andreaea* together with the leafy hepatics in the same genus *Lichenastrum*. Later Linnaeus (1753) again placed the then known species of *Andreaea* together with all the leafy liverworts, but in *Jungermannia* L, the only leafy liverwort genus he recognised. The placement in *Jungermannia* was based on the longitudinal dehiscence of capsules, which is also the primary mode in liverworts. It was Ehrhart (1778), who eventually established the new genus *Andreaea* in honour of J. G. R. Andreae. However, Ehrhart included only one species (*A. petrophila,* formerly *Jungermannia alpina* L.) in the genus and retained it among the liverworts.

In 1801, Hedwig finally placed the genus among the mosses based on the presence of a columella, an aerial calyptra and the lack of elaters that characterise the Hepatics. He however described the 4 valves of its capsule as peristome teeth united to a persistent operculum. Since Hedwig's treatment, the genus has been placed among the mosses with no changes except those of infra-generic groupings (outlined below), reconsideration of species limits, and descriptions of new species and sub-specific taxa.

Infra-generic classification

Over the years, *Andreaea* has been divided into two and later three subgenera (Dixon, 1929). In the process there have been many varied subgeneric and sectional classification changes as briefly exemplified in Table 1. Currently, the genus is arranged in three subgenera (Braithwaite, 1887, Dixon, 1929, Sainsbury, 1955, Murray, 1988b) namely: subgenus *Andreaea,* subgenus *Chasmocalyx* (Braithw.) Limpr and subgenus *Acroschisma* (Hook.) Wils. Subgenus *Andreaea* is distinguished by large convolute perichaetia and deeply 4-fid capsule and includes the large sections *Andreaea* and *Nerviae*. Subgenus *Chasmocalyx*, lacking any evident perichaetial

leaves and with deeply 6 valved capsules, includes *A. nivalis* Hook, *A. nitida*, *A. rigida* Wilson, *A. fuegiana* (Cardot) S.W. Greene, *A. pachyphylla* (C. Mull.) Broth.

Table 1. Summary of the generic and sub-generic taxa in which species of the genus *Andreaea* have been placed since its establishment. The information was obtained from the *Missouri Botanical Garden - TROPICOS Nomenclatural Data Base – (05 Jul 2000)* as well as a number of other literature sources listed in the reference section.

Taxon	Publication Year	Published in
Andreaea subgenus *Acroschisma* Hook. f. & Wilson.	1844	*London Journal of Botany* 3: 536.
Acroschisma (J. D. Hooker & W. Wilson) Lindl.	1846	*The Vegetable Kingdom* 63.
Andreaea section *Andreaea* Müll. Hal.	1848	*Synopsis Muscorum Frondosorum omnium hucusque Cognitorum* 1: 6.
Andreaea section *Acroschisma* (Hook. f. & Wilson) Müll. Hal.	1848	*Synopsis Muscorum Frondosorum omnium hucusque Cognitorum* 1: 11.
Andreaea section *Euandreaea* Lindb. ex Braithw.	1880	*The British Moss-flora* 1: 6. 1880.
Andreaea section *Chasmocalyx* Lindb. ex Braithw.	1880	*The British Moss-flora* 1: 15.
Andreaea subgenus *Euandreaea* (Lindb. ex Braithw.) Lindb.	1885	*Die Laubmoose Deutschlands, Oesterreichs und der Schweiz* 1: 139.
Andreaea subgenus *Chasmocalyx* (Lindb. *ex* Braithw.) Lindb.	1885	*Die Laubmoose Deutschlands,* esterreichs und der Schweiz 1: 152.
Andreaea section *Nivales* Lindb.	1897	*European and N. American Bryineae (Mosses)* 2: 392.
Andreaea section *Petrophilae* Lindb.	1897	*European and North American Bryineae (Mosses)* 2: 391.
Andreaea section *Rupestres* Lindb.	1897	*European and North American Bryineae (Mosses)* 2: 392.
Andreaea section *Nerviae* Cardot.	1908	*Wissenschaftliche Ergebnisse der Schwedischen Südpolar-Expedition* 1901--1903 4(8): 55.
Andreaea section *Enerviae* Cardot.	1908	*Wissenschaftliche Ergebnisse der Schwedischen Südpolar-Expedition* 1901--1903 4(8): 51.
Neuroloma Cardot.	1911	*Revue Bryologique et Lichénologique* 38: 50.
Andreaea subsection *Costatae* Mönk.	1927	*Die Laubmoose Europas* 124.
Andreaea subsection *Ecostatae* Mönk.	1927	*Die Laubmoose Europas* 124.
Neuroloma Cardot ex A. Donat.	1936	*Revista Sudamericana de Botánica* 3: 67.
Andreaea section *Depressinerves* W. Schultze-Motel.	1970	*Wildenowia* 6: 30.
Andreaea section *Nitidae* W. Schultze-Motel.	1970	*Wildenowia* 6: 29.

and *A. australis* Muell, whereas subgenus *Acroschisma* with the capsule cleft only at the upper end into 6 — 8 valves and very large and convolute perichaetial leaves comprises only of *A. wilsonii*. The current designation of species studied here to various subgenera and sections is given in appendix 1. For species delimitations, the *A. rupestris* complex (*A. rupestris* var. *rupestris*, *A. rupestris* var. *papillosa*, *A. alpestris*, *A. obovata* var. *hartmannii*, and *A. obovata* var. *obovata*) has been the most difficult to delimit (Sainsbury, 1955, Murray 1988b, Chiang, 1998) due to their possession of a lot of polymorphic characters. Similar difficulty is to be expected in inferring phylogenetic relationships on morphological basis.

Relationships with other Bryophytes

The genus *Andreaea* has features linking it not only to the other mosses but also to the other two major groups of bryophytes, the Anthocerotophyta (hornworts) and the Marchantiophyta (liverworts).

The longitudinal dehiscence of capsules, valves with secondary thickening and the absence of a seta (possesses a pseudopodium) are features that *Andreaea* shares with both the liverworts and the hornworts (Sim, 1926, Watson, 1971; Murray, 1988a). The common tendency of valves to remain attached apically is another feature linking the liverworts to *Andreaea* (Murray, 1988a).

Within the Bryophyta, the Andreaeaceae have been considered to be in various respects as related to all the major lineages, though apparently not closely related to any other groups of the mosses (Schofield, 1985). They have been considered (e.g. by Braithwaite, 1887 and Cavers, 1911) to be intermediate between the Sphagnales and Bryales especially in the structure and development of the sporangium. Steere and Murray (1976) also suggested that on account of its possession of a seta, Andreaeobryaceae linked Andreaeaceae to the Bryopsida (The "True" Mosses). Many other authors (e.g. Dixon, 1932 and Schofield, 1985) have placed it at the start of their classifications, thus emphasising its putatively primitive features. The widely held view is that the Andreaeopsida are an isolated line characterised by many generalised as well as peculiar features, and are thus both markedly different from,

but still share a number of other similar features with all major groups of bryophytes (Murray, 1988a). The sporangium of *Andreaea* for example has a unique combination of characters, some of which are peculiar to the genus whereas other structural and embryological features are shared with the Sphagnales and Bryales (Cavers, 1911, Murray, 1988a).

For the Sphagnales, the presence of a pseudopodium in *Andreaea*, the nature of the sporogenous layer (dome shaped over the columella), the capsule that is at first enclosed in a large saccate calyptra, and then elevated on an elongated pseudopodium and a prothalium that is of a somewhat lobate form (Braitwaite, 1887), are very similar features linking them to the Andreaeaceae (Schofield, 1985; Murray, 1988a). The Andreaeaceae, however, differs from the Sphagnales in characters such as the mode of formation of the sporogenous layer (Schofield, 1985) and in the longitudinal dehiscence of the capsule.

Further, in a comparison of features of *Andreaeobryum*, *Andreaea* and *Takakia* S. Hatt. & Inoue with those of other bryophytes, Murray (1988a) concluded that the Andreaeopsida (sensu Murray, 1988a) are cladistically and patristically primitive and that their closest relative is the even more primitive *Takakia* (Takakiaceae). Thus the genus *Takakia* has also been related to *Andreaea* through *Andreaeobryum* and according to Murray (1988a), the widely held view that the class Takakiopsida represents the most primitive bryophyte known is a valid one. Murray (1988a) considers *Takakia* as the taxon that is most closely related to Andreaeopsida. Interestingly, some morphological features that separate the genus *Andreaeaobryum* from *Andreaea* unite *Andreaeaobryum* and *Takakia* and appear to be otherwise unknown among the Bryophytes (Murray, 1988a). Later Smith and Davidson (1993) even included Takakiaceae as another family within the Andreaeopsida. However, recent molecular studies (e.g. Hedderson et al., 1998 and Newton et al., 2000) suggest a closer relationship of *Takakia* to *Sphagnum* than to either *Andreaea* or *Andreaeobryum*.

The foregoing outline of the similarities of *Andreaea* to so many bryophyte groups as well as its peculiarities from them all is fittingly summed up in Murray (1988a)'s description of the genus as "isolated on the one hand, prototypic and synthetic on the

other". This certainly makes the genus *Andreaea* a very interesting group phylogenetically. The presence in *Andreaea* of so many shared features with all the major bryophyte groups suggests that these features may be pleisiomorphic for Bryophytes (Murray, 1988a).

Understanding the phylogenetic relationships of this genus is therefore a vital piece in understanding the advent of many characteristics of the various present day bryophyte taxa. However, as amply demonstrated by the brief history presented here, there have been varied views of not only what species limits should be upheld within *Andreaea* but also differences in opinions of delimitations of infra generic taxa (e.g. Table 1), and of the characters (synapomorphies) on which these should be based. A phylogenetic analysis of various data types, utilising a number of phylogeny inference approaches would help evaluate species relationships as well as the monophyly of the various traditional sub-generic groups within *Andreaea*.

Phylogenetic studies in which the species of *Andreaea* have been included (e.g. Mishler et al., 1992; Garbary et al., 1993; Garbary and Renzaglia, 1998; Hedderson, et al., 1996; Hedderson et al., 1998, Newton et al., 2000) have all been higher-level studies dealing with relationships at family level or above, and only including very few species (i.e. not more than three). Phylogenetic relationships within *Andreaea* are therefore not clearly understood (Newton et al., 2000). This study though indeed preliminary in investigating phylogenetic relations for the genus aims at including representative species of all the major sub-groupings within the genus. The study has also been clearly focussed towards highlighting some plausible evolutionary character changes for the species of the *Andreaea*.

GENERAL OBJECTIVE

The purpose of this study was mainly to determine phylogenetic relationships of the genus *Andreaea* using morphological and molecular sequence data (from *rps4* and *trnL-F* chloroplast gene regions) and various phylogeny inference approaches.

SPECIFIC OBJECTIVES

Some questions addressed in this study were as follows.

1. Is the genus *Andreaea* monophyletic i.e. are the other genera of Andreaeopsida sensu Smith and Davidson, 1993 (i.e. *Andreaeobryum* and *Takakia*) embedded within this genus or not?
2. Are the major putative infra-generic taxa (as in introduction) valid?
3. Are costate species (mostly section *Nerviae*) more closely related to each other than to ecostate species (section *Andreaea*) and vice versa or not?
4. Which character states (e.g. gametophytic and sporophytic) are pleisiomorphic and which ones are derived, within the genus *Andreaea*?

CHAPTER 2

MORPHOLOGY AND CLADISTIC ANALYSIS OF THE GENUS *ANDREAEA*

INTRODUCTION

The genus *Andreaea* consists of autoicous or dioicous plants that occur in reddish-brown to blackish tufts or cushions (Scott, et al., 1976). The slender plants are brittle and fragile when dry with the lower parts of plants often much worn. The fragments broken off can act as propagules (Scott, et al., 1976). The branched protonema is attached to the substratum by rhizoids, both structures ranging from uniseriate to multiseriate (Schofield, 1985). The usually branched stems of *Andreaea* bear spirally arranged leaves that are erect to squarrose or falcate to falcate-secund (Schofield, 1985). A range of leaf areolation types occur as will be detailed in the later sections. Costae may be present or absent, weak or strong, ending below the apex, percurrent, excurrent or decurrent. Axillary hairs usually have 1-2 quadrate basal cells and usually one elongate hyaline terminal cell. Perichaetial leaves are often differentiated, convolute and sheathing (Murray, 1988b), sometimes sheathing but not convolute. Capsules dehiscence results in 4 to 8 longitudinal valves. Annuli, opercula and peristomes are absent. The commonly spherical spores range from small (*ca.*9 μm) to large (*ca.*110 μm). The calyptra is campanulate-mitrate, apical and very small (Murray, 1988a, Cao & Chien. 1995).

Indeed, due to the probable antiquity of *Andreaea*, there are many possibilities of morphological character evolution so that it is difficult and, according to Smith (1986), perhaps impossible to determine [on the basis of morphology alone] whether polyphyly or monophyly, or analogies or homologies are at hand, or in which direction changes have occurred.

However, a few suggestions have been made regarding the polarity of character evolution, or at least the order of change of certain characters. For example in the case of leaf differences, Schofield (1985) viewed the costate species of *Andreaea* as more generalised (i.e. possessing pleisiotypic characters states), especially those in

which perichaetial leaves are basically like the vegetative leaves. According to Schofield (1985), species show an apparent reduction series in sporophytic features, with an increase in specialisation of the sporangium for spore dispersal. In *A. nivalis*, for example, the dehiscence lines extend the entire length of the sporangium whereas in *A. morrisonensis* they are reduced and confined to the upper half (Schofield, 1985). In *A. wilsonii*, the dehiscence lines are more numerous and further reduced to near the apex of the sporangium (Schofield, 1985).

The current taxonomic classification of the genus *Andreaea* has been based on morphological characters. However as mentioned above, there have been differences in opinions of which morphological characters are important for setting species limits. For example, according to Sainsbury (1955) fruiting characters in the genus are not helpful for specific distinctions and the bracts and leaves are the bases on which characters should be founded. For costate species (section *Nerviae*), Schulzte-Motel (1970) also considered the costa-lamina relationship in the upper part of the leaf as the most important character, defining taxa. Murray (1988b) however has esteemed sporophyte characters, especially spore size, over gametophyte characters, more so for delimiting the costate taxa of *Andreaea*. Though characters that delimit taxa may not necessarily be useful for phylogeny inference, all the different views expressed by the various researchers have served as a useful guide in this study to the possible characters (synapomorphies) for inferring phylogenetic relationships of the genus. The morphological analysis in this study was thus largely based on gametophyte, especially leaf, characters that have also featured prominently as the basis for much of the traditional species delimitations. A number of sporophyte characters were also included.

AIMS OF THE CHAPTER

The aims of this chapter were to examine morphological variation (gametophytic and sporophytic) for the genus *Andreaea* in a cladistic context and from this information, determine phylogenetic relationships using the maximum parsimony approach. The monophyly of the various infra-generic groups was also tested based on this morphological evidence.

MATERIALS AND METHODS

Specimen sampling

This study utilised 61 herbarium specimens for morphological analysis (see appendices 1A and 1B) comprising 19 species and 8 subspecies of *Andreaea,* the only species of *Andreaeobryum*, 1 species of *Takakia*, 2 species of the genus *Tetraphis* and 2 species of *Sphagnum*.

The species included in the study were selected mainly on the basis of availability of herbarium specimens. Taxa representing the presumed phylogenetic divergences within *Andreaea* (i.e. the three main subgenera and sections within these) were included, and an attempt was made to represent character variation that has previously been utilised to postulate groupings (as indicated in the introduction). Species were also selected to include representatives of different geographic regions around the world. However, due to time and resource limitations, it has not been possible to include species from every major geographic region or many representatives of the different infra-generic groups. The included species however are sufficient to give useful insights for the questions being addressed in this thesis.

To determine the polarity of character evolution within *Andreaea*, the outgroup method was utilised (Jefferies, 1979; Nixon and Carpenter, 1993). Based on the earlier discussion (chapter 1) of the similarities of *Andreaea* to the other major groups of bryophytes and also on results of a number of molecular studies (e.g. Hedderson et al., 1998; Newton et al 2000), *Tetraphis*, *Sphagnum*, *Takakia* and *Andreaeobryum* are potential candidates as outgroup taxa. The trees in the separate analyses were therefore rooted to the genus *Sphagnum* with the inclusion of *Tetraphis Takakia* and *Andreaeobryum*.

Specimen examination

The morphological work involved the study of both gametophyte and sporophyte characters. As far as possible, information was obtained from herbarium specimens rather than literature, to avoid problems associated with inconsistency in

interpretation by various authors. However where unavoidable, due to lack of specimens, or where confirmation was necessary, the available literature was consulted. Whole herbarium specimens were examined by teasing out a few individual plants from masses of herbarium material. Specimens were mounted in water for easy examination. These were then examined using the Zeiss Stemi SV6 or Leica MS5 stereomicroscopes. In many cases a number of specimens were examined for each taxon and these formed the basis for character states assigned to that particular taxon.

Gametophyte and sporophyte parts were mounted in Hoyers' solution on microscope slides, and examined under the Zeiss Standard 25 compound microscope. For many specimens where sporophytes were not present, spores were easily found among the perichaetial and terminal leaves. However in other instances the measurements were obtained from literature. The characters listed and described below were scored for cladistic analysis. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision) or the Leitz Diaplan compound microscope (Leitz Wetzler, German), using a Zeiss AxioCam camera (Carl Zeiss Vision), connected to a Pentium III computer. The AxioVision 2.0.5.3 service Pack 2 software (Carl Zeiss Vision GmbH, 1998, 1999) was utilised for acquiring images.

Character coding

The method of character coding adopted in a study is of importance to the outcome of the analyses performed on the dataset. Pleijel (1995) pointed out that if we assume consistency in the phylogeny reconstruction methodology, we should expect correctly identified homologies to form a pattern of congruence and convergence on the single true tree topology and false homologies to point in all possible directions and constitute noise in the analysis. However the assumption of randomness is violated if a number of characters are linked (Pleijel, 1995). For example presence of an operculum in the genera *Sphagnum* and *Tetraphis* is biologically linked to the transverse dehiscence of the capsule (i.e. a capsule dehisces transversely by means of an operculum) whereas possession of valves in *Andreaea* is linked to the longitudinal dehiscence (i.e. a capsule dehisces longitudinally to give valves). The use of all these as separate characters would overweight the evidence provided by this variation

(Wilkinson 1995). To avoid the problems of linked characters, Pimetel and Riggins (1987), have advocated multistate coding (Composite coding of Wilkinson, 1995). However where multistate characters have independent components (e.g. as in leaf shape), another form of coding, reductive coding has the advantage of making it easier for homoplasy to be revealed than under composite coding (Wilkinson 1995). Reductive coding is aimed at reducing complex multistate characters, where possible, by dividing the observed variation into meaningful independent (non-linked) binary characters (Pleijel, 1995 and Wilkinson, 1995). Reductive coding partitions the variation of character complexes into simpler characters, each describing variation in a particular component of a character complex (Wilkinson, 1995). Reductive coding for biologically independent characters also helps to avoid limiting the construction of internal nodes (Wilkinson, 1995).

The problem that may arise from reductive coding however is the escalation of inapplicable characters and their consequent problems (Nixon & Davies, 1991; Platnick et al., 1991). In some cases, this can be circumvented by addition of a presence-absence character (Type C of Pleijel, 1995). As much as possible, the method of coding adopted here is reductive coding (Wilkinson, 1995) equivalent to types C and D of Pleijel (1995). However, where multistate coding has the advantage outlined above it is the preferred coding option.

The coding utilised in the current analysis therefore has aspects of each of these approaches and is only an attempt to try and achieve the best coding for each character type, and thus avoid the treatment of every character as identical in nature to all other characters in the dataset.

Coding for Takakia

Takakia presented special problems for character coding. Regardless of recorded similarity to Andreaeopsida in certain characters, the structures of *Takakia* that are referred to as leaves may not be homologous to the leaves of other bryophytes and may have evolved independently from leaves in other bryophytes (Murray 1988a). However, even if these were homologous, it would be difficult to compare them with the rest of the taxa due to their unusual shape. *Takakia* has terete leaves whereas the

rest of the species studied have flat leaves, hence the difficulty in discerning homologies (e.g. which leaf cells are marginal and which ones are central). Many other characters were similarly difficult to code for *Takakia*, being specific to flat leaves, making it inevitable to code them as inapplicable for *Takakia*.

List and description of morphological characters

Some habit characters such as colour of the plant or leaves were not utilised as only herbarium specimens were used for the study. Fresh specimens that might give a good indication of some habit characters were generally not available. Other characters not included in this study are those pertaining to growth form and ecology. Due to the availability of herbarium specimens only, some of which have no such information, inclusion of these characters would result in much missing data. *Andreaeobryum* specimens could not be obtained for this study. However, detailed information on gametophyte and sporophyte characters was available in the literature (Murray 1988b). The distribution matrix for characters is given in Table 2.

Gametophyte characters

1. *Mucilage papillae on protonemata*: **(0)** Present; **(1)** Absent.

Mucilage papillae occur on the gametophores of *Andreaea*, *Andreaeobryum*, *Takakia* and *Tetraphis*. However these are present on protonemata only in *Andreaeobryum* (Murray 1987, 1988a) and *Takakia*. Mucilage papillae are absent in *Sphagnum*.

2. *Mucilage papillae on gametophores*: **(0)** Present, **(1)** Absent

In mosses the mucilage papillae present on gametophores are usually axillary (see also character 59). In *Andreaea* species, these axillary mucilage hairs consist of one short brownish basal cell and 1 or rarely 2 long hyaline apical cells (e.g. plates 1A and 1B). They are usually associated with perigonial leaves (Schofield and Hebant 1984). Axillary hairs are also present in the genera *Tetraphis*, *Andreaeobryum* and *Takakia*. However they are absent in *Sphagnum*. *Takakia lepidozioides* possesses 2 celled beaked mucilage papillae (plates 1C and 1D) that are similar to those in *Andreaeobryum* (Murray, 1988a). These are however, not only axillary (See description on character 1)

3. *Beaked mucilage papillae:* **(0)** Present; **(1)** absent.

Beaked mucilage papillae occur in *Takakia* (plates 1C and 1D) and *Andreaeobryum* (Murray, 1988a).

4. *Shape of protonemal appendages*: **(0)** terete (cylindric); **(1)** dorsiventral.

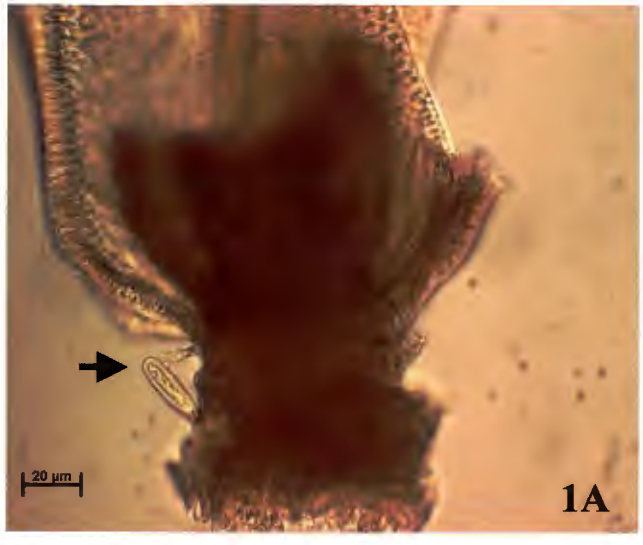
Protonemal appendages are frequent and terete in *Andreaeobryum*, whereas they are rare and dorsiventral in *Andreaea*, *Tetraphis* and *Sphagnum* (Murray 1988a). Protonemal appendages are also terete in *Takakia* (Smith and Davidson, 1993).

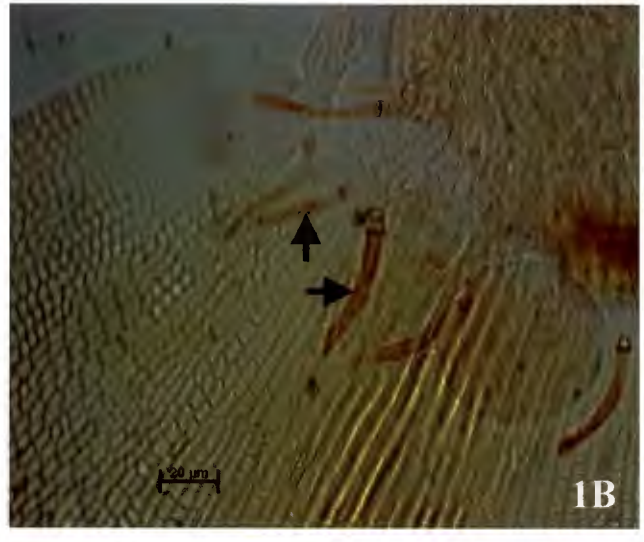
5. *Rhizoids on the stem*: **(0)** present; **(1)** absent.

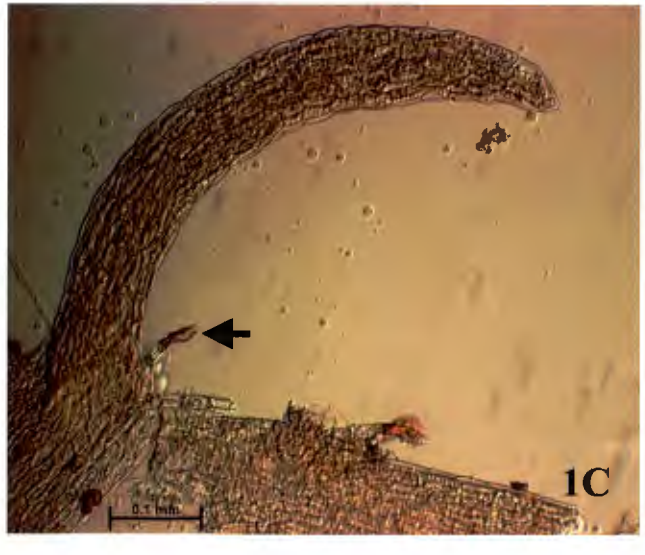
With the exception of *Sphagnum* and *Takakia* (Schofield and Hébant, 1984, Murray, 1988a), most moss gametophores (e.g. those of *Andreaea*, *Andreaeobryum and Tetraphis*) bear rhizoids on the stem (Schofield and Hebant 1984). In *Andreaea*, the basal creeping portions of the stems give rise to numerous rhizoids mostly consisting of cylindrical or plate like outgrowths, the former growing into the rock crevices, whereas the latter becoming applied closely to the rock surface (Cavers, 1911).

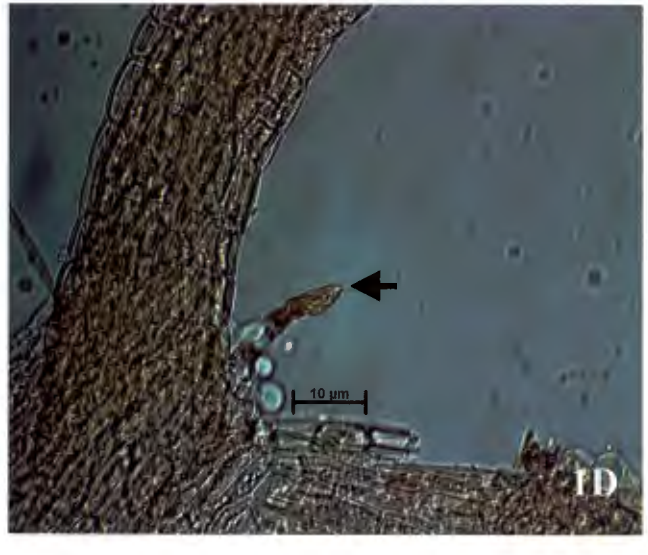
6. *Sexuality*: **(0)** monoicous; **(1)** dioicous.

More than 50% of mosses are dioicous (Schofield and Hebant 1984). Plants that are monoicous generally have the antheridia and archegonia in terminal groups, on separate branches (Parihar, 1965). Presence of both female and male sexual organs on the same plant definitely indicates a monoicous taxon. Where only one sex was


20 µm
1A


20 µm
1B


1C


10 µm
1D

Plates 1: Specimens were rehydrated prior to mounting in Hoyers' solution on microscope slides. Photographs were taken from the Leitz Diaplan compound microscope (Leitz Wetzler, German).

1A: *A. rothii*; axillary mucilage papilla on apical part of branch, with one basal quadrate cell and a broadly rectangular apical cell.

1B: *A. nitida*; axillary mucilage papillae, with one or two basal quadrate cells and a narrow elongated apical cell.

1C: *Takakia lepidozioides*; beaked mucilage papilla in axil of "leaf".

1D: *Takakia lepidozioides*; enlarged view of beaked mucilage papilla in axil of "leaf".

found the sexuality status was confirmed in literature. Most *Andreaea* species studied are monoicous and only a few are dioicous (i.e. *A. blyttii, A. nivalis, A. subulata* Harv, *A. gainii* Card, *A. australis, A. bistratosa* Magill, and *A. nitida*). *Andreaeobryum*, *Tetraphis* and *Takakia* are dioicous, whereas some species of *Sphagnum* are dioicous and others monoicous.

7. *Plants branched*: **(0)** Present; **(1)** Absent.

Except for *Tetraphis* species, all taxa studied are branched (e.g. plates 2A –2D).

8. *Type of Branching*: **(0)** monopodial; **(1)** sympodial.

In *Sphagnum* species, branching is sympodial, whereas *Andreaea* species are monopodially branched. *Andreaea* and *Andreaeobryum* species are mostly irregularly branched by subapical innovations.

LEAF CHARACTERS

In most of the studied genera, the mature main stem leaves and branch leaves are uniform. However in *Sphagnum* leaves on the main stem are different from the branch leaves. Therefore for all the specimens, leaf characters were determined from mature branch leaves.

9. *Leaf arrangement*: **(0)** spiral, regular and arranged singly; **(1)** spiral and irregular, in-groups of two or three.

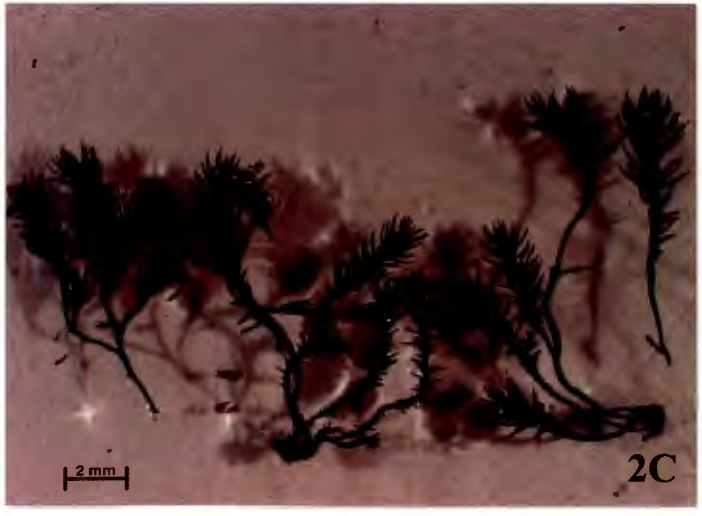
Takakia like *Andreaea*, *Andreaeobryum, Tetraphis* and *Sphagnum* has spirally arranged "leaves". However, unlike the others, the "leaves" of *Takakia* are arranged in loose irregular clusters of two to three (plate 3B). This irregular phyllotaxy (arrangement of leaves in regard to the axis) is one peculiar feature of the genus *Takakia* (Hattori & Mizutani 1958).

Characters 8-11 (Leaf Orientation)

Leaf orientation is a difficult character to code for many taxa due to its polymorphic nature within some of the individual species studied. However, two approaches taken here have circumvented this problem; first, the type of coding adopted (as outlined

2A

2B



2C

2D

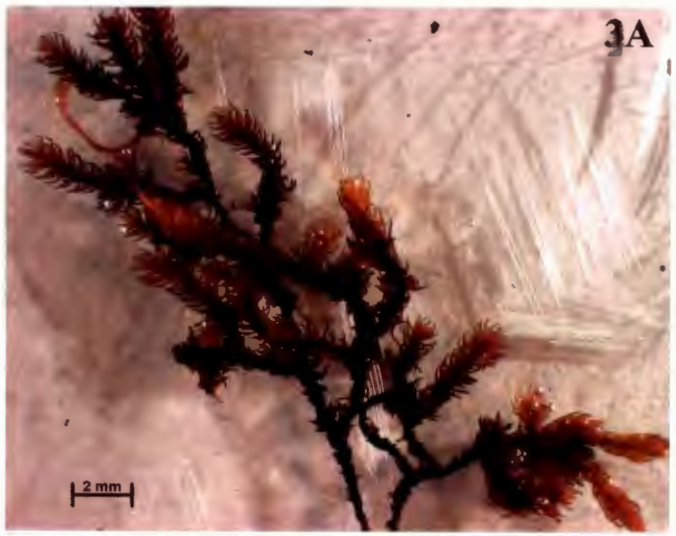
Plates 2: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

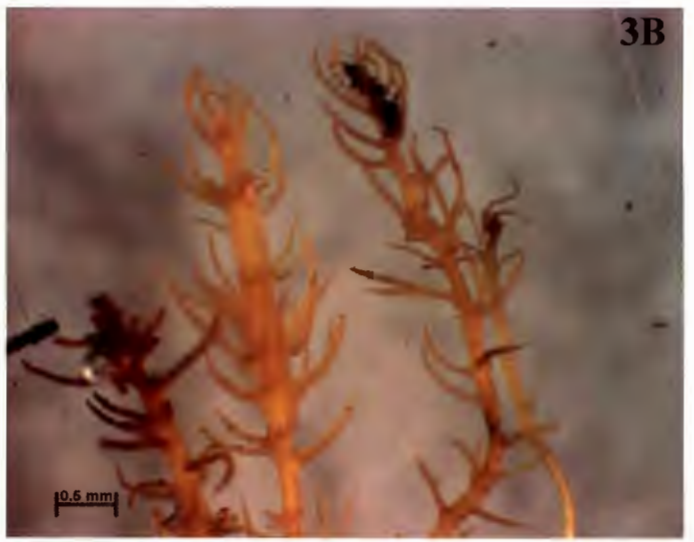
2A: habit of *A. nivalis* showing branched plant with spirally arranged, falcate second leaves.

2B: habit of *A. alpina* showing branched plant with spirally arranged, erect spreading to wide spreading leaves.

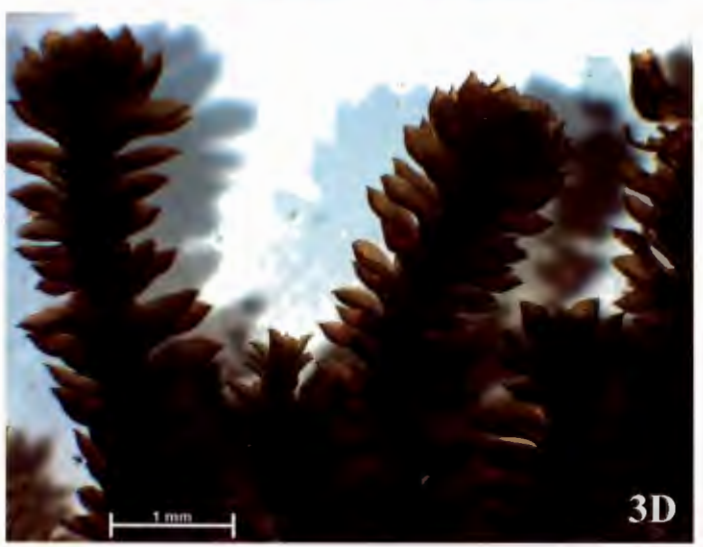
2C: habit of *A. bistratosa* showing branched plant with spirally arranged, wide spreading leaves.

2D: habit of *A. rupestris* var. *rupestris* showing branched plant with spirally arranged, erect spreading to wide spreading leaves.


3A
2 mm


3B
0.5 mm


.73 mm
3C


1 mm
3D

Plates 3: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

3A: habit of *A. wilsonii* showing branched plant with spirally arranged wide spreading leaves.

3B: habit of *Takakia lepidozioides* showing branched plant with an irregular phyllotaxy and terete "leaves".

3C: habit of *A. australis* showing branched plant with spirally arranged falcate leaves.

3D: habit of *A. nitida* showing branched plant with spirally arranged leaves that are wide spreading to squarrose.

above) and second, the consideration of the type of orientation occurring in at least approximately 75% of the leaves as the state for any particular specimen.
Plants with wide spreading to squarrose leaves were considered as those that have most of the leaves (at least ca. 75%) oriented between 60 degrees and 90 degrees. These include A. rupestris (plate 2D), *A. obovata*, *A. alpina* (plate 2B), *A. gainii*, *A. alpestris*, *A. wilsonii*, *A. sinuosa*, *A. acutifolia*, *A. bistratosa* (plate 2C) and *A. nitida*. *Takakia lepidozioides* also exhibits wide spreading to squarrose leaves that however, have an irregular orientation (plate 3B).

Plants with erect spreading to wide spreading leaves are those that have most of the leaves (at least ca. 75%) oriented at about 15 to 60 degrees. These include species of *Tetraphis*.

Secund leaves are those that arise spirally round the stem with most of them (at least ca. 75%) inclined towards the same direction (e.g. plates 2A and 4D). This tendency is especially strong in taxa such as *A. nivalis* Hook. and mostly weak in taxa such as *A. blyttii* Schimp, *A. subulata* and *A. frigida* Huebener.

Imbricate leaves that are more or less stem clasping, are found in the genus *Sphagnum*, whereas species of the genus *Tetraphis* exhibit erect spreading to wide spreading leaves.

10. *Leaves wide spreading to squarrose*: **(0)** Present; **(1)** Absent.

11. *Imbricate leaves*: **(0)** Present; **(1)** Absent.

12. *Mostly erect spreading to wide spreading leaves*: **(0)** Present; **(1)** Absent.

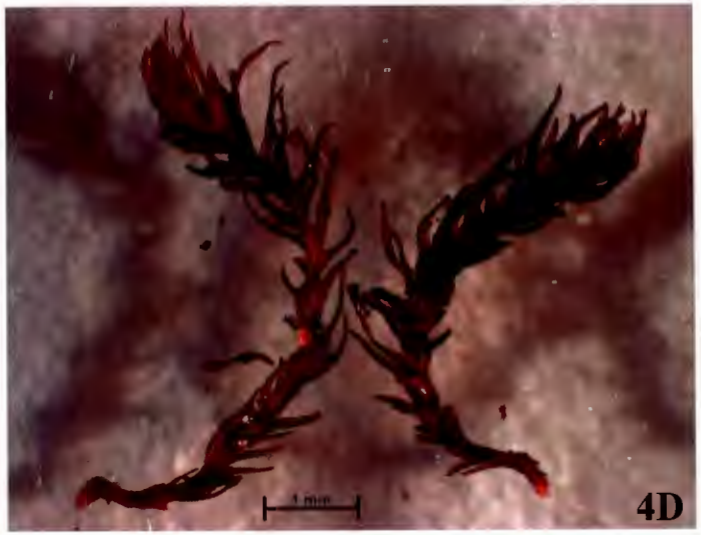
13. *Secund leaves*: **(0)** Present; **(1)** Absent.

14. *Flat leaves at least partly appressed to the stem*: **(0)** Present; **(1)** Absent.
Many species of *Andreaea* e.g. *A. rupestris*, *A. obovata* and *A. megistospora* have leaves with about the basal 1/2 of the leaf appressed to the stem. Leaves of *A. nivalis*, *A. subulata*, *A. australis* and *A. frigida* are not basally appressed to the stem.


2 mm
4A


2 mm
4B


2 mm
4C


1 mm
4D

Plates 4: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

4A: habit of *A. subulata* showing plant with spirally arranged falcate-secund leaves.

4B: habit of *A. frigida* showing branched plant with spirally arranged falcate leaves.

4C: habit of *A. megistospora* ssp. *megistospora* showing small, branched plant with spirally arranged leaves.

4D: habit of *A. schofieldiana* showing branched plant with spirally arranged falcate-second leaves.

However *A. nivalis* leaves apparently expose more of the stem than those of other *Andreaea* species. *A. frigida* leaves, though appressed, appear intermediate in this character between basally appressed leaves and those that are not. Leaves of *Sphagnum* are entirely appressed to the stem

15. *Extent to which leaves are appressed to the stem*: **(0)** Up to ½ the basal part of the leaves; **(1)** entirely.

See description for character 12.

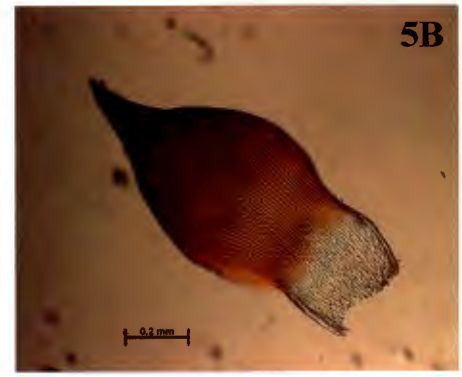
16. *Form of leaf lamina*: **(0)** Flat; **(1)** Terete.

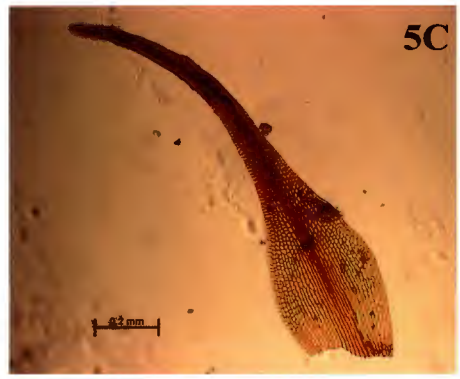
The outgroup taxon *Takakia lepidozioides* has terete stems, branches and leaves (Plate 3B). Terete leaves are one of the several features that have been considered peculiar to *Takakia* among Bryophytes (Hattori & Mizutani 1958). All other mosses have flat leaves (e.g. plates 5A-5H).

Characters 17 — 25 (Leaf Shape)

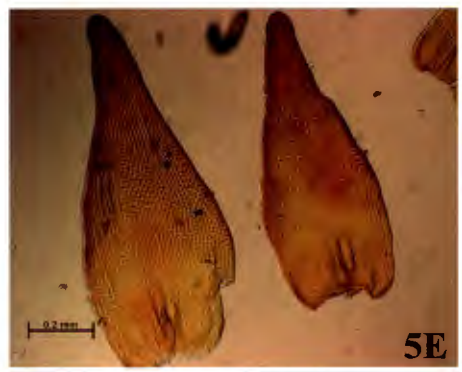
Leaves in *A. blyttii*, *A. nivalis*, *A. subulata*, *A. rothii* Web et Mohr, *A. crassinervia* Bruch, *A. megistospora* and *A. schofieldiana* are falcate (i.e. have a tendency to be curved like a sickle; see plates 2A, 4A, 4D, 5A and 5C). Leaves of *Tetraphis pellucida* Hedw. and *T. geniculata* Girg. ex Milde though not curved, are oblique. Panduriform leaves (e.g. plate 5B) are found in *A. obovata*, *A. alpina* and *A. gainii*. Mature leaves in *A. obovata* are actually only very slightly panduriform, but reflexed backwards and slightly in-rolled at the "hip" and the apical part curved slightly forwards. This is what gives them a panduriform appearance (with the wider part above rather than below) when observed dorsiventrally, without flattening (i.e. without flattening with a cover slip). When flattened, most of the leaves are actually more ovate than panduriform. *A. alpina* and *A. gainii* have clearly panduriform leaves with the apical part of leaf broader than the basal part, and strongly reflexed backwards at the hip and forward apically. The panduriform shape is still clearly apparent in *A. alpina* and *A. gainii*, even when flattened.

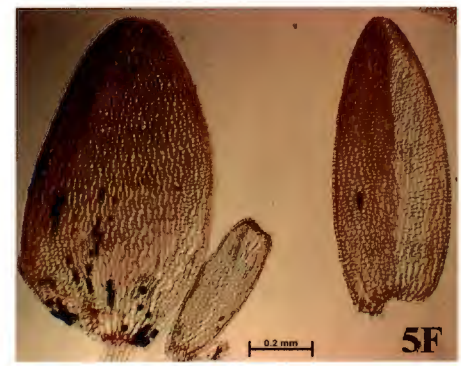

5A
0.2 mm

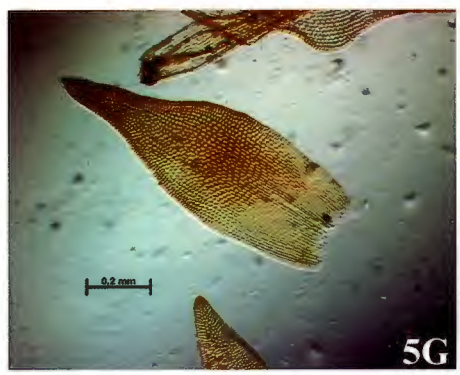

5B
0.2 mm

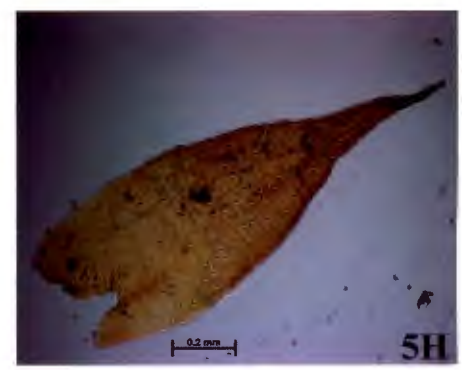

5C
0.2 mm


5D
0.2 mm


0.2 mm
5E


0.2 mm
5F


0.2 mm
5G


0.2 mm
5H

Plates 5: Specimens were rehydrated prior to mounting in Hoyers' solution on microscope slides. Photographs were taken from the Leitz Diaplan compound microscope (Leitz Wetzler, German).

5A: *A. rothii*, falcate ovate lanceolate, subulate leaf

5B: *A. alpina*, panduriform leaf with broad apical part.

5C: *A. blyttii*, falcate ovate lanceolate subulate leaf with basally indistinct costa.

5D: *A. nitida*, obovata leaf, with umbonate apex and costa fanning out apically.

5E: *A. wilsonii*, oblong lanceolate leaf with narrowly obtuse apex.

5F: *Sphagnum pylaesii*, broadly ovate leaves with truncate tips.

5G: *A. rupestris* var. *rupestris*, oblong lanceolate with a non-subulate apex that is more abruptly narrowed from an ovate base than in var. *papillosa*.

5H: *A. rupestris* var. *papillosa*, oblong lanceolate with a non- subulate apex that is gradually narrowed from an ovate base.

Leaves of *A. rupestris*, *A. alpestris* and *A. wilsonii* are oblong lanceolate [e.g. plate 5E) with var. *rupestris* having a non subulate apex that is relatively abruptly narrowed from an ovate base whereas var. *papillosa* has a non subulate apex that is gradually narrowed from an ovate base.

Ovate lanceolate leaves occur in *A. subulata*, *A. acutifolia*, *A. blyttii*, *A. nivalis*, *A. rothii*, *A. crassinervia*, *A. megistospora* and *A. frigida*. *A. bistratosa* is slightly constricted in about the lower third of the leaf. However the overall appearance is very similar to the linear lanceolate leaves in *A. sinuosa*. *Andreaeobryum macrosporum* leaves are also linear lanceolate. Concave, broadly boat shaped (cymbiform) leaves (plates 3D and 5D) occur in *A. nitida* J. D. Hooker et Wilson, whereas *A. australis* has oblong leaves with a rugose appearance.

Tetraphis species have ovate lanceolate oblique leaves and broadly ovate leaves are found in *Sphagnum* (plate 5F).

17. *Flat falcate leaves*: **(0)** Present; **(1)** Absent.

18. *Ovate lanceolate leaves:* **(0)** Present; **(1)** Absent.

19. *Linear lanceolate leaves*: **(0)** Present; **(1)** Absent.

20. *Oblong lanceolate leaves*: **(0)** Present; **(1)** Absent.

21. *Oblong leaves*: **(0)** Present; **(1)** Absent

22. *Lanceolate oblique leaves*: **(0)** Present; **(1)** Absent.

23. *Cymbiform leaves*: **(0)** Present; **(1)** Absent.

24. *Broadly ovate leaves*: **(0)** Present; **(1)** Absent.

25. *Panduriform leaves*: **(0)** Present; **(1)** Absent.

26. *Long subula:* **(0)** Present; **(1)** Absent.

A long subula was considered present in leaves that have a long narrow apical part that is at least half the length of the leaf (e.g. plates 5A and 5C). A long subula is present in *A. subulata*, *A. acutifolia*, *A. blyttii*, *A. nivalis*, *A. rothii* and *A. crassinervia* *A. megistospora* and *A. schofieldiana* and *A. frigida*.

27. *Margins of flat mature leaves*: **(0)** with 'entire' portions; **(1)** at least partially crenate to serrate.

Leaves of all species were observed for this character at x 160 magnification. In *A. rupestris* ssp. *papillosa* the apex appears crenulate only due to papillose cells. *A. alpina* is serulate in about the basal half of the leaf and entire apically. *A. gainii* is mostly serrulate basally, but very rarely apically as well. *A. nivalis* is serrulate throughout. This is especially distinct in 'newly formed' mature leaves. The rest of the species have leaves with entire margins.

CHARACTERS 28-34 (LEAF APICES)

Though most of the species of *Andreaea* studied have acuminate apices the range of variation includes acute, acuminate, obtuse, cuspidate and umbonate apices.

Species with apices greater than 45° and less than 90° were considered as acute, whereas acuminate ones were considered as those gradually narrowing to a long sharp point at less than 45 degrees. *A. rupestris* ssp. *rupestris* and *A. alpestris* have sub-acute leaf apices drawn from a non-subulate and narrow, more basal part.

Cuspidate apices occur in *A. alpina* and *A. gainii*. *A. nivalis* has apical cells that gives a more or less triangular shape.

A. *nitida* has convex leaves with abrupt, rounded, central points (umbonate).

Narrowly obtuse apices continuing from a broad basal part are present in *A. obovata*, *A. wilsonii* and *Andreaeobryum macrosporum*.

For *Sphagnum girgenshonii* Russ, the leaf apices (for branch leaves) generally appear similar to acuminate apices but are narrowly obtuse, truncate and 3-fid split tips whereas *S. pylaesii* Brid. has leaf apices that are broadly truncate and 5-pointed.

28. *Leaves with acuminate apices*: **(0)** Present; **(1)** Absent.

29. *Leaves with cuspidate apices*: **(0)** Present; **(1)** Absent.

30. *Leaves with umbonate apices*: **(0)** Present; **(1)** Absent.

31. *Apices narrowly obtuse from a broad relatively basal part*: **(0)** Present; **(1)** Absent.

32. *Leaf apices narrowly obtuse from a narrow relatively basal part*: **(0)** Present ; **(1)** Absent.

33. *Leaf apices sub-acute from a non-subulate narrow more basal part*: **(0)** Present; **(1)** Absent.

34. *Leaf apices truncate:* **(0)** Present; **(1)** Absent.

35. *Truncate apex*: **(0)** with a tri-fid split; **(1)** with a five-fid split.

36. *Leaf costae on flat leaves*: **(0)** present; **(1)** Absent.

Some *Andreaea* species e.g. *A. nivalis, A. blyttii, A. rothii, A. frigida, A. crassinervia, A. megistospora, A. australis* and *A. nitida* are nerved (costate) whereas other species such as *A. rupestris, A. obovata, A. alpina, A. alpestris, A. wilsonii, A. gainii* and *A. bistratosa* lack a costa (e.g. plates 5G and 5H). Apparently intermediate forms are present in which the nerve is sometimes not discernible in the basal part e.g. *A. blyttii* (see also character 41). *Andreaeobryum macrosporum* and *Tetraphis* species are costate, whereas *Sphagnum* species are ecostate.

37. *Costa prominence*: **(0)** strong: **(1)** weak.

Weak costae are considered those that are not easily distinct in the leaf sometimes difficult to notice in certain parts of the leaf blade, for some of the leaves. Strong costae are considered here as those that are always easy to see in all parts of the leaf where they occur and on all mature leaves. Most of the costate species studied here

have strong costae. The costae in *A. blyttii* and *A. subulata* are weak. In *A. blyttii* it is so faint that it almost fades off especially at the base. This is also characteristic of *A. heneimanii* Hampe & C. Mull. (not available for this study).

38. *Costa fanning out and faded apically*: **(0)** yes: **(1)** No.

A. nitida has a broad costa that characteristically fans out apically (plate 5D).

39. *Decurrent costa*: **(0)** Present: **(1)** Absent.

A decurrent costa is present in *A. australis* and the genus *Tetraphis*.

40. *Apical part of costa*: **(0)** extending to just below apex; **(1)** Percurrent; **(2)** Excurrent.

In *A. rothii*, *A. frigida*, *A. nitida* and the outgroup taxon *Tetraphis* the costa ends just below the apex, whereas *A. blyttii, A. crassinervia* and *Andreaeobryum macrosporum* have excurrent costae. Percurrent costae occur in *A. nivalis*, *A. schofieldiana*, *A. subulata*, *A. australis* and *A. megistospora*. The rest of the species are ecostate.

41. *Costa that is sometimes wanting basally*: **(0)** Present; **(1)** Absent.

A. blyttii (plate 4C) , exhibits a costa is sometimes not apparent in the basal part of the lamina. This feature is also characteristic of *A. heinemanii* (Murray, 1988b), which however was not available for this study.

42. *Lamina layers of flat leaves*: **(0)** unistratose; **(1)** bistratose.

A. rupestris, *A. obovata*, *A. australis* and *A. nitida* have unistratose leaves. *A. blyttii*, *A. alpestris*, *A. nivalis*, *A. subulata*, *A. alpina, A. rothii, A. wilsonii, A. megistospora*, *A. gainii*, and *Andreaeobryum macrosporum* have uni and bi -stratose parts of leaves, whereas *A. crassinervia* has uni to multistratose parts of the leaves. *A. bistratosa* has uniformly bistratose leaves. *Takakia* has terete leaves that are therefore considered multistratose. Both *Tetraphis* and *Sphagnum* species have unistratose leaves

43. *Leaves with multistratose areas*: **(0)** Present; **(1)** Absent.

See description for character 42.

44. *Pitted leaf cells:* **(0)** present; **(1)** absent.

Leaves of *A. rupestris*, *A. obovata*, *A. alpestris, A. alpina*, *A. wilsonii*, *A. frigida*, *A. gainii*, *A. acutifolia* and *A. sinuosa* have clearly pitted leaf cells. *A. wilsonii* appears to have basal cells that are pitted both laterally and longitudinally (plate 6A) In certain cases, it appears that some cells have actually fused. *A. sinuosa* usually has more than one lateral pit per cell (plate 6C) The rest of the *Andreaea* species in the study, including *Andreaeobryum macrosporum*, *Tetraphis* and *Sphagnum* species, have cells that are not pitted.

Characters 45 — 50 (Basal Marginal Cells)

Basal marginal cells are longitudinally rectangular and followed by laterally rectangular cells in *A. obovata*, *A. alpestris* and *A. gainii*, whereas they are rectangular and immediately followed by quadrate cells in *A. rupestris*, *A. alpina, A. blyttii*, *A. sinuosa*, *A. acutifolia*, *A. bistratosa* and *A. wilsonii. A. australis* and *A nitida* have laterally rectangular marginal cells throughout the entire length of the leaves (plates 6B and 6C). Quadrate basal marginal cells occur in *A. nivalis*, *A. subulata*, *A. rothii*, *A. frigida*, *A. crassinervia*, *A. schofieldiana*, *A. nitida*, *A. megistospora* and *Andreaeobryum macrosporum.* All leaf cells appear rectangular in the outgroup *Takakia lepidozioides*, whereas they are rhomboidal in *Sphagnum* species.

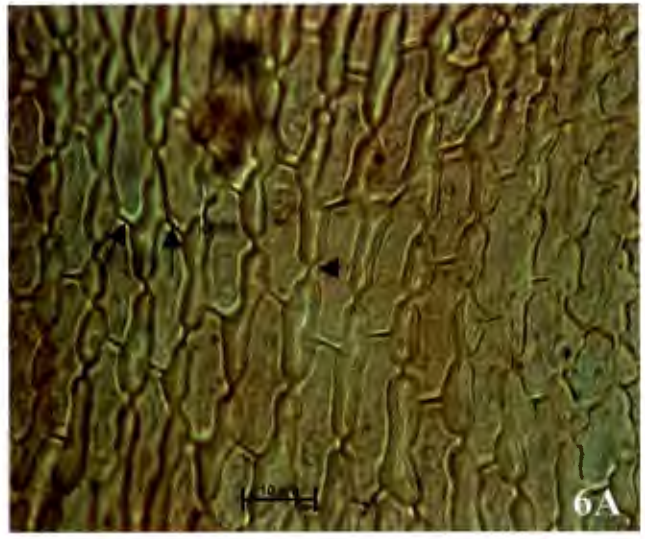
45. *Quadrate basal marginal cells*: **(0)** Present; **(1)** Absent.

46. *Rhomboidal basal marginal cells*: **(0)** Present; **(1)** Absent.

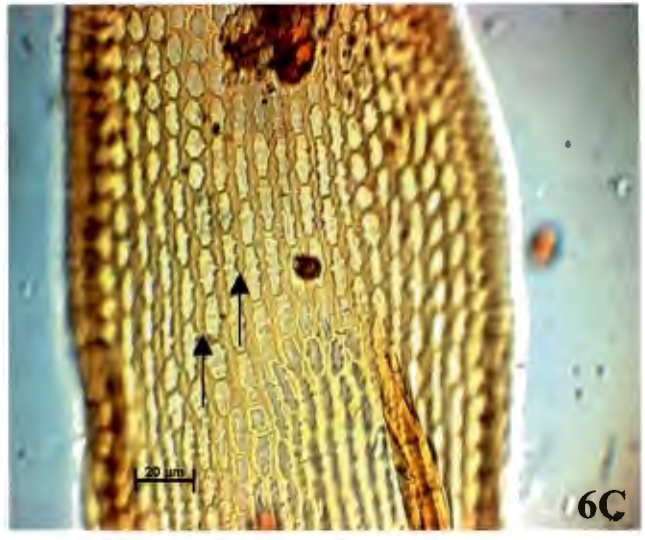
47. *Longitudinally rectangular basal marginal cells*: **(0)** Present; **(1)** Absent.

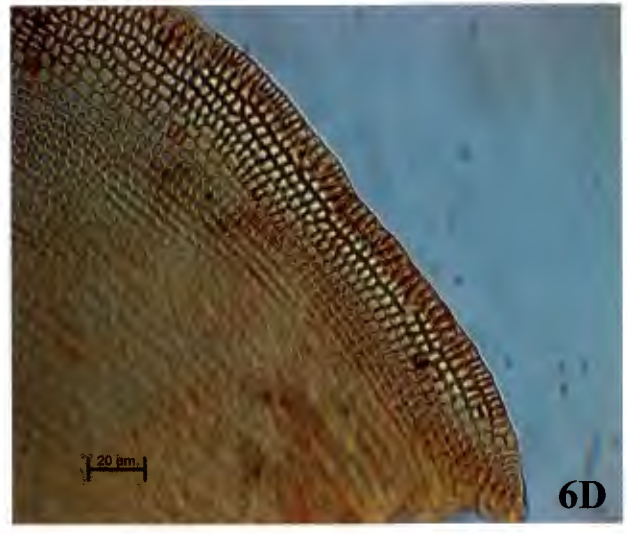
48. *Laterally rectangular cells following longitudinally rectangular basal marginal cells.* **(0)** Present; **(1)** Absent.

49. *Quadrate cells immediately following longitudinally rectangular basal marginal cells*: **(0)** Present; **(1)** Absent.


6A


0.1 mm
6B


20 µm
6C


20 µm
6D

Plates 6: Specimens were rehydrated prior to mounting in Hoyers' solution on microscope slides. Photographs were taken from the Leitz Diaplan compound microscope (Leitz Wetzler, German).

6A: *A. wilsonii*, showing pitted leaf cells that are apparently pitted both laterally and longitudinally.

6B: *A. australis* showing quadrate central leaf cells and laterally rectangular marginal cells.

6C: *A. sinuosa* showing sinuous basal leaf cells with multiple lateral pits.

6D: *A. nitida* showing quadrate central leaf cells and laterally rectangular marginal cells.

50. *Laterally rectangular basal marginal cells*: **(0)** Present; **(1)** Absent.

Characters 51 — 54 (Central Basal Leaf Cells)

Central basal leaf cells were considered here as those found in the basal third of the leaf between the costa and the margin.

A. rupestris, *A. obovata*, *A. alpestris*, *A. alpina*, *A. wilsonii*, *A. gainii*, *A. bistratosa*, *A. acutifolia* and *A. sinuosa* have rectangular basal central cells. The rectangular basal central cells in *A. sinuosa* are characteristically sinuous. In *A. rothii*, *A. frigida*, *A. crassinervia*, *A. schofieldiana*, *A. megistospora*, *A. australis* and *A. nitida*, basal central cells are rounded quadrate or hexagonal. *A. nivalis* has only quadrate or short rectangular basal central cells. *Andreaeobryum* and *Tetraphis* also have rounded or hexagonal basal central cells whereas *Sphagnum* species have sigmoid to fusiform basal central leaf cells. In *A. blyttii*, *A. subulata*, *A. frigida* and *A. schofieldiana*, juxtacostal basal cells are rectangular whereas the other central basal cells are quadrate, rounded or hexagonal.

51. *Most basal central leaf cells rectangular:* **(0)** Present; **(1)** Absent.

52. *Mostly rounded or hexagonal basal central leaf cells:* **(0)** Present; **(1)** Absent.

53. *Quadrate basal central leaf cells:* **(0)** Present; **(1)** Absent.

54. *Sigmoid to fusiform basal central leaf cells:* **(0)** Present; **(1)** Absent.

55. *Quadrate, hexagonal or rounded cells that are laterally adjacent to Juxtacostal rectangular central basal leaf cells*: **(0)** Present; **(1)** Absent.

56. *Sinuous leaf cells:* **(0)** Present; **(1)** Absent.

57. *Rhomboidal cells forming network with chlorophyllous cells*: **(0)** yes; **(1)** no.

Rhomboidal cells forming network with chlorophyllous cells are a characteristic of the genus *Sphagnum*.

58. *Transition in cell type from leaf base to apex:* **(0)** abrupt or sub-abrupt; **(1)** gradual.

A number of *Andreaea* species have cells that differ in the basal part of the leaf from those of the apical (distal) part. The transition between the different cell types ranges from gradual to abrupt. An abrupt change is considered here as change from one cell type to another within one or two cells distance, whereas sub-abrupt is considered change within at least about 5 cells, though not necessarily at the same level along the width of the leaf. The transition is considered gradual if within more than about five cells distance with a wider region of mixed or intermediate shaped cells.

An abrupt or sub-abrupt transition of cell shape, as described above, is found only in *A. blyttii.* The transition is gradual in *A. rupestris*, *A. obovata*, *A. alpestris*, *A. wilsonii*, *A. alpina*, *A. gainii*, *A. bistratosa*, *A. acutifolia* and *A. sinuosa*.

In *A. subulata*, *A. rothii*, *A. crassinervia*, *A. megistospora*, *A. schofieldiana*, *A. frigida* and *A. nivalis*, *A. australis*, *A. nitida*, and *Andreaeobryum macrosporum*, including the outgroup taxa *Tetraphis* and *Sphagnum*, there is basically no transition in cell shape though a change in size is usually apparent. This character was therefore coded as inapplicable for these taxa.

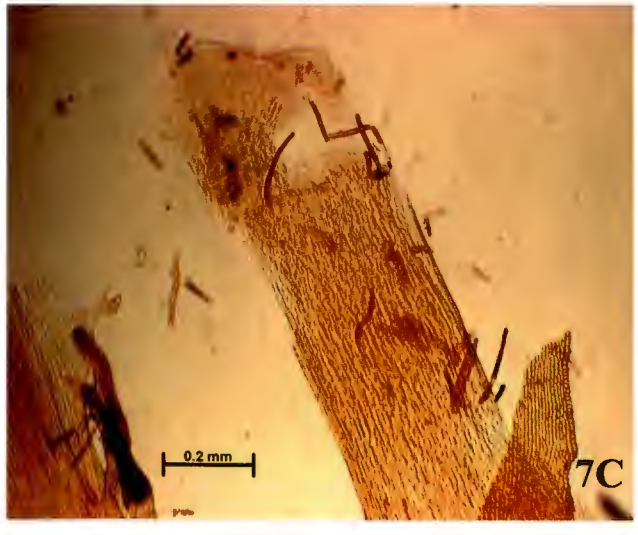
59. *Axillary hairs*: **(0)** Present; **(1)** Absent.

Axillary hairs are present in *Andreaea, Andreaeobryum*, *Takakia* and *Tetraphis*. In most species of *Andreaea* e.g. *A. rupestris*, *A. alpina* and *A. rothii*, these axillary hairs consist of one short brownish basal cell and one long hyaline apical cell. *Takakia Lepidozioides* and *Andreaeobryum* posses 2 celled beaked mucilage papillae that are however not only axillary. In *Sphagnum*, axillary hairs are absent (Schofield and Hebant 1984).

60. *Mucilage hairs on upper pseudopodium*: **(0)** present; **(1)** absent.

7A

7B



7C

7D

Plates 7: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

7A: *A. rupestris* var. *papillosa*; differentiated, sheathing and convolute perichaetial leaves.

7B: *A. nitida*, sheathing but non-convolute perichaetial leaves, pseudopodia and capsule.

7C: *A. nitida*, mucilage hairs on pseudopodium.

7D: *A. frigida*, differentiated, sheathing and convolute perichaetial leaves surrounding pseudopodia.

In most *Andreaea* species, the axillary hairs occur only a short way up the lower part of the pseudopodium. However in *A. nivalis* and *A. nitida* (plate 7C), axillary hairs are also present on the upper part of the pseudopodium. Unfertilised archegonia and small bracteoles also occur high on the pseudopodium of *A. nivalis* and *A. nitida* (Murray, 1988b). Other *Andreaea* species lack any bracteoles or leaves on the pseudopodium.

Characters 61 — 65 (Perichaetial Leaves).

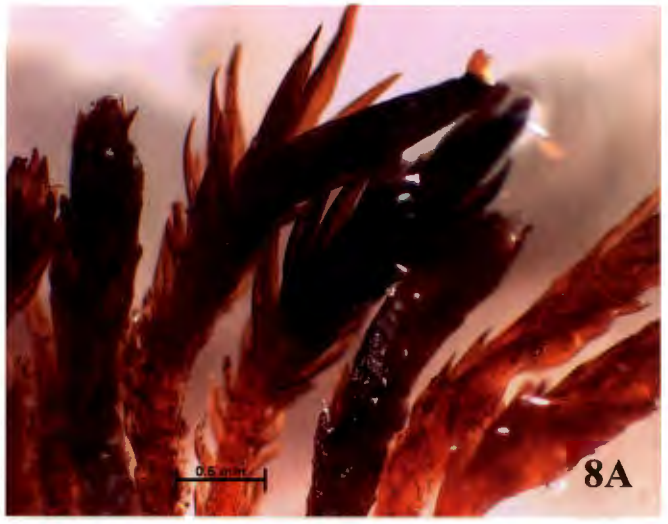
In *Andreaea* each group of sexual organs is surrounded by a number of large involucral leaves or perichaetial leaves (Schofield, 1985). In most species of *Andreaea*, the perichaetia and capsule develop at the terminal part of the stem (acrocarpous as defined by La Farge-England, (1996)) and new branches develop by means of sub apical innovations. *Tetraphis* species are also acrocarpous whereas *Sphagnum* species are cladocarpous (sexual organs develop only on lateral branches with juvenile development similar to that on main branches (La Farge-England, 1996).

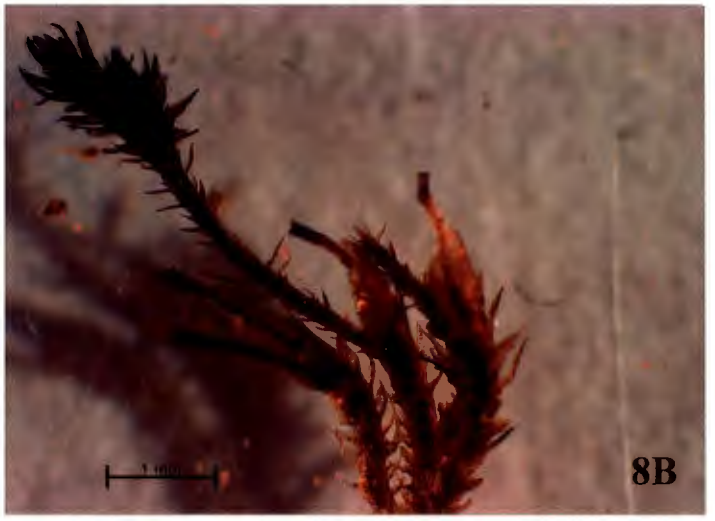
Perichaetial leaves in many *Andreaea* species are generally larger and differentiated from the mature leaves (e.g. plates 7A, 8A, 8B and 8D). However, in *A. nivalis* and *A. australis* the perichaetial leaves though larger than the mature leaves are not differentiated from them (plate 8C). Perichaetial leaves are sheathing and convolute in most *Andreaea* species whereas they are sheathing but not convolute in *A. subulata*, *A. nitida* (plate 7B) and *A. australis* (plate 7D and 9D). However, in *A. nivalis*, they are neither convolute nor sheathing. They are sheathing in the outgroup *Sphagnum* but spreading to squarrose and flexuous in the *Tetraphis*. Unlike in most mosses the perichaetium of *Andreaea* is developed in part prior to fertilization, whereas that of *Andreaeobryum* is developed after fertilization (Murray, 1988a).

61. *Perichaetia Position*: **(0)** acrocarpous; **(1)** pleurocarpous.

62. *Differentiated perichaetial leaves*: **(0)** Present; **(1)** Absent.

63. *Arrangement of flat perichaetial leaves*: **(0)** sheathing; **(1)** non-sheathing.


0.5 mm
8A


8B


8C


1 mm
8D

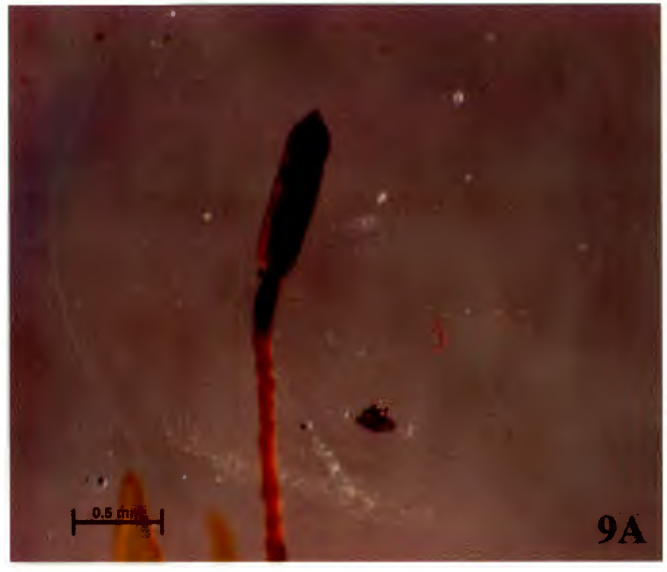
Plates 8: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

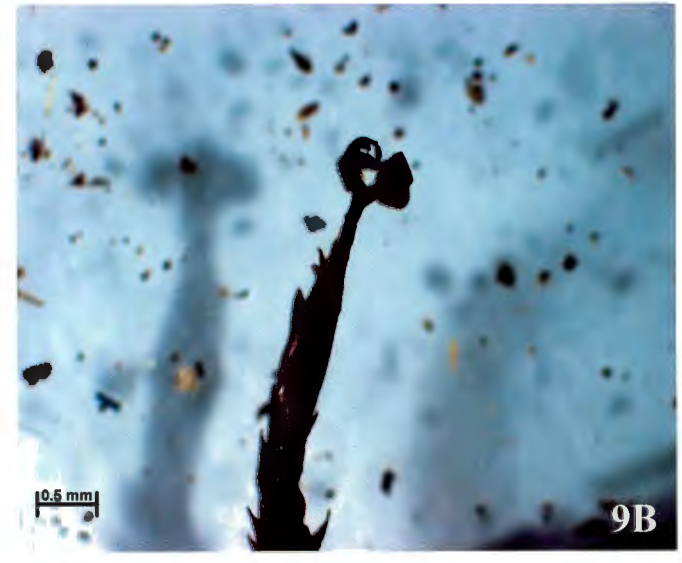
8A: *A. sinuosa*; differentiated, sheathing and convolute perichaetial leaves.

8B: *A. acutifolia*; differentiated perichaetial leaves and pseudopodia.

8C: *A. nivalis*; undifferentiated, non-sheathing and non-convolute perichaetial leaves.

8D: *A. wilsonii*; differentiated convolute and sheathing apical leaves.


0.5 mm
9A


0.5 mm
9B


0.5 mm
9C


0.5 mm
9D

Plates 9: Specimens were mounted in water for easy examination. Photographs were taken from the Zeiss M400 dissecting microscope (Carl Zeiss Vision).

9A: *A. wilsonii*; cylindrical capsule, elevated on elongated pseudopodium.

9B: *A. alpina*; dehisced four-valved capsule.

9C: *A. australis*; non-convolute perichaetial leaves and ovoid capsule elevated on pseudopodium.

9D: *A. australis*; dehisced four-valved capsule.

64. *Convolute perichaetial leaves*: **(0)** Present; **(1)** Absent.

65. *Development of perichaetia prior to fertilisation:* **(0)** Present; **(1)** Absent.

66. *Paraphyses among antheridia*: **(0)** Present; **(1)** Absent.

These are numerous filamentous paraphyses among the antheridia in species of *Andreaea* as well in *Tetraphis*. Paraphyses are, however, absent in *Takakia*, *Andreaeobryum* (Schofield, 1985; Murray, 1988a) and *Sphagnum* (Schofield and Hebant 1984).

Sporophyte characters

Species of the genus *Andreaea* fruit quite frequently (Scott, et al., 1976). The sexual organs are borne in terminal groups on separate branches usually on the same plant. (Cavers, 1911). The sporangium of *Andreaea* has an interesting mixture of characters, some of which are peculiar to the genus whereas other structural and embryological features of the sporophytes are indeed shared with the Sphagnales and the Bryales (Cavers, 1911; Schofield, 1985; Murray, 1988a).

67. *Structure elevating the capsule:* **(0)** seta; **(1)** Pseudopodium.

The capsule in mosses is usually supported by a structure called a seta, which is a part of the sporophyte, formed from the "seta meristem" through a brief period of intercalary division (Crandall-Stotler, 1984) and thus clearly differentiated from the main stem. The seta is present in *Andreaeobryum* and *Tetraphis* but absent in *Andreaea* and *Sphagnum*. Instead, the capsule in *Andreaea* and *Sphagnum* is elevated on a prolongation of the gametophyte, the pseudopodium. Though the pseudopodium resembles a seta in appearance as well as in function (Parihar, 1965), it continues from the vegetative stem without a joint (Scott, et al., 1976; Schofield 1985).

68. *Length of pseudopodium*: **(0)** less than 4mm; **(1)** more than 4 mm **(2)** more than 6 mm.

All the species of *Andreaea* studied except *A. wilsonii* have pseudopodia that are less than 4 mm long. *Sphagnum girgensohnii* has a pseudopodium that is between 4 and 6 mm long. The pseudopodium in *A. wilsonii* and *S. pylaesii* is more than 6 mm long.

69. *Geniculate seta*: **(0)** Present; **(1)** Absent.

Unlike in *Tetraphis pellucida* and *Andreaeobryum macrosporum*, the seta of *Tetraphis geniculata* is geniculate, serrate or papillose above the geniculate point and smooth below (Murray, 1988b).

Characters 70 — 73 (Shape of the capsule)

Most *Andreaea* species have ovoid capsules (plate 9D). However, *A. wilsonii* (plate 9A) and *Tetraphis* have cylindrical capsules with tapered ends. In *Sphagnum*, the capsules are globose (barrel shaped). *Andreaeobryum macrosporum* possesses an obtuse conic capsule (Murray 1988a).

70. *Ovoid capsules*: **(0)** Present; **(1)** Absent.

71. *Cylindrical capsules*: **(0)** Present; **(1)** Absent.

72. *Globose capsule*: **(0)** Present; **(1)** Absent.

73. *Obtuse conic capsule*: **(0)** Present; **(1)** Absent.

74. *Capsule dehiscence*: **(0)** Transversely by means of an operculum; **(1)** Longitudinally by means of valves.

Andreaea, *Andreaeobryum* and *Takakia* differ from the Sphagnales and other mosses in capsule dehiscence, which is longitudinal (e.g. plates 9B and 9C) and resembles that of some Hepaticae (Sim, 1926 and Schofield, 1985). In *Andreaea*, the dehiscence occurs along rows of thin walled suture cells (Parihar, 1965, Murray, 1988a), spores being shed via the gradual splitting and closing of the capsule by means of longitudinal valves. This occurs when the capsule elongates and shortens in response to changing moisture conditions. In the genus *Andreaeobryum* however, there is little

change in shape with changes in moisture (Murray, 1988a). Suture cells are absent in *Takakia* (Smith and Davidson, 1993). An operculum is absent in the genera *Andreaea*, *Andreaeobryum* and *Takakia* (see also character 76). Among the studied genera, the operculum is present only in *Tetraphis* and *Sphagnum*.

75. *Extent of longitudinal capsule dehiscence*: **(0)** within upper third; **(1)** least from lower 1/3 and up to near the apex (below upper ¼).

In most of the *Andreaea* species studied the capsule splits at least from lower 1/3 and up to near the apex (below upper ¼). In *A. wilsonii*, however, dehiscence is confined to the upper third of the capsule.

76. *Peristome*: **(0)** Present; **(1)** Absent.

In a number of mosses a single or double circle of teeth, the peristome, occupies the inside of the capsule beneath the operculum, surrounding its mouth after operculum has been shed (van Rooy, 1997). This structure is absent in *Andreaea* and the outgroup taxon *Sphagnum* but present in *Tetraphis*.

Characters 77 — 80 (The Calyptra)

In mosses a structure called the calyptra (which is strictly speaking, not a sporophyte character on account of its gametophytic origin) forms a sheathing cap that protects the elongating tip of the sporophyte (Schofield and Hébant, 1984). Usually as the moss capsule matures, the calyptra is ruptured and its upper portion carried on the apex of the capsule (Parihar, 1965). In most mosses the calyptra persists until maturity (Schofield, 1985). However in *Andreaea*, the mitrate calyptra ruptures early, before the sporangium is mature (Schofield, 1985). The upper part of this calyptra (in *Andreaea*) is stunted covering only less than 1/3 of the capsule, whereas in *Sphagnum* it is larger and tattered covering about 1/3 or more, but not the entire capsule (Janzen, 1917; Murray, 1988a). The calyptra of *Takakia* is mitriform and covers the upper ¼ or less of the capsule (Smith and Davidson, 1993), whereas that of *Andreaeobryum* and *Tetraphis*, covers the entire capsule (Sim 1926, Murray 1988a).

77. *Sporophyte development to maturity, within calyptra*: **(0)** Present; **(1)** Absent.

78. *Stunted calyptra, covering less than apical 1/3 of capsule*: **(0)** Present; **(1)** Absent.

79. *Tattered and large calyptra, covering upper 1/3 to 1/2*: **(0)** Present; **(1)** Absent.

80. *Large calyptra, covering entire capsule*: **(0)** Present; **(1)** Absent.

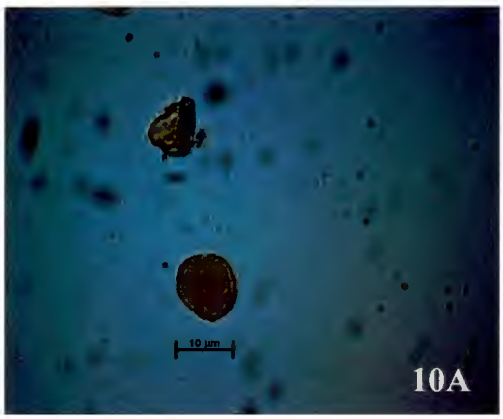
81. *Spore size*: **(0)** mostly between 9 and 20 um; **(1)** mostly between 21 and 35 um; **(2)** mostly between 36 and 59 um; **(3)** mostly between 60 and 95 um; **(4)** mostly between 88 and 100 um.

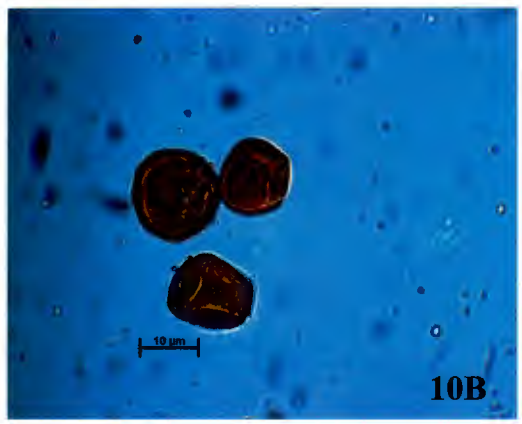
Spore size in *Andreaea* shows variation within the different species. However, the ranges of variation within species seem to be consistent for each species with very few outliers. The spore size ranges for the studied *Andreaea* species were mainly in four categories as indicated below: Most of the *Andreaea* species have spore size ranging between 21 and 35 μm in diameter (e.g. plates 10B and 10D). The smallest spores (between 9 and 20μm) occur in *A. blyttii* and *A. sinuosa* (i.e. plates 10A and 10C). *A. rothii* has spore size ranging between 36 and 59μm (plate 10E). The largest *Andreaea* spores (between 60 and 95μm) are found in *A. megistospora* (plate 10F). Though no capsules were found on some specimens, spores were easily found among the perichaetial and terminal leaves. Sizes were also confirmed with literature. *Andreaeobryum* has spores that are mostly between 88 and 100 um.

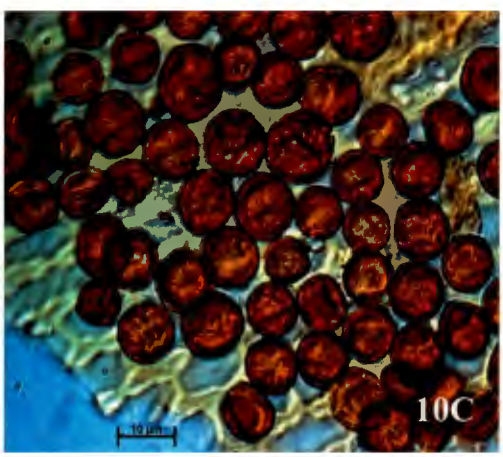
82. *Endosporic germination:* **(0)** Present; **(1)** Absent.

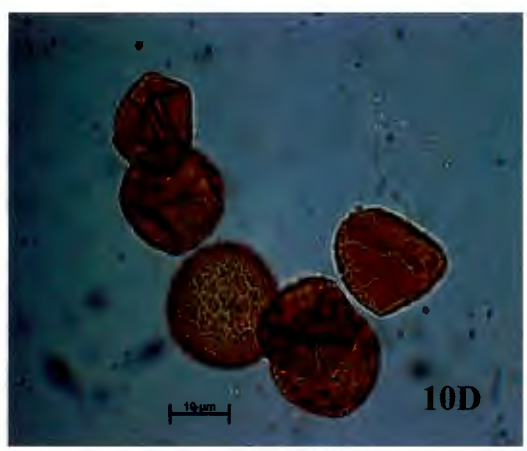
Spores in *Andreaea* mature very earlier and usually undergo cell division forming a globular cellular mass (Parihar, 1965) before the exospore (spore coat) is ruptured (Cavers, 1911, Schofield, 1985). They germinate to form either a thalloid or a branched ribbon like protonema (Scott, et al., 1976). The phenomenon of endosporic germination is absent in *Takakia, Andreaeobryum, Sphagnum* and *Tetraphis* (Murray 1988a).

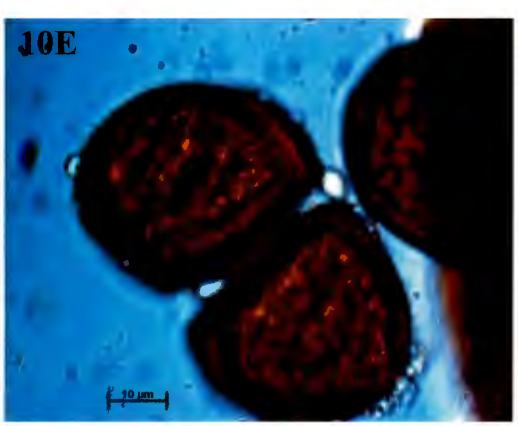
83. *Longitudinal capsule dehiscence*: **(0)** straight; **(1)** spiral.

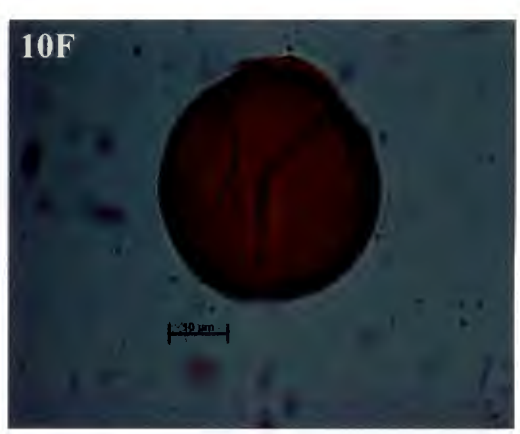

10 µm
10A


10 µm
10B


10 µm
10C


10 µm
10D


10E
10 µm


10F
30 µm

Plates 10: Specimens were rehydrated prior to mounting in Hoyers' solution on microscope slides. Photographs were taken from the Leitz Diaplan compound microscope (Leitz Wetzler, German).

10A: *A. blyttii*; spore (ca. 10 µm, usually between 9 — 20 µm).

10B: *A. rupestris* var. *rupestris*; spores (usually between 21 — 35 µm).

10C: *A. sinuosa*; spores (usually between 9 — 20 µm).

10D: *A. gainii*; spores (usually between 21 — 35 µm).

10E: *A. rothii*; spores (usually between 36 and 59 µm).

10F: *A. megistospora* spp. *megistospora*; spores (usually between 60 and 95 µm).

Capsule dehiscence is straight in *Andreaea* and *Andreaeobryum*, and spiral in *Takakia* (Smith and Davidson, 1993). In *Sphagnum* and *Tetraphis* it is transverse by means of an operculum.

84. *Valves that are split to give tooth like projections:* **(0)** Present; **(1)** Absent.
In *Andreaea* (most species), *Takakia* and *Andreaeobryum* the slits of capsule valves do not extend to the apex and hence the ends of the valves remain attached (Parihar, 1965). However *A. wilsonii* has valves that do not remain apically attached, but split to give tooth-like projections.

85. *Valve arrangement*: **(0)** regular; **(1)** irregular.
This is regular in *Andreaea* and irregular in *Takakia and Andreaeobryum* (Murray 1988a). The valve arrangement is also more or less spiral in *Takakia* (Smith and Davidson, 1993).

86. *Number of capsule valves*: **(0)** up to 4; **(1)** up to 6; **(2)** up to 8; **(3)** 2.
The number of capsule valves varies in the species of *Andreaea* studied, from 4 to 12. Most species however have 4 valves. Species with 4 valves include *A. rupestris*, *A. obovata*, *A. blyttii*, *A. alpestris*, *A. subulata*, *A. alpina*, *A. rothii*, *A. frigida*, *A. crassinervia* and *A. megistospora*, *A. gainii*, *A. australis*, *A. nitida*, *A. acutifolia* and *A. sinuosa*. In *A. wilsonii* valves vary from 6 to 12, whereas *A. nivalis* has up to 6 valves. *A. wilsonii*, the valves are initially attached together apically and later split up to give tooth-like projections.

Table 2: Distribution of characters states among the species of *Andreaea* and the outgroup taxa. Missing characters were coded as "?", inapplicable characters as "-" and multi-state characters as "{ / }".

A rupestris var rupestris	1011000000 1110001110 111111 01 1 1 1100-1---- - 01 0 110101 0111-11110 0000010-01 1111-10111 00100
A rupestris var papillosa1	1011000000 1110001110 111111 00 1 1 1111-1---- - 01 0 110101 0111-11110 0000010-01 1111-10111 00100
A rupestris var papillosa2	1011000000 1110001110 111111 00 1 1 1111-1---- - 01 0 110101 0111-11110 0000010-01 1111-10111 00100
A obovata var obovata1	1011000000 1110001111 111101 01 1 1 0111-1---- - 01 0 110011 0111-11110 0000010-01 1111-10111 00100
A obovata var hartmannii	1011000000 1110001111 111101 01 1 1 0111-1---- - 01 0 110011 0111-11110 0000010-01 1111-10111 00100
A obovata var obovata2	1011000000 1110001111 111101 01 1 1 0111-1---- - 01 0 110011 0111-11110 0000010-01 1111-10111 00100
A alpina	1011000000 1110001111 111101{01}101 1111-1---- -{01}10110101 0111-11110 0000010-01 1111-10111 00100
A gainii	1011010000 1110001111 111101{01}101 1111-1---- -{01}10110011 0111-11110 0000010-01 1111-10111 00100
A alpestris	1011000000 1110001110 111111 01 1 1 1101-1---- -{01}10110011 0111-11110 0000010-01 1111-10111 00100
A wilsonii	1011000000 1110001110 111101 01 1 1 0111-1---- -{01}10110101 0111-11110 0000011-10 1110-10110 00002
Andreaeobryum macrosporum	0000010001 1100000101 111111 01 1 1 0111-00112 1{01}11011--1 1001111--0 111110-111 1011-01104 00112
A sinuosa	1011010000 1110001101 111111 00 1 1 1111-1---- - 01 0 110101 0111-01110 0000010-01 1111-10110 00100
A subulata	1011010001 1101-00011 111110 00 1 1 1111-01111 1{01}11011--1 1001111-10 0000010-01 1111-10111 00100
A acutifolia	1011000000 1110001011 111110 00 1 1 1111-1---- - 01 0 110101 0111111110 0000010-01 1111-10111 00100
A blyttii	1011010001 1100000011 111110 00 1 1 1111-01112 0{01}11110101 1001011010 0000010-01 1111-10110 00100
A nivalis	1011010001 1101-00011 111110 10 1 1 1111-00111 1{01}11011--1 1101111-00 1110010-01 1111-10111 00101
A rothii ssp rothii	1011000001 1100000011 111110 00 1 1 1111-00110 1{01}11011--1 1001111-10 0000010-01 1111-10112 00100
A rothii ssp falcata	1011000001 1100000011 111110 00 1 1 1111-00110 1{01}11011--1 1001111-10 0000010-01 1111-10112 00100
A crassinervia	1011000001 1100000011 111110 00 1 1 1111-00112 1{01}01011--1 1001111-10 0000010-01 1111-10111 00100
A megistospora ssp megistospora	1011000001 1100000011 111110 00 1 1 1111-00111 1{01}11011--1 1001111-10 0000010-01 1111-10113 00100
A megistospora ssp epapillosa	1011000001 1100000011 111110 00 1 1 1111-00111 1{01}11011--1 1001111-10 0000010-01 1111-10113 00100
A schofieldiana	1011010001 1100000011 111110 00 1 1 1111-00111 1{01}11011--1 1001011-10 0000010-01 1111-10111 00100
A frigida	1011000001 1101-01011 111110 00 1 1 1111-00110 1{01}11011--1 1001011-10 0000010-01 1111-10111 00100
A bistratosa	1011010000 1110001101 111111 00 1 1 1111-1---- - 11 1 110101 0111-11110 0000010-01 1111-10110 00100
A nitida	1011010000 1110001111 110111 01 1 0 1111-00010 1 01 1 011--1 1001111-00 0010010-01 1111-10111 00100
A australis	1011010001 1111-00111 011111 00 1 1 1111-00101 1 01 1 111--0 1001111-10 1010010-01 1111-10111 00100
Takakia lepidozioides	0000110010 111--1---- ------ -- - - -0-------- - -0 1 ------ -----11--- ---110-101 1111-?0111 11113
Tetraphis geniculata	-011011-01 1011-01111 101111 00 1 1 1111-00100 1 01 1 111--1 1011111--0 011100-010 110-011100 1----
Tetraphis pellucida	-011011-01 1011-01111 101111 00 1 1 1111-00100 1 01 1 111--1 1011111--0 011100-110 110-011100 1----
Sphagnum girgensohnii	-111110101 0110101111 111011 01 1 1 111001---- - 01 1 101--1 1110-10-11 0001111-11 010-10101? 1----
Sphagnum pylaesii	-111110101 0110101111 111011 01 1 1 111021---- - 01 1 101--1 1110-10-11 0001112-11 010-10101? 1----

CLADISTIC ANALYSIS OF MORPHOLOGICAL DATA

The morphological data were assembled in Nexus data editor software (Page, 2000). The data (Table 3) were analysed using parsimony analysis as implemented in PAUP version 4.0b4a for Macintosh (Swofford, 1998).

To test for the presence of significant phylogenetic signal in the datasets, the g_1 statistic (Heusenbeck, 1991), based on the distribution of 100,000 random tree lengths (constructed using the RANDOM TREES procedure) was utilised. The skewness of the tree length distribution is a good indicator of the quality of the character data, with a left skewed tree length distribution indicating the presence of significant phylogenetic signal (Heusenbeck, 1991).

Owing to the large size of the data set it was not possible to perform an exact search method (Kitching, et al., 1998). The most parsimonious trees were therefore generated using the heuristic search option of PAUP*, using the tree bisection–reconnection (TBR) algorithm (with random addition sequence; 1000 replicates; branches having maximum length zero collapsed to yield polytomies; MULPARS and STEEPEST DESCENT options in effect, and with an initial 'maximum trees saved' setting of 50,000). The TBR branch-swapping algorithm divides the initial tree into two subsets by bisecting a branch between nodes and then pruning both resulting free branches leaving two disjoint subtrees. The two subtrees are then reconnected by creating a linking branch between them. All possible bisections and reconnections are then evaluated. The branch lengths of optimal trees were calculated using ACCTRAN optimisation (places character state changes on the tree as close to the root as possible) (Farris, 1970). All transformations were weighted equally (Fitch parsimony). Fitch parsimony (Fitch, 1971, Hartigan, 1973) allows free transformation of a state into any other state with the cost of only one additional step in tree length, thus permitting free reversibility of transformations.

Strict Consensus Tree

In cladistic analysis, multiple parsimonious trees are often obtained due to different character state optimisations of homoplastic characters or from the choice of which

characters should be homoplastic (Anderberg and Tehler, 1990). To synthesise one phylogeny from these many hypotheses, the strict consensus was calculated as implemented in PAUP*. It illustrates components common to all the equally parsimonious cladograms of the analysis (Anderberg and Tehler, 1990).

Successive weighting

In an attempt to obtain a single or few most parsimonious trees, successive weighting (Farris, 1969, Carpenter, 1988) was applied using the maximum value of rescaled consistency index. Successive weighting was also utilised as a check on whether homoplasy (i.e. low consistency index) could obscure character information (Farris 1969). This assigns weights to characters such that the characters that have a low consistency index (incongruent with the other characters) receive low weight and those with a high consistency index (i.e. congruent with the other characters) are given higher weighting. However use of the consistency index to assign weights does not result in completely homoplasious characters being assigned a value of '0', and hence the use of the rescaled consistency index (rc = ri x ci) as the value of weight (Farris 1989) in the current analysis. This approach achieves a value of zero for completely homoplasious characters.

Character state optimisation

The supporting character state changes were optimised, using ACCTRAN optimisation, onto one of the most parsimonious trees chosen arbitrarily from among the 18 most parsimonious trees obtained from the analysis of successively weighted data.

Bootstrap analysis

To assess the support for the recovered tree topology, bootstrap analysis (Felsenstein, 1985) was performed in PAUP*. The algorithm proceeds by random sampling with replacement of the characters in a data set to build up a bootstrap data set of the same size as the original data set, which is analysed to give a tree or a number of trees. This procedure is repeated at least 100 times and the percentage of occurrence of a

particular node among the trees of the sample data sets is considered as an index of support. However this does not give true confidence limits in a statistical sense.

In this analysis bootstrapping was conducted using the Fast Heuristic search (TBR algorithm with MULPARS effected), with 10,000 random replicates. All groups (nodes) with a frequency of greater than 50% were retained in the bootstrap consensus tree.

Jackknife analysis

To further assess the stability of each clade obtained from analysis of the dataset, a jackknife analysis (Davis, 1993) was performed in PAUP* version 4.0b4a (Swofford, 1998), with 10,000 replicates and 33.7% character deletion, and via the "Fast" stepwise addition option. Jackknife proceeds by random deletion of characters without replacement. After each character sampling, the general heuristic search is run followed by the consensus tree computation. The percentage of times that a clade appears in the replicate analyses is the support for that particular clade. In the current analysis, clades with frequency greater than 50% were retained in the consensus tree.

RESULTS OF MORPHOLOGICAL ANALYSES

Equally weighted data

Cladistic analysis of the unweighted morphological dataset yielded 411 equally most parsimonious trees (Length = 157 steps, consistency index (C.I) = 0.592, C.I excluding uninformative characters = 0.536, retention index (R.I.) = 0.782 and rescaled consistency index (RC) = 0.463. A g_1 statistic of 0.417 suggested presence of a phylogenetic signal in the morphological dataset; i.e. a significant ($P<0.05$) level of structure existed within the dataset (Hillis and Heulsenbbeck, 1992).

The fairly resolved strict consensus tree of the 411 most parsimonious trees recovered from the analysis of unweighted morphological data is presented in Figure 1. In all the 411 most parsimonious trees *Andreaea* was resolved without statistical support as a monophyletic sister to *Takakia*, with no bootstrap or jackknife support. Two major clades were distinct in the ingroup; the first fairly well supported clade (node 13;

bootstrap = 71% and Jackknife = 77%), consisting of all the ecostate species and the second, poorly supported clade (node 7; bootstrap <50% and Jackknife = 56%), consisting of costate species. All other nodes for the ingroup were poorly supported. Within the clade containing the ecostate species, the deepest node (13) was polytomous consisting of the *A. sinuosa / A. bistratosa* clade, *A. acutifolia*, and a well-supported clade containing the rest of the ecostate species of *Andreaea* that were studied (node12). *A. wilsonii* (traditionally within subgenus *Acroschisma*) was resolved within section *Andreaea* of the subgenus *Andreaea*, whereas *A. nivalis* (traditionally within subgenus *Chasmocalyx*) was placed in an unresolved clade with species of section *Nerviae* (subgenus *Andreaea*) and *A. australis* (subgenus *Chasmocalyx*). Positions within this costate species clade (node 7), excluding *A. nitida*, were poorly resolved.

Successively weighted data

The successive weighting was performed twice before the tree stabilised. A total of 9 trees (length = 721, consistency index (C.I) = 0.769, C.I excluding uninformative characters = 0.692, retention index (R.I.) = 0.874 and rescaled consistency index (RC) = 0.672) were recovered from the analysis of the successively weighted morphological data. The strict consensus tree of the 9 most parsimonious trees recovered from this analysis is presented in Figure 2. In all the 9 most parsimonious trees the monophyly of the *Andreaea* clade is well supported (bootstrap = 88% and jackknife = 94%). Two major clades were distinct for the ingroup taxa (the *Andreaea* clade), the first, well supported clade (node 20; bootstrap = 84% and jackknife = 89%) including all the ecostate species and a second unsupported clade (node 17) comprising the costate species excluding *A. nitida*, which is resolved on a polytomy with these two major clades. The clade of costate species excluding both *A. australis* and *A. nitida* (node 15) is however well supported (bootstrap = 69% and jackknife = 80%). Within the clade consisting of ecostate species, *A. bistratosa*, *A. sinuosa*, and *A. acutifolia* form an unsupported grade basal to the rest of the species (node 15 of Figure 2). This clade (node 17), excluding *A. bistratosa*, was weakly supported (bootstrap = 59% and jackknife = 65%). All other positions are similarly of poor support.

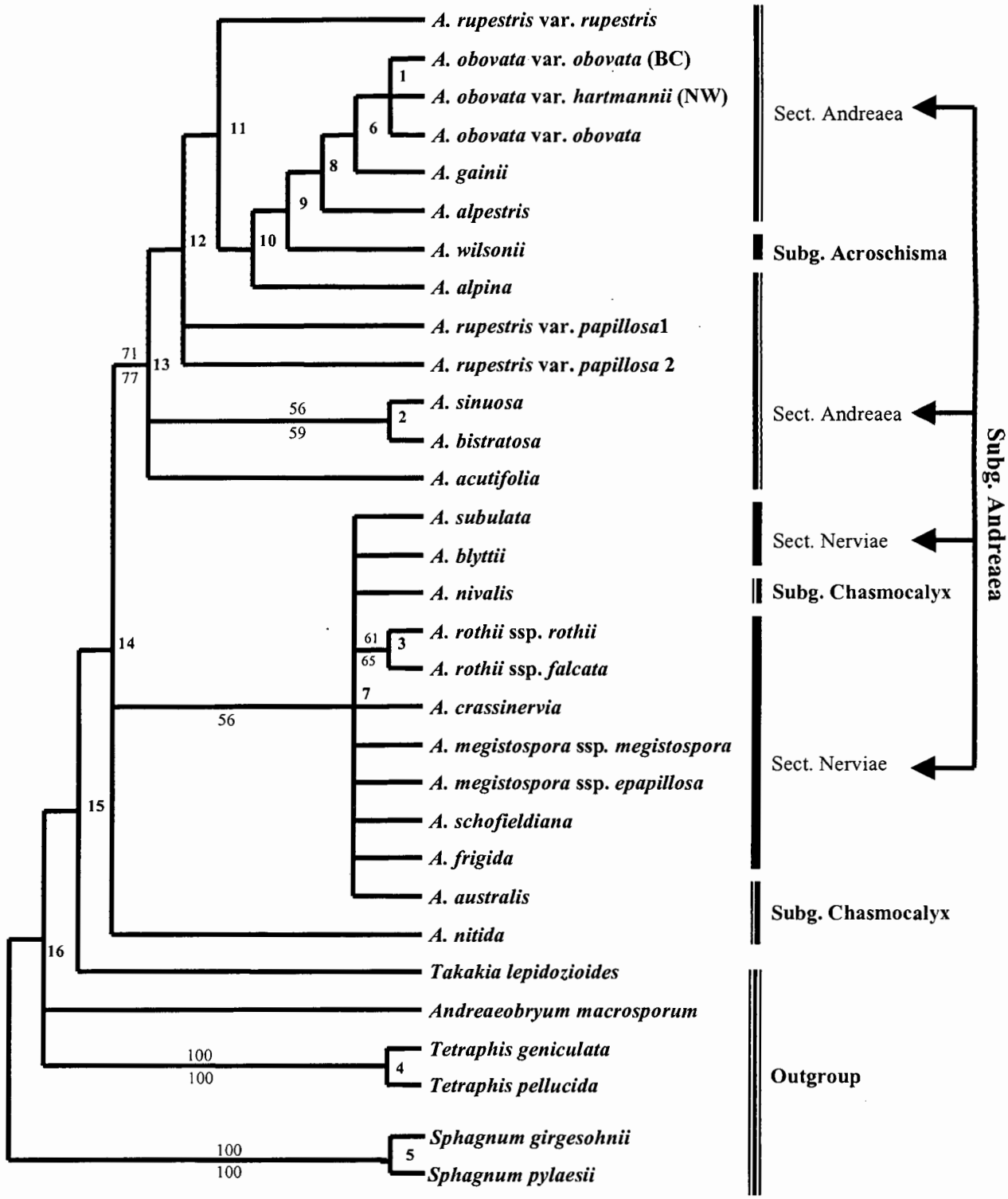


Figure 1. Strict consensus of 411 most parsimonious trees recovered from an analysis of the unweighted morphological dataset. Numbers above are bootstrap values whereas those below are jackknife values. Node numbers are bold values inside the nodes. The classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

In the clade comprising most of the costate species, except *A. nitida* (i.e. node 17), *A. australis* is placed coordinate to a well-supported (bootstrap = 69% and Jackknife = 80%) clade (node 15). The rest of the relationships in the clade are either unresolved or poorly supported.

Optimisation of character state transformations

Figure 3 shows one of the most parsimonious trees recovered from the analysis of weighted morphological data, with all unambiguous character changes optimised onto it. Of the 20 synapomorphies observed for the ingroup taxa, only 11 were uncontradicted (Table 3), 4 of which are for the whole *Andreaea* clade, 2 for the clade consisting of all ecostate species and 1 each, for 4 other clades as indicated in table 3. Forty-four homoplasious character changes were observed, whereas 20 autapomorphies were observed, 6 for *A. wilsonii* and 3 each for *A. nivalis*, *A. australis* and *A. nitida*. The other autapomorphies were observed in *A. blyttii* (2), *A. sinuosa* (1) and *A. bistratosa* (1).

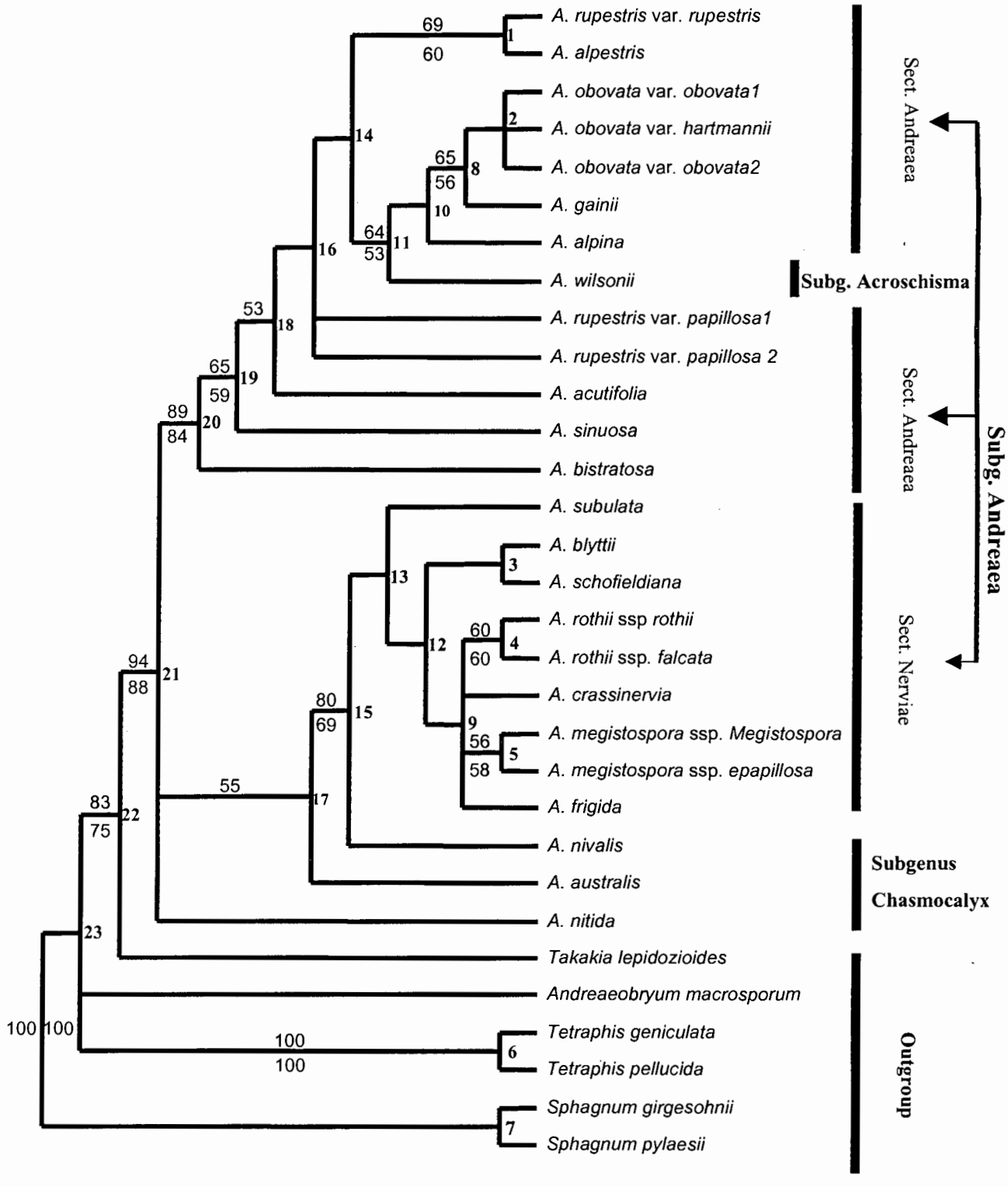


Figure 2: strict consensus tree of the 9 most parsimonious trees recovered from an analysis of the weighted morphological data set. The numbers above are bootstrap values whereas the ones below are jackknife values. Nodes numbers are bold values inside the nodes. The classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

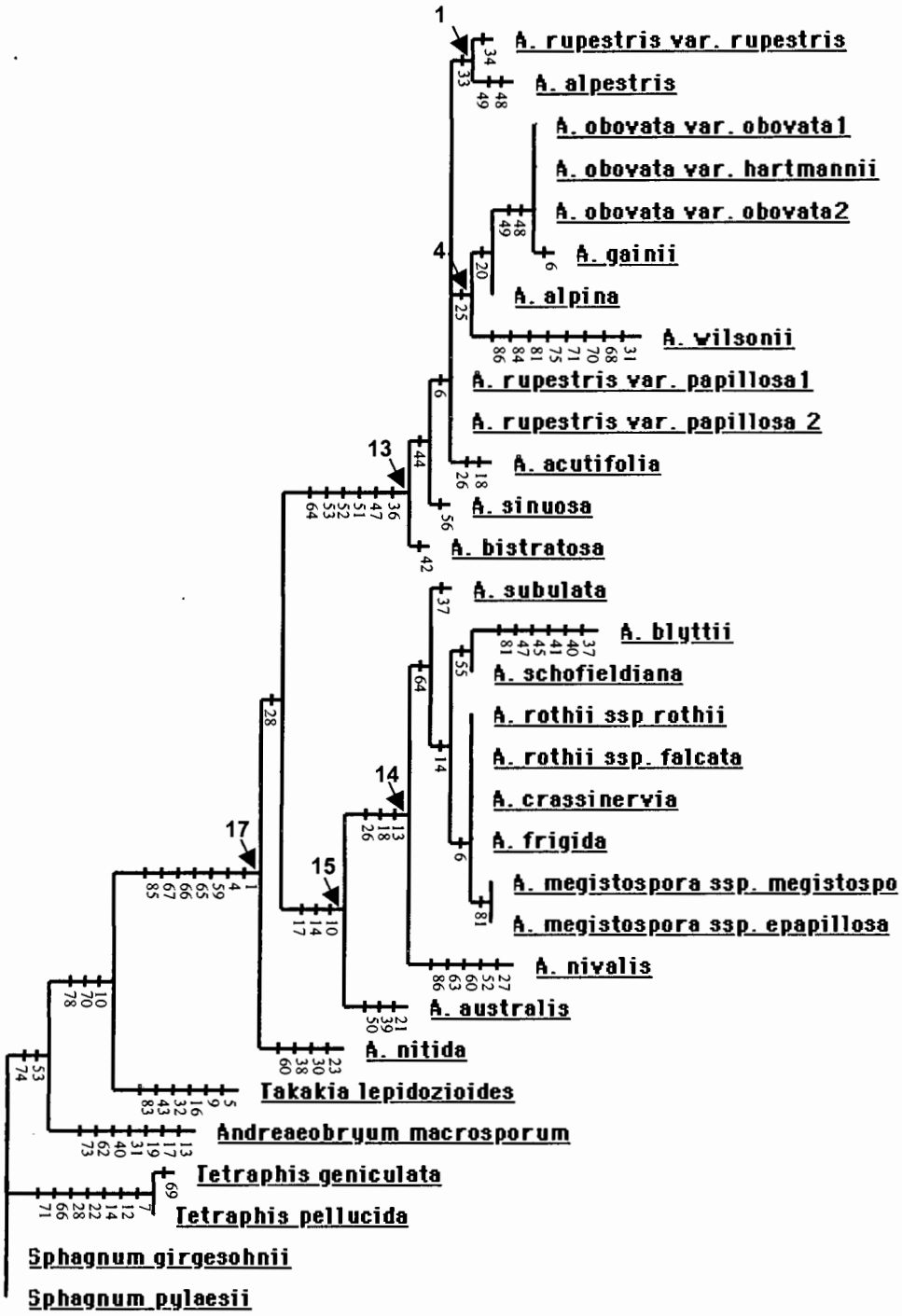


Figure 3: One of the most parsimonious trees (chosen arbitrarily) obtained from an analysis of successively weighted morphological data with character state changes optimized onto it via ACCTRAN. Nodes referred to in the text, are indicated by numbers with arrows.

Table 3: Eleven uncontradicted synapomorphies for clades in the genus *Andreaea*. Node numbers refer to figure 3.

CLADE	NODE NUMBER	SYNAPOMORPHIES
The whole *Andreaea* clade.	17	1 4 65 85
Ecostate species clade.	13	36 51
***A. rupestris, A. alpestris* clade.**	1	33
***A. obovata, A. gainii, A. alpina, A wilsonii* clade.**	4	25
Costate species, excluding *A. nitida.*	15	17
Costate species, excluding *A. nitida* and *A. australis.*	14	13 26

DISCUSSION OF THE MORPHOLOGICAL ANALYSES

The analyses of unweighted morphological data showed very poor support (i.e. <50% bootstrap and jackknife values) for the genus *Andreaea* as a whole. None of the traditional infra-generic groupings was resolved as monophyletic. What emerged from the analyses was that *A. wilsonii*, sometimes afforded generic status in earlier treatments (e.g. Reimers, 1954), was resolved in a clade of ecostate species of subgenus *Andreaea* whereas *A. nivalis* (traditionally subgenus *Chasmocalyx*) was resolved within the section *Nerviae* of the same subgenus. Relationships of many species and assemblages were not resolved in the recovered phylogeny, and positions for most of the resolved groups were not supported by either the bootstrap or jackknife indices. This however is probably due to the high level of homoplasy in the dataset as indicated by the low CI-excluding uninformative characters (0.536). Only 34% of the characters were found to be non-homoplasious by optimisation of character state transformation onto one of the most parsimonious trees (figure 3). The high level of homoplasy may be due to different lineages, with highly labile characters, responding to similar selective pressures with consequent convergent

evolution. There are very few uniquely shared and uncontradicted synapomorphies in the dataset, for most of the clades. There is thus lack of morphological synapomorphy to unite groups within the genus. For example only one character was optimised on the node supporting the costate species of *Andreaea* excluding *A. australis*. Similarly, only 2 characters were optimised on the node supporting the clade consisting of the nerveless species of *Andreaea*. It is worth noting that *A. wilsonii*, *A. nivalis* and *A. blyttii* exhibit considerable morphological divergence compared to the other taxa.

Further corroboration of the influence of homoplasy in the dataset was shown by changes that occurred in the recovered topology upon character re-weighting. The monophyly of the genus *Andreaea* became strongly supported and the overall resolution improved. This further indicated that the homoplasious character changes had obscured phylogenetic relationships in the unweighted dataset. However there is still little agreement in these results with earlier suggestions: for example Murray, (1988) suggested that both *A. blyttii* and *A. nivalis* may posses pleisiomorphic features within *Andreaea*. Only section *Nerviae* of the subgenus *Andreaea* was resolved (with no support) as monophyletic in the analysis of weighted morphological data. The monophyly of the other subgenera and sections were not supported by this analysis. Members of the subgenus *Chasmocalyx* were resolved as non monophyletic (polyphyletic), with *A. nivalis* and *A. australis* forming a grade that was coordinate to the clade of section *Nerviae* and *A. nitida* occupying a position on a polytomy with the two major clades resolved of the rest of the species of *Andreaea*. *A. wilsonii* was resolved in the section *Andreaea* of Subgenus *Andreaea*. The results of the morphological analysis therefore only approximated the earlier views (e.g. Braithwaite, 1887, Schultze-motel, Murray, 1988b) of infra-generic assemblages.

However, the phylogeny recovered in the current analysis was not a very reliable one as indicated by the low bootstrap and jackknife support values. At this stage evidence for such generic and sectional groups is weak, and it is therefore not possible to make final conclusions about the monophyly or non-monophyly of infrageneric groups or about species relationships. Whether for example costate or ecostate species are a natural group or not, or whether taxa like *A. wilsonii*, *A. nitida* and *A. nivalis* should belong to various subgeneric assemblages distinct from the subgenus *Andreaea* may

therefore not be clearly evident before further analyses (i.e. molecular analysis and a combined morphological and molecular analysis) in the proceeding chapters. The inclusion of data from molecules in the latter part of the thesis may give a better understanding of the overall phylogeny as well as of any natural assemblages within *Andreaea*. Due to the non-reliability of the morphological results alone, the discussion of morphological character evolution is also reserved for the latter part of the thesis.

CONCLUSIONS

Though the analysis of successively weighted morphological data strongly supports the monophyly of the genus *Andreaea*, it has only provided weak suggestions of phylogenetic relationships and of sub-generic and sectional groupings in the genus *Andreaea*. The high level of homoplasy in the dataset has definitely obscured information from the other characters regarding the various infra-generic assemblages, resulting in weakly or unsupported clades and grades. The non-recovery of strongly supported relationships and clades has further been aggravated by lack of uncontradicted synapomorphies for various presumed infra-generic groups. Other evidence, independent of the morphological characters utilised here, such as ultra-structural, ontogenetic, or molecular data is therefore required to shed more light on, and strengthen the understanding of phylogenetic relationships within the genus. In the proceeding chapters however, only molecular sequence data have been employed to add more phylogenetic information and further inquire into the evolutionary relationships of the genus.

CHAPTER 3

CLADISTIC ANALYSIS OF THE GENUS *ANDREAEA* BASED ON *TRNL–F* AND *RPS4* SEQUENCES

INTRODUCTION

In certain cases, morphological evidence alone may not be adequate for providing information on the relationships at certain taxonomic levels. Further, morphological characters are often polymorphic for some taxa. There are also not so many other morphological alternatives to utilise for inferring phylogenies, when a given set of morphological data seems inadequate for resolving relationships. Thus the morphological data alone sometimes prove inadequate as the sole information for answering phylogenetic questions in taxa such as the genus *Andreaea* where the level of polymorphism in morphological characters is apparently very high and where uncontradicted synapomorphies for various groups are rare.

The development of molecular techniques such as DNA sequencing has provided another source of data from which phylogenies can be established independently. Molecular techniques therefore serve as an independent means of evaluating the evolution of morphological and other characters (Donoghue and Sanderson, 1992). It is therefore hoped that molecular analyses employed in the current study would serve to provide insights into the evolution of morphological characters in the genus *Andreaea*. A molecular approach to phylogeny estimation is also seen here as valuable in testing the validity of the putative sub-generic and sectional groupings in the genus (i.e. by testing their monophyly). This would thus also confirm or invalidate the use of certain characters utilised in the past for postulating these groupings.

For the current study, two chloroplast gene regions, the ribosomal protein S4 (*rps4*) and the chloroplast DNA *trnL–F* region were utilised. This two genes approach was aimed at capturing phylogenetic information about the species of *Andreaea* from molecular datasets with varying rates of sequence evolution.

The *trnL–F* intergenic spacer has been shown in a number of studies (e.g. Kohjyouma et al., 2000; Bakker et al., 2000, Terry et al., 2000, Brouat et al., 2001) to be informative for resolving infra-generic relationships. The *rps4* gene, a short region (~ 600bp), which encodes for the S4 subunit of the plastid ribosome, has also been shown to have phylogenetic potential (Soltis and Soltis, 1998). It has been used mostly for answering supra-generic phylogenetic questions (e.g. Nadot, 1994; Cox and Hedderson, 1999; Buck et al., 2000). Not so many studies have utilised the gene for lower level (i.e. infra-generic) studies. However, this gene is generally easy to amplify and sequence, and a few initial sequences obtained for *Andreaea* during this study showed that it would be useful in providing information especially for the deeper nodes in the *Andreaea* phylogeny as well as for linking the genus with outgroup taxa. Together with the *trnL–F* sequences, it can offer useful information for inferring phylogenetic relationships of the genus *Andreaea*.

In this chapter both separate analysis of the *trnL–F* and combined analysis of the *trnL–F* and *rps4* datasets were performed. Whether to perform separate analysis or integrate multiple datasets in cladistic analysis is a highly contentious issue. Various authors have advocated different approaches to the problem of handling diverse datasets. Separate analysis of datasets has been suggested by a number of authors e.g. Marshall (1992), de Queiroz (1993) whereas others, e.g. Miyamoto (1985), Kluge (1989), Barrett et al. (1991), Donoghue and Sanderson (1992), have advocated combining all available data prior to phylogenetic analysis (character congruence or total evidence approach). Bull et al., (1993) and de Queiroz, (1993) however, have suggested separate analysis of subsets of total available data and allowing combination of data only if the resulting trees from separate analysis are congruent. Another similar approach advanced by some authors (e.g. Mickevich, 1978; Hillis, 1987; Swofford, 1991; de Queiroz, 1993) is taxonomic congruence, which entails the separate analysis of diverse datasets followed by combination of the independent estimates using consensus methods.

Some of the possible advantages of combining datasets in phylogenetic analysis include the following; (1) Poorly supported but similarly resolved clades may become more robust with the combined data, (2) speed of computation usually increases appreciably with combined datasets, (3) weak signals, masked by noise in separate

datasets may be additive and rise above the noise (e.g. Barrett et al., 1991), (4) combining various datasets may substantially improve the resolution of the tree (Hillis, 1987), if for example different character classes are useful in resolving specific areas of the tree but are uninformative for others. One character set may resolves nodes closer to the tips of the tree and another may be more useful for basal resolution. The fundamental aim of the combined analysis therefore, is to recover a phylogeny with maximum resolution and support.

In the current study various possibilities have been considered. Firstly it is inappropriate to ignore the possibility of utterly uncombinable datasets (e.g. datasets with significantly heterogeneous evolutionary models, or too many missing characters can not be appropriately subjected to a combined ML analysis). Second it is worth bearing in mind the potential value of combining datasets, in certain cases (e.g. Hillis, 1987). Therefore, in the current analysis I used both combined analyses of all the available datasets as well as separate analyses of the individual datasets. Phylogeny inferences were then derived from the combined as well as the separate analyses. For the maximum likelihood analysis, data were not combined since the *rps4* had too much missing information for a number of taxa and the species were therefore not sufficiently corresponding in the *trnL* and *rps4* datasets.

AIMS OF THE CHAPTER

This chapter presents the methods utilised for obtaining the cpDNA sequences *rps4* and *trnL–F* for the genus *Andreaea*. Molecular variation based on these sequences is documented and used to infer phylogenetic relationships of the genus *Andreaea* using maximum parsimony or maximum likelihood analyses of the separate and combined molecular datasets. The monophyly of the various infra-generic groups was also thus tested based on the molecular evidence. For reasons explained later, the *rps4* sequences were only utilised for combined analyses.

MATERIALS AND METHODS

Taxon Sampling

In the molecular study sequences from 39 specimens were utilised for phylogenetic analysis. These consisted of sequences for 16 ingroup taxa (*rps4* and *trnL–F*) obtained from dried herbarium specimens and 13 outgroup taxa (already available from GenBank) representing some of the suborders in all the 5 orders of the Bryales (Vitt, 1984). These groups of mosses share enough molecular synapomorphies for the sequences utilised in this study (Newton et al., 2000), and can therefore be easily related to each other in cladistic analyses. The exemplar outgroup taxa included; the only species of *Andreaeobryum*, two species of *Sphagnum* and one species each of *Encalypta* Hedw., *Funaria* Hedw., *Hedwigia* P. Beauv., *Disphyscium* Mohr., *Dawsonia* R. Br., *Polytrichum* Hedw., *Buxbaumia* Hedw., *Dicranum* Hedw. and *Tetraphis* Hedw. Addition of these outgroup taxa was aimed at facilitating the determination of the relative position of *Andreaea* among the mosses. Information for the specimens of these species utilised in the study, together with the GenBank accession numbers (for accessioned sequences), is given in Tables 4 and 5. Even though many members of the speciose genus *Andreaea* were not included in the study, all its current infra-generic groups were represented. The species of *Andreaea* included in the molecular analyses were selected also on the basis of availability of material suitable for DNA analysis of sequences for *trnL–F* and *rps4* gene regions. Other limitations are pointed out later.

DNA extraction

Total genomic DNA was extracted from herbarium specimens using either the Qiagen DNeasy Plant Mini kit (PE Biosystems), following manufacturer's instructions, or the small-scale CTAB extraction method (Gawel and Jarret, 1991).

To check whether DNA was successfully extracted for each sample, agarose gel electrophoresis was performed in 1% agarose gels stained with ethidium bromide. The gels were photographed using a UV Transilluminator (UVP, inc), with a video camera connected to a CyberTeck CSI computer and a Video copy processor.

TABLE 4: List of the taxa included in the present study. The classification was derived from Dixon 1932; Schultze-motel 1970; Vitt, 1984 and Murray 1988b. All the sequences have not yet been accessioned.

TAXON	VOUCHER SPECIMEN	AVAILABLE SEQUENCES	
		*rps*4	*trnL–F*
Subgenus *Andreaea*			
Section *Andreaea*			
A. rupestris var. *rupestris*	Hedderson T. A 10439	✓	✓
A. rupestris var. *papillosa* (Lindb.) Podp.	Hedderson T. A 6640	—	✓
A. rupestris var. *papillosa* (Lindb.) Podp.	Einar Heegaard, 262	✓	✓
A. acutifolia Hook et Wils.	Schofield 108339 With Talbot	—	✓
A. alpestris (Thed.) Schimp.	Hedderson T. A 4755	—	✓
A. obovata var. *hartmannii*	Einar Heegaard, 247	✓	✓
A. obovata var. *obovata* Thed.	Einar Heegaard, 263	✓	✓
Section *Nerviae*			
A. blyttii Schimp.	Hedderson T.A 4976	—	✓
A. crassinervia	Einar Heegaard, 239	✓	—
A. frigida Huebener	Einar Heegaard, 252	✓	✓
A. megistospora B. Murr. ssp. *megistospora*	Schofield 83690	—	✓
A. rothii ssp. *falcata*	Einar Heegaard, 38	—	✓
A. rothii ssp. *rothii*	Einar Heegaard, 54	—	✓
A. subulata Harv.	Hedderson T. A 13229	✓	✓
Subgenus *Chasmocalyx*			
A. nivalis Hook	Einar Heegaard, 253	✓	✓
Subgenus *Acroschisma*			
A. wilsonii Hook.	C. J. Cox and B. Goffinet No. 668/00	✓	✓

Table 5: List of the outgroup taxa included in the present study. the classification was derived from Vitt (1984) and Murray (1988b).

Taxon	GenBank Accession numbers	
	*rps*4	*trnL–F*
Sphagnales		
Sphagnum palustre L.	AF231892	AF231902
Sphagnum cuspidatum Brid.	AF231893	AF231903
Takakiales		
Takakia Lepidozioides Hatt. et Inoue	AF231894	AF231904
Tetraphidales		
Tetraphis pellucida Hedw.	AF231896	AF231908
Polytricales		
Polytricaceae		
Dawsonia papuana Schlieph. &Geheeb	AF208419	AF246704
Polytricaceae		
Polytrichum commune Hedw.	AF208428	AF231907
Bryales		
Buxbaumiaceae		
Buxbaumia aphylla Hedw.	AF231897	AF231909
Dicranaceae		
Dicranum scoparium Hedw.	AF234158	AF234159
Encalyptaceae		
Encalypta rhaptocarpa Schwaegr.	AF23777	AF023717
Funariaceae		
Funaria hygrometrica Hedw.	AF023776	AF023716
Hedwigiaceae		
Hedwigia ciliata (Hedw.) P. Beauv.	AF005517	—
Buxbaumiaceae		
Disphyscium foliosum (Hedw.) Mohr.	AF223034	AF229891

Polymerase Chain Reaction (PCR)

Polymerase Chain Reactions (Saiki et al., 1985) were performed to obtain sequencing templates from the genomic DNA. The primers used for amplifying the *rps4* region were forward primer *rps5* (5^{1} ATGTCCCGTTATCGAGGACCT 3^{1}) and reverse primer *trnAS* (5^{1} TACCGAGGGTTCGAATC 3^{1}) whereas the ones used for amplifying the *trnL–F* region were forward primer *trnL* (5^{1} CCA AAT CGG TAG ACG CTA GG 3^{1} and reverse primer *trnF* (5^{1} TTT GAA CTG GTG ACA CGA G 3^{1}).

Amplification was carried out in a Sprint thermal cycler (Hybaid Limited) programmed for denaturation at 97°C for 2 min, followed by 30 cycles of the following thermal sequence; 1 min at 97°C, 1 min at 52°C and 2 min at 72°C, with a 7 minute extension step after the 30 cycles.

To check whether the amplification was successful and whether there were any contaminants (i.e. if the negative controls had any undesirable bands), the PCR products were subjected to gel electrophoresis, as in the case of the extracted DNA.

The PCR products were purified using the QIAquick PCR purification Kit (QIAGEN) following the manufacturers' instructions.

DNA sequencing

The DNA templates were sequenced in 10µl, fluorescent dye-labelled reactions, using the same primers utilised for amplification and in conjunction with the ABI Prism™ Dye Terminator cycle sequencing ready reaction kit (P.E. Applied Biosystems). The cycle sequencing reaction tubes were placed in a PCR Sprint thermal cycler (Hybaid Limited) and subjected to the following thermal sequence for 25 cycles: 96°C for 30 seconds (denaturation), 50°C for 15 seconds (annealing) and 60°C for 4 minutes (extension). Sequence products were resolved on an ABI 373 automated sequencer according to the manufacturer's instructions (Applied Biosystems).

Assemblage and Alignment of sequences

The resulting sequences were assembled in SeqMan software of the Lasergene System Software Package (DNASTAR Inc, 1994). The raw data on each trace file were checked for base calling errors by the computer associated with the sequencer. Initially, preliminary alignments were performed for four or three sequences using the Clustal algorithm as implemented in the MegAlign alignment package (Lasergene System Software, DNA Star Inc.). This alignment was then adjusted by eye and all the remaining sequences added to this initial alignment. The aligned sequences were saved in Nexus format to allow compatibility with MacClade (Maddison & Maddison, 1992) and PAUP* version 4.0b4a (Swofford, 1998). Since the initial sequences were easily alignable to the already available sequences, it was evident that they were the targeted sequences and not contaminant sequences (e.g. bacterial or fungal). They were therefore not submitted to GenBank for confirmatory BLAST searching.

The amplified PCR products for the *trnL–F* region were approximately 540 bp. long. Only 470 positions of the final alignment were used, since some regions at the ends were missing for many taxa. Three phylogenetically informative indels varying in length from 3 to 29 bases long were present in the final *trnL–F* alignment. An AAATAAG repeat was shared by *A. obovata* var. *obovata*, *A. obovata* var. *hartmannii*, *A. rupestris* ssp. *rupestris*, *A. rupestris* ssp. *papillosa*, *A. alpestris* and *A. acutifolia* in position 94 relative to *A. blyttii*. *A. blyttii* and *A. nivalis* shared a 29 base pair insertion at position 187 relative to *A. blyttii*, whereas *A. rothii* ssp. *falcata*, *A. rothii* ssp. *rothii* and *A. megistospora* ssp. *megistospora* shared a TTT repeat in position 149 relative to *A. blyttii*. The CG content for all the sequences (all positions included) was 27% (G = 14%, C = 13%). Therefore, the *trnL–F* sequences for species of *Andreaea* are AT rich.

The amplified PCR products for the *rps4* region were approximately 900 bp long. This unusually long *rps4* region is apparently characteristic of the *Andreaea* species. In most other bryophyte species it is only about 600 bp long (Soltis and Soltis, 1998). The long *rps4* region in *Andreaea* is only due to the extension of the non-coding region. The coding region is similar in length to other bryophyte species.

Most of the reverse direction cycle sequencing reactions produced sequences that were unsatisfactory and due to limitations of time, it was not possible to repeat the sequencing. Most of these reverse direction reactions were therefore not utilised for sequence assembly. From the few complete sequences available it is evident that this long non-coding region will provide many potentially informative characters for *Andreaea*. Because the non-coding region sequences were excluded, only 660 bp were aligned. The final region used comprised 514 positions, 28 phylogenetically informative substitutions and no phylogenetically informative indels. Again regions at the ends were not included for many taxa. The CG content for all the sequences (all positions included) was 28%. The *rps4* sequences of *Andreaea* are therefore AT rich.

The aligned sequences for both the *trnL–F* and the *rps*4 sequences are provided at the back of the thesis, (on diskette) in the form of combined data files with the different datasets written in it. Photocopied versions of the thesis may not contain the diskette and the data can therefore be obtained from the author on request.

Cladistic Analysis

Because of the paucity of the available *rps4* dataset (i.e. sequences for a number of representative species of some of the putative sub-generic taxa were unavailable), it could not be subjected to separate analyses. Therefore only the *trnL–F* dataset was analysed separately. However, both datasets were included in the combined analyses. Species that had only *trnL* sequences were scored as having missing characters for the *rps4* genes sequences.

The molecular data had regions that were not easily alignable across the outgroup taxa and the species of *Andreaea*. Therefore, two types of combined *rps4/trnL* datasets were utilised for the analyses. For "complete" combined dataset (including all the species for which molecular data was available), the non-alignable portions of the sequences were excluded leaving only 819 alignable sequence positions compared to 984 available sequences for *Andreaea*. The excluded sequences (mostly *trnL*

sequences) were however, useful for resolving relationships within *Andreaea* and presented few alignment problems for the species of the genus. Therefore, a second, "reduced" combined dataset was assembled, including all the available sequences but with only *Andreaea* species and *Andreaeobryum* (as outgroup) added. This dataset comprised 984 sequence positions. *Andreaeobryum* was used as an outgroup based on results of the "complete" *trnL/rps4* data analysis (see results section), which placed it firmly outside the *Andreaea* clade, confirming that it is not embedded within the genus *Andreaea*.

Maximum parsimony analysis of the "complete" *rps4 / trnL–F* dataset

To test the monophyly of the genus *Andreaea*, the "complete" *rps4/trnL–F* dataset was analysed using the maximum parsimony method. The g_1 statistic (Heusenbeck, 1991), based on the distribution of 10,000 random tree lengths (constructed using the RANDOM TREES procedure) was utilised to test for the presence of significant phylogenetic signal in the "complete" *rps4/trnL–F* dataset. The analysis was conducted using the heuristic search option of PAUP* (test version 4.0b4a), using the tree bisection–reconnection (TBR) algorithm (with random addition sequence; 1000 replicates; branches having maximum length zero collapsed to yield polytomies; MULPARS and with STEEPEST DESCENT options in effect). The branch lengths of optimal trees were calculated using ACCTRAN optimisation (Farris, 1970) and all transformations were weighted equally.

Successive weighting, using the maximum value of the rescaled consistency index, was also utilised as a check on whether homoplasy (i.e. low consistency index) could obscure character information (Farris 1969) and also as an attempt to obtain a single or few most parsimonious trees (Farris, 1969, Carpenter, 1988). The successively weighted data was analysed using similar settings to those used for the un-weighted data. Support for the recovered phylogenies was evaluated using the jackknife support index as explained in chapter 2, for the morphological analyses.

Maximum likelihood analysis of the reduced combined *rps4/trnL–F* data could not be performed due to missing *rps4* data for some species. Analysis of such an incomplete combined dataset by maximum likelihood method would not be feasible because

models based on maximum likelihood approaches are computationally intensive and the analysis gets considerably slower with incomplete datasets.

Maximum parsimony analysis of the "reduced" *rps4 / trnL–F* dataset

A combined MP analysis of the "reduced" *trnL/rps4* molecular datasets was performed to determine what intra-generic phylogenetic relationships would be recovered from the combined molecular evidence, including all the available sequences for the species of *Andreaea*.

Owing to the advent of fast computers, it is now possible, even with relatively large datasets (>20), to perform exact methods such as the branch and bound search, which do not require all possible trees to be evaluated (Kitching *et al.*, 1998). In branch and bound search a tree is initially calculated usually by one of the heuristic methods (in this case stepwise addition sequence, where taxa are added to a developing tree), the length of which is taken as the upper bound for trees subsequently generated by the branch and bound process. The sequence of tree building and reconstruction then proceeds as for exhaustive search but the length of the tree is calculated as each new taxon is added. As soon as a tree is encountered where the length exceeds the upper bound, the path is abandoned because addition of more taxa can only further increase the length of the tree (Kitching *et al.*, 1998). The search then back tracks one node and tries a different path. In this way, the number of trees that must actually be evaluated is considerably reduced (Kitching *et al.*, 1998). The Branch swapping option was performed using settings similar to those used for the combined "complete"*trnL/rps4* analysis. Successive weighting was also performed as explained above and the support for the recovered phylogeny evaluated using bootstrap and Jackknife support indices as explained in chapter 2 for the morphological analysis.

Maximum Parsimony Analysis of the *trnL–F* sequences

The *trnL–F* data analysis was performed to determine the phylogenetic relationships within *Andreaea* based on this region alone. The sequence data for the *trnL–F* sequences were analysed separately using parsimony analysis as implemented in PAUP* test version 4.0b4a (Swofford, 1998). The g_1 statistic (Heusenbeck, 1991),

based on the distribution of 10,000 random tree lengths (constructed using the RANDOM TREES procedure) was determined to test for the presence of significant phylogenetic signal in the dataset.

The analysis was conducted using the branch and bound search option of PAUP* (test version 4.0b4a), with settings similar to those applied for the analysis of the "reduced" *trnL–F/rps4* dataset as explained above. Successive weighting was also performed as explained above. The support for recovered phylogenies was evaluated using Bootstrap and Jackknife support indices.

Maximum likelihood analysis of the *trnL–F* dataset.

When data involves moderate to large amounts of change (as is common in molecular data), parsimony methods can fail in their estimate of phylogeny (Felsenstein, 1981). Use of parsimony methods would be well justified if the rates of change along lineages are at least approximately equal, but this may not be true for molecular data (i.e. the rates of evolution may differ in different lineages), resulting in inconsistent estimates of the evolutionary tree when amounts of change are sufficiently unequal (Felsenstein, 1978, 1981).

Therefore the maximum likelihood (ML) approach (Felsenstein, 1981, 1983) has been utilised to further check the results of parsimony analyses and basically as a better approach for molecular sequence data as it takes into account the different rates of change along branches (Lewis, 1998). ML utilises appropriate probabilistic models that also allow the calculation of confidence intervals for branch lengths, in this way, also allowing for the possibility of heterogeneity of branch lengths. These models are therefore suitable for specific patterns and rates of molecular sequence evolution for different datasets (Lewis, 1998). An assumption of equal branch lengths as implied by MP analysis would lead to problems of long branch attraction (e.g. Felsenstein, 1978, Hendy and Penny, 1989), a phenomenon that occurs when a dataset contains some taxa that have accumulated relatively more changes than other taxa in the dataset, resulting in the MP algorithms grouping the long branches together even if this grouping does not reflect the actual phylogenetic history. Further ML analyses

allow for exploration of the data before estimating parameters of the model from the data.

Basically the principle of ML involves evaluating the probability that a particular tree could have generated a given dataset. The evaluation proceeds as follows: for any given site the probability (likelihood) of having particular character states (an A, a C, a T or a G) are computed. Having computed the likelihoods at each site, the joint probability that the tree and model confer upon all sites is computed as the product of the individual site likelihoods. The result with the highest likelihood is taken as the best estimate of the phylogeny (Felsenstein, 1981)

Maximum likelihood (ML) analyses of the *trnL–F* sequences was performed in PAUP* using the general time reversible model of nucleotide substitution (GTR).

The rate variation among sites was assumed to follow a gamma (Γ) distribution ('dG'). Obtaining likelihoods by integrating over the Γ distribution is computationally intensive (Yang, 1994). Therefore, the Γ distribution was divided into four approximate rate categories (Yang, 1994) that produce boundaries such that each category has equal probability. The median (middle, other than mean rate) was used to represent all the rates within a particular category. The gamma distribution parameter α (the shape parameter) is an indication of the rate of variation across sites (Yang et al., 1995).

Tree scores for the most parsimonious tree obtained from the analysis of the *trnL–F* molecular dataset were evaluated under the GTR + dG substitution model, and the model parameter estimates of the most parsimonious tree were fixed in subsequent ML analyses. This allowed the ML analysis to be performed under a model that matches the pattern and rates of nucleotide substitution present in the *trnL–F* dataset.

Support for the estimated phylogeny using the ML approach, was determined using likelihood Jackknife technique as implemented in PAUP* version 4.0b4a (Swofford, 1998).

RESULTS

Maximum parsimony analysis of the "complete" *trnL–F/rps4* dataset

The distribution of tree lengths for 10, 000 randomly generated trees was significantly left skewed (g_1 = -0.648; $p < 0.05$), suggesting the presence of a phylogenetic signal in the "complete" *trnL–F/rps4* dataset (Heulsenbeck, 1991). Of the 812 characters included, 409 were variable and 217 were parsimony informative. The analysis of the un-weighted complete *trnL–F/rps4* data resulted in 5354 trees (L = 757; CI = 0.686; CI excluding uninformative characters = 0.559; RI = 0.704; RC = 0.484). One of the most parsimonious trees recovered is presented in figure 4. The genus *Andreaea* was resolved, and very strongly supported (jackknife = 87%), as monophyletic. *Andreaeobryum* was resolved, without any jackknife support, as sister to the peristomate mosses. Successive re-weighting however resulted a strongly supported (jackknife = 87%) placement of *Andreaeobryum* in a clade with the peristomate mosses (see figure 4). Support for the monophyly of the genus *Andreaea* was also improved (to Jackknife = 92%) by the successive reweighting. There were no other interesting changes in the topology upon successive reweighting. Only some of the relationships within the genus *Andreaea* clade were resolved differently without any change in support. The relationship of *Andreaea* with other genera remained the same.

Maximum parsimony analysis of the "reduced" *trnL–F/rps4* dataset

The distribution of tree lengths for 10, 000 randomly generated trees was significantly left skewed (g_1 = -0.58; $p < 0.05$), suggesting the presence of a phylogenetic signal in the "reduced" *trnL–F/rps4* dataset (Heulsenbeck, 1991). Of the 984 characters included, 124 were variable and 53 were parsimony informative. Maximum parsimony analysis of un-weighted the "reduced" *trnL–F/rps4* dataset yielded 36 most parsimonious trees (length = 216, consistency index (CI) = 0.752, CI excluding uninformative characters = 0.562, rescaled consistency index (RC) = 0.562, and retention index (RI) = 0. 747). One of these most parsimonious trees is presented in Figure 5.

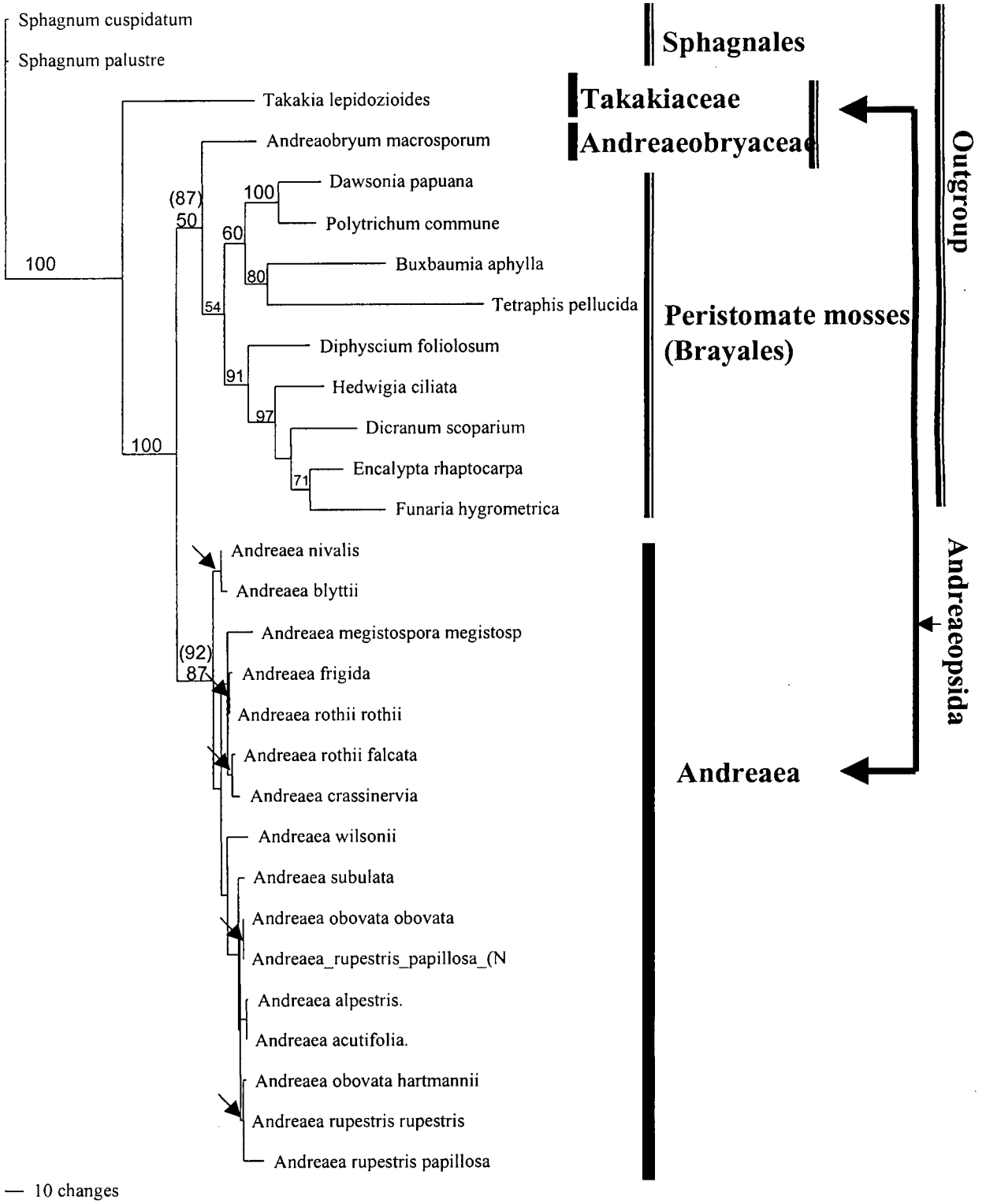


Figure 4: A phylogram chosen arbitrarily from the 4354 most parsimonious trees obtained from a combined analysis of the "complete" datasets of *trnL* and *rps4* regions (i.e. with all the available ingroup and outgroup taxa included). The numbers above branches are Jackknife values of interest, whereas those in parenthesis are jackknife support values after successive re-weighting. Arrows indicate nodes that collapse in the strict consensus tree. The classification is according to Vitt (1984) and Smith and Davidson (1993).

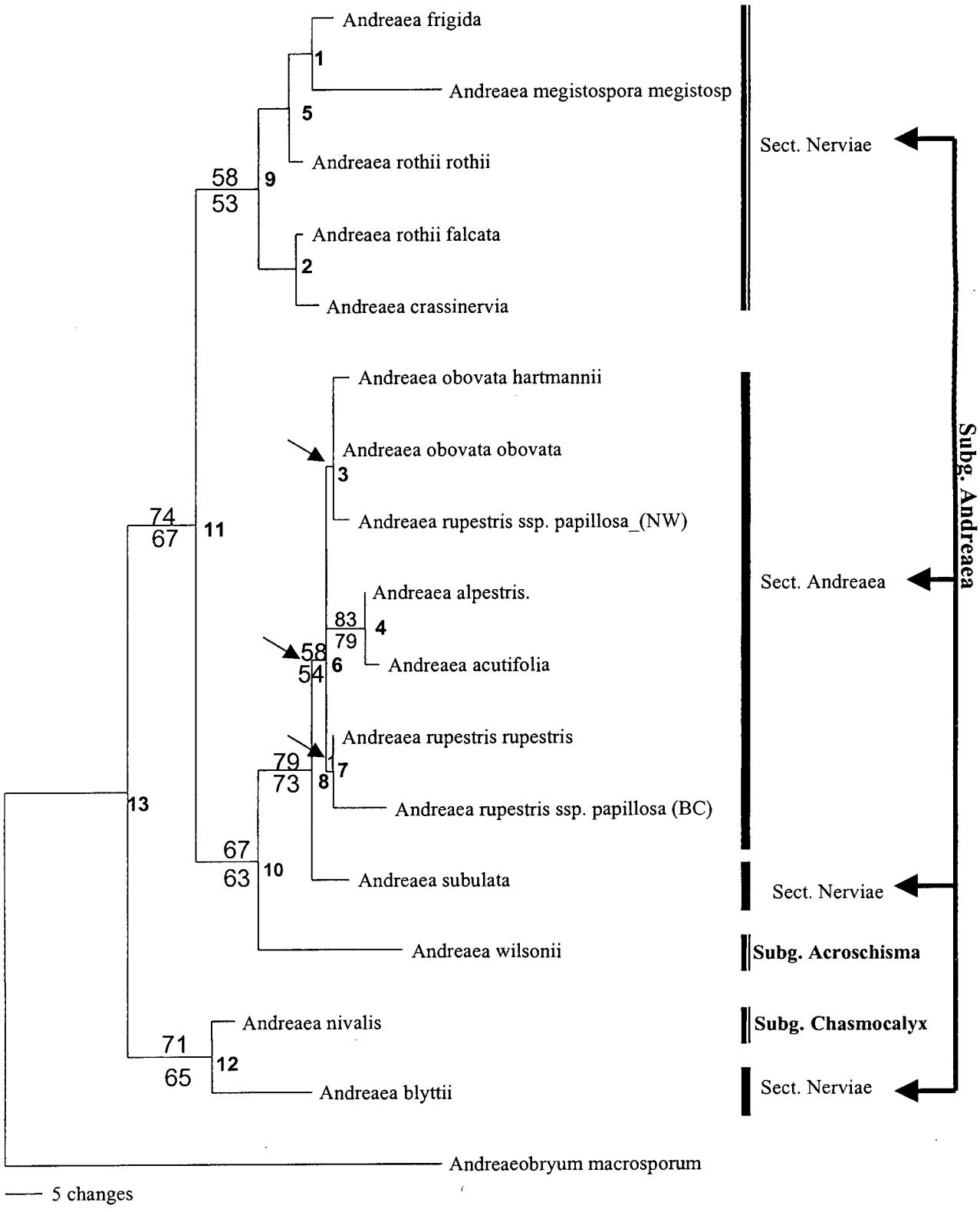


Figure 5: One of the trees, chosen arbitrarily from the 36 MPTs obtained in an analysis of unweighted "reduced" *trnL/rps4* molecular data. Numbers above branches jackknife values, whereas those below are bootstrap values. Node numbers are indicated. Arrows indicate nodes that collapse in the strict consensus. The classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

Three main clades were resolved within *Andreaea*. A moderately supported lineage (jackknife = 71%; Bootstrap = 65) comprising *A. nivalis / A. blyttii* clade (node 12), is coordinate to a well-supported clade (jackknife = 67%; Bootstrap = 74%) comprising the rest of the taxa. Within this latter clade two large groups are resolved, with the larger and fairly supported clade (node 10) comprising a grade of *wilsonii* and *A. subulata*, the latter being coordinate to the species of the section *Andreaea* (node 6). Within the section *Andreaea* clade the sister relationships of *A. alpestris* and *A. acutifolia* is strongly supported (jackknife = 83% and bootstrap = 79%), whereas other relationships in the clade are unsupported. The second and smaller group (node 9) is weakly supported and comprises *A. frigida, A. megistospora* ssp. *megistospora, A. crassinervia*, and the two subspecies of *A. rothii*. All relationships within the clade are however, not supported by either bootstrap or jackknife values. Successive weighting had no effect on the topologies recovered from an analysis of the *trnL–F/rps4* data. A number of species exhibited relatively long branch lengths as shown in Figure 5, i.e. *A. megistospora* (17), *A. wilsonii* (19), *A. nivalis* + *A. blytii* (11) and *A. blytii* (13). It was however not possible to perform a maximum likelihood analysis (see methods section), which as explained in Chapter 3 would be appropriate to test the placement of these taxa (especially the *A. nivalis / A. blyttii* clade) by maximum parsimony analysis.

Maximum parsimony analyses of the *trnL–F* dataset

Of the 470 characters included in the *trnL–F* data, 362 were constant and 36 were parsimony informative. The distribution of tree lengths for 10,000 randomly generated trees was significantly left skewed ($g_1 = -0.72$; $p < 0.05$), suggesting the presence of phylogenetic signal in the *trnL–F* dataset. Maximum parsimony analysis of the *trnL–F* dataset yielded 45 most parsimonious trees (length = 138, consistency index (CI) = 0. 8696, CI excluding uninformative characters = 0.714, Rescaled consistency index (RC) = 0.6832 and retention (RI) = 0.7857. One of the most

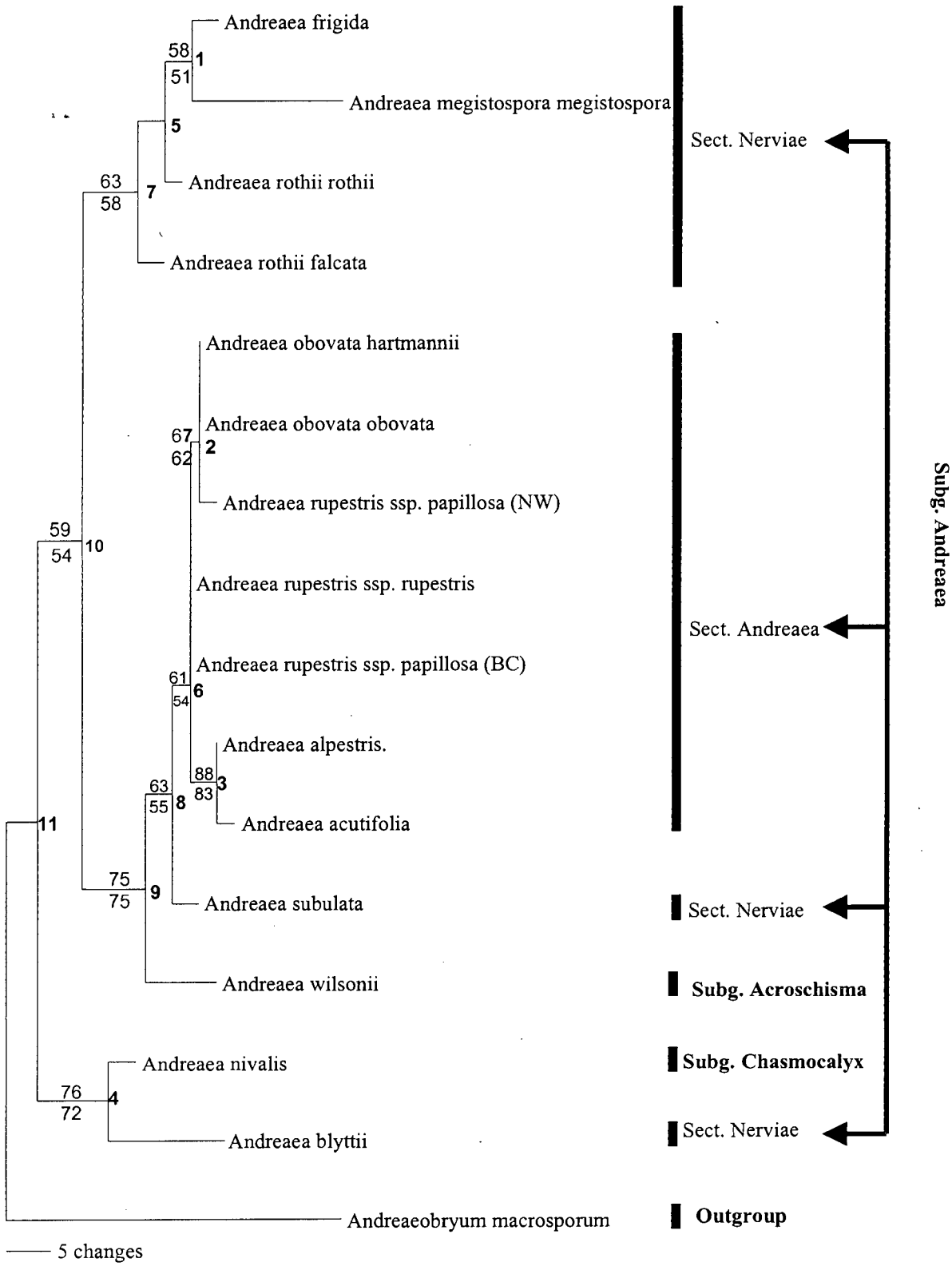


Figure 6: phylogram of one of the most parsimonious trees recovered from an analysis of the unweighted *trnL-F* dataset. This tree is exactly the same as the consensus tree. Numbers above the branches are Jackknife values whereas those below are bootstrap values. Node numbers are indicated inside the nodes. The classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

parsimonious trees recovered from the analysis of unweighted *trnL–F* data is presented in Figure 6. This tree was exactly the same as the strict consensus tree. This tree is also similar to the tree recovered from the analysis of the "reduced" *trnL/rps4* dataset. The two differ in the resolution of the clade comprising ecostate species and *A. subulata*. The clade (node 9) consisting of all the ecostate species (section *Andreaea*), *A. wilsonii* and the costate *A. subulata* (putatively of section *Nerviae*) was better supported (bootstrap 75% and Jackknife = 75%) in the separate *trnL–F* MP analysis than in the reduced *trnL/rps4* analysis. The *A. obovata/A. rupestris ssp papillosa* clade was also better supported (bootstrap = 62% and Jackknife = 67%) in the *trnL–F* MP analysis than in the reduced *trnL/rps4* analysis. Successive reweighting had no effect on the topologies from the *trnL* MP analysis.

Maximum likelihood analysis of the *trnL–F* dataset

Evaluation of the *trnL–F*–MPT provided the following estimates of the GTR+dG model parameters: r-matrix = (6.272, 10.781, 1.730, 4.584, 14.805, 1) corresponding respectively to A-C, A-G, A-T, C-G, C-T, and G-T substitution types (G-T was arbitrarily set to 1), proportion of invariant sites (pinvar) = 0.773, shape parameter of gamma distribution (Shape) = 0.29 (PAUP* notation). Figure 7 shows the most likely tree (Ln-Likelihood = -1410). This tree differs from the *trnL–F* MPT in a number of aspects. *A. nivalis* is not resolved in a clade with *A. blyttii*, but the two form a grade with *A. blyttii* being sister to the remaining species (node 9). Within the major clade resolved (node 9), *A. rothii* ssp. *falcata* and *Andreaea megistospora* are sister to clade in which *A. frigida* and *A. rothii* ssp. *rothii* are coordinate with the remaining taxa (node 6). In this latter group, *A. wilsonii* is sister to the rest of the species. Unlike the results of parsimony analysis, *A. subulata* is resolved within the presumed section *Andreaea*. However, the only supported clade in the whole topology is the *A. alpestris* / *A. acutifolia* clade.

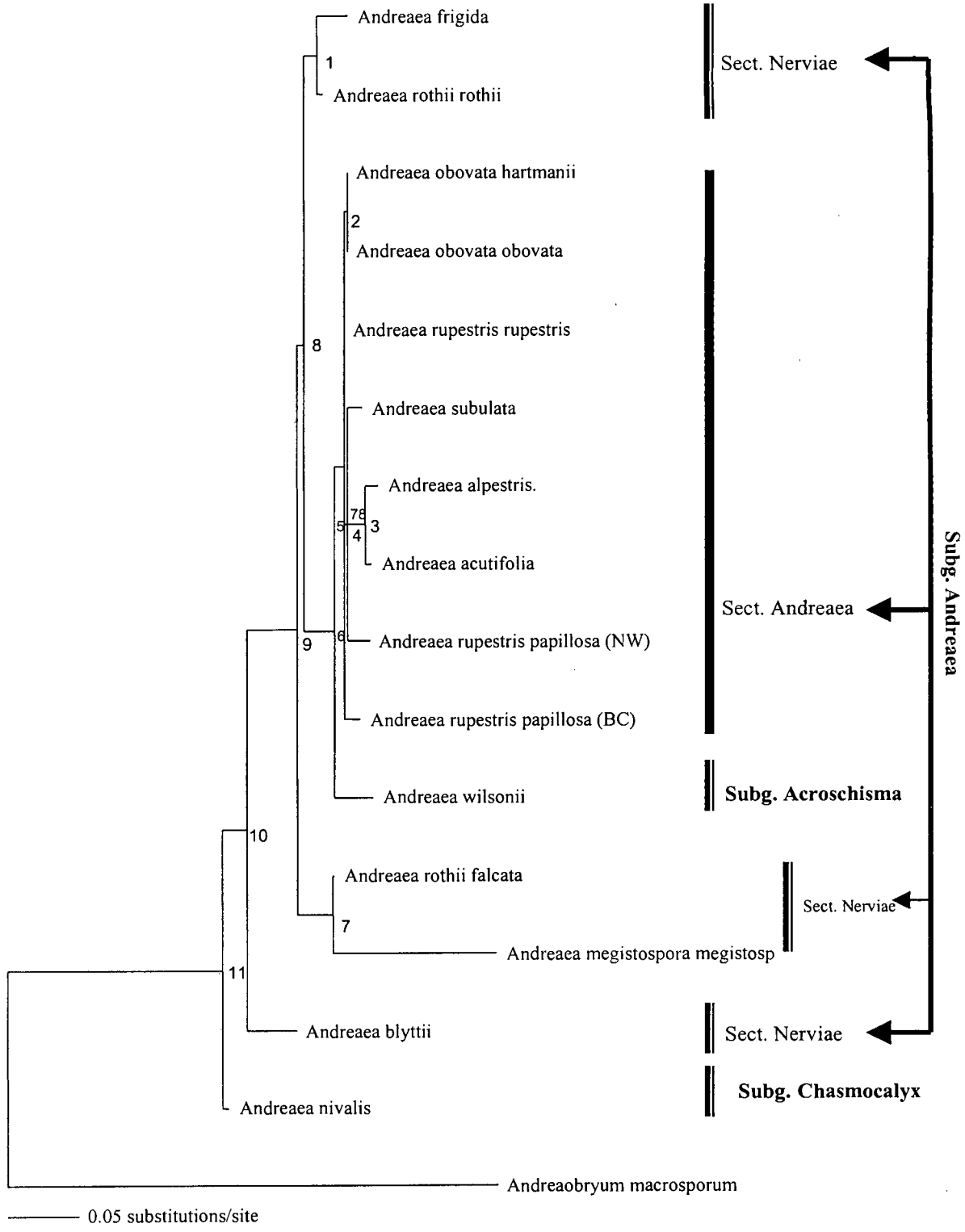


Figure 7: The maximally likely (Ln-likelihood = -1410) tree obtained from the analysis of *trnL* sequences from 15 species of the *Andreaea* and 1 species of *Andreaeobryum*. Number above the branch supporting *A. acutifolia* / *A. alpestris* clade is a Jackknife percentage (based on 100 replicates). Node numbers are shown inside each node (except for node 5, outside the node). Classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

DISCUSSION OF THE MOLECULAR ANALYSES

The analysis of the "complete" *trnL/rps4* was aimed at determining the position of *Andreaea* relative to the other moss groups. The genus *Andreaea* was strongly supported as monophyletic by the analysis of the molecular data. However its close relationship to the other members of Andreaeopsida sensu Smith and Davidson (1993), was not supported. *Andreaeobryum* emerged as strongly related to the peristomate mosses.

The maximum parsimony analysis of both the reduced *trnL/rps4* as well as the separate maximum parsimony and maximum likelihood analyses of the *trnL–F* data were mainly aimed at recovering relationships within the genus *Andreaea*.

Analysis of the "reduced" *trnL–F/rps4* dataset did not support a number of putative infra-generic relationships (Figure 5). Within the subgenus *Andreaea*, the presumed section *Nerviae* was not supported. This section was instead resolved as polyphyletic. The section *Andreaea* of the same subgenus was weakly supported as monophyletic (node 6 of figure 5) by the "reduced" *trnL–F/rps4* dataset analysis.

The subgenus *Chasmocalyx*, represented by *A. nivalis* was not resolved within any of the other subgeneric groups by the *trnL–F/rps4* analysis, but as a separate lineage that was moderately supported (bootstrap = 65%, jackknife =71%) as sister to *A. blyttii*, with the *A. nivalis* / *A. blyttii* clade being resolved as sister to the clade containing the rest of the ingroup species.

The currently monotypic subgenus *Acroschisma* (represented by *A. wilsonii*) was resolved as a separate lineage coordinate to the well supported (bootstrap = 73%, jackknife = 79%) *A. subulata*/section *Andreaea* clade (node 8 of Figure 5), suggesting that it may be a separate lineage closely related to the ecostate species. The costate *A. subulata* (section *Nerviae*) was interestingly resolved as coordinate to the section *Andreaea* in this analysis and this may suggest that it may also probably represent another separate lineage coordinate to the ecostate species of the genus.

The lack of confident resolution of a number of relationships by the *trnL/rps4* may be due to lack of information in the dataset for the resolution of these relationships. There were only 53 parsimony informative characters in the dataset. It is possible that most of these are only informative about relationships in certain parts of the topology and not others. Addition of more information from other datasets can be expected to improve the support for the right topology. In addition further sampling among the species of *Andreaea* may improve resolution.

The topology recovered from the MP analysis of the *trnL–F* dataset (Figure 6) was very similar to the topology from the reduced *trnL/rps4* analysis (figure 5). The only difference was that in the *trnL–F* analysis, the weakly supported *A. obovata/A. rupestris* ssp. *papillosa* (from Norway (NW)) clade was present in the section *Andreaea* clade, whereas for the reduced *trnL/rps4*, it was collapsed in the strict consensus tree. Whereas the clade consisting of the whole section *Andreaea*, *A. subulata* and *A. wilsonii* was moderately supported by the *trnL–F/rps4* analysis, it was better supported (bootstrap = 75%, jackknife = 75%) in the *trnL–F* maximum parsimony analysis. Other parts of the topology though well resolved were not well supported. This was probably due to the low level of phylogenetically informative characters since the homoplasy level was not too high to obscure character information. Re-weighting did not affect results of either the reduced *trnL–F/rps4* or the *trnL–F* maximum parsimony analysis, indicating that homoplasy had not obscured any phylogenetically useful information in the un-weighted analysis. The section *Nerviae* is apparently the most incoherent group, as members of the group are resolved in several lineages. There is thus no strictly dichotomous split of costate and ecostate species.

Unlike the *trnL–F* maximum parsimony analysis, the maximum likelihood analysis (figure 7) did not resolve *A. subulata* as basal or coordinate to the traditional section *Nerviae*, but as embedded within the section *Andreaea*. However this position is not supported by ML-jackknife. Further interestingly ML analysis did not resolve the two subspecies of *A. rothii* within one clade, but in different lineages (node 7 and node 1). Unlike the MP analysis ML did not resolve the *A. nivalis / A. blyttii* relationship as a clade but a grade. All these resolutions were however also un-supported by ML

jackknife. Both MP and ML analyses nevertheless, suggested that some members of section *Nerviae* possess pleisiotypic character states.

As shown in Figure 6, a number of species exhibited relatively long branch lengths. However since most of the changes involved are autapomorphic, the placement of the taxa may not be affected by the long branch attraction problem (Felsenstein, 1978), except for the *A. nivalis* + *A. blyttii* clade. It would have been expected that a maximum likelihood analysis would have given better support for the most likely relationships. However, most parts of the *trnL–F* ML topology are unsupported.

CONCLUSIONS

The genus *Andreaea* is monophyletic. Its presumed closer relationship to *Andreaeobryum* and *Takakia* than to other mosses (e.g. Murray, 1988a, Smith and Davidson, 1993) is contradicted by the combined molecular evidence, more strongly so for *Takakia*. Most of the currently recognised infra-generic groups are resolved as non-monophyletic. Three major groups emerge from both the "reduced" trnL/rps4 and *trnL–F* maximum parsimony analyses, though only two are firmly resolved in this analysis; the clade including *A. wilsonii*, *A. subulata* and all the ecostate species (section *Andreaea*) and the *A. nivalis* / *A. blyttii* clade. The clade consisting of *A. rothii*, *A. megistospora*, *A. crassinervia* and *A. frigida* was not strongly supported. However since the other clades in which these species could alternatively be resolved are strongly supported, it is possible that with more information this weakly supported clade could be corroborated. Definitely species of section *Andreaea* (subgenus *Andreaea*) share most recent common ancestry. What was unresolved was whether the clade is paraphyletic or not, since *A. subulata* (section *Nerviae*) was weakly embedded in this section by ML analysis and relationships between this group, *A. wilsonii* and *A. subulata* are not firmly resolved. It is clear that *trnL–F* and the coding region of the *rps4* sequences provide some but not enough characters on their own to unambiguously resolve all the relationships within the genus *Andreaea*.

CHAPTER 4

CLADISTIC ANALYSIS OF *ANDREAEA* BASED ON COMBINED MORPHOLOGICAL AND MOLECULAR DATA

INTRODUCTION

The possible advantages of combining datasets have already been alluded to in chapter 3. The current chapter presents analyses that utilised the total evidence of the study to determine the relationships of the genus *Andreaea* to the other major mosses groups and also to determine the infra-generic relationships. The monophyly of the genus as well as that of the infra-generic groups was thus also tested based on these total evidence analyses.

MATERIALS AND METHODS

Taxon sampling and datasets

The same data utilised for the *trnL/rps4* analyses in chapter 3 were combined with the morphological data for this analysis. Species for which only morphological data were available (i.e. *A. australis*, *A. nitida*, *A. bistratosa*, *A. megistospora* ssp. *epapillosa*, *A. alpina* and *A. gainii*) were also included. As in the analyses of combined molecular data in chapter 3, two types of combined data were assembled, the first one (the "reduced" *trnL/rps4*/morphology dataset) containing all the available sequences for *Andreaea* and *Andreaeobryum* only, and a second one ("complete" *trnL/rps4*/morphology) containing only the sequences that were easily alignable across the ingroup and all the outgroup taxa included in this study (see table 5).

Cladistic analysis of the combined morphological and molecular datasets.

For both the "complete" and "reduced" *trnL/rps4*/morphology datasets, the g_1 statistic (Heulsenbeck, 1991), based on the distribution of 10,000 random tree lengths (constructed using the RANDOM TREES procedure) was utilised to test for the presence of significant phylogenetic signal. Combined analyses of morphological and molecular datasets were performed, in PAUP* (test version 4.0b4a), using the branch and bound search option for the "reduced" combined analysis and the heuristic search option for the larger "complete" combined dataset. In both cases, the branch swapping option was performed using the tree bisection–reconnection (TBR) algorithm, via the random addition sequence; with 1000 replicates, the MULPARS option in effect and STEEPEST DESCENT on. Branch lengths of optimal trees were calculated using ACCTRAN optimisation (Farris, 1970). All transformations were weighted equally (Fitch parsimony; Fitch, 1971, Hartigan, 1973).

Successive weighting was performed using the maximum value of the rescaled consistency index. The successively weighted data were analysed using similar settings to those used for the un-weighted data. Support for the recovered phylogenies was evaluated using the bootstrap and jackknife support indices as explained in the preceding chapters.

RESULTS OF COMBINED MORPHOLOGY AND MOLECULAR ANALYSES

Analysis of the "complete" *trnL–F/rps4*/morphology dataset

The distribution of tree lengths for 10, 000 randomly generated trees was significantly left skewed (g_1 = -0. 63; $p < 0.05$), suggesting the presence of a phylogenetic signal in the "complete" *trnL–F/rps4* / morphology dataset. Of the 905 characters included, 497 were variable and 215 were parsimony informative. The analysis of the un-weighted "complete" *trnL–F/rps4*/morphology data resulted in 269 trees (L = 912; CI = 0.666; CI excluding uninformative characters = 0.550; RI = 0.708; RC = 0.471). One of the most parsimonious trees is given in figure 8. The genus *Andreaea* was

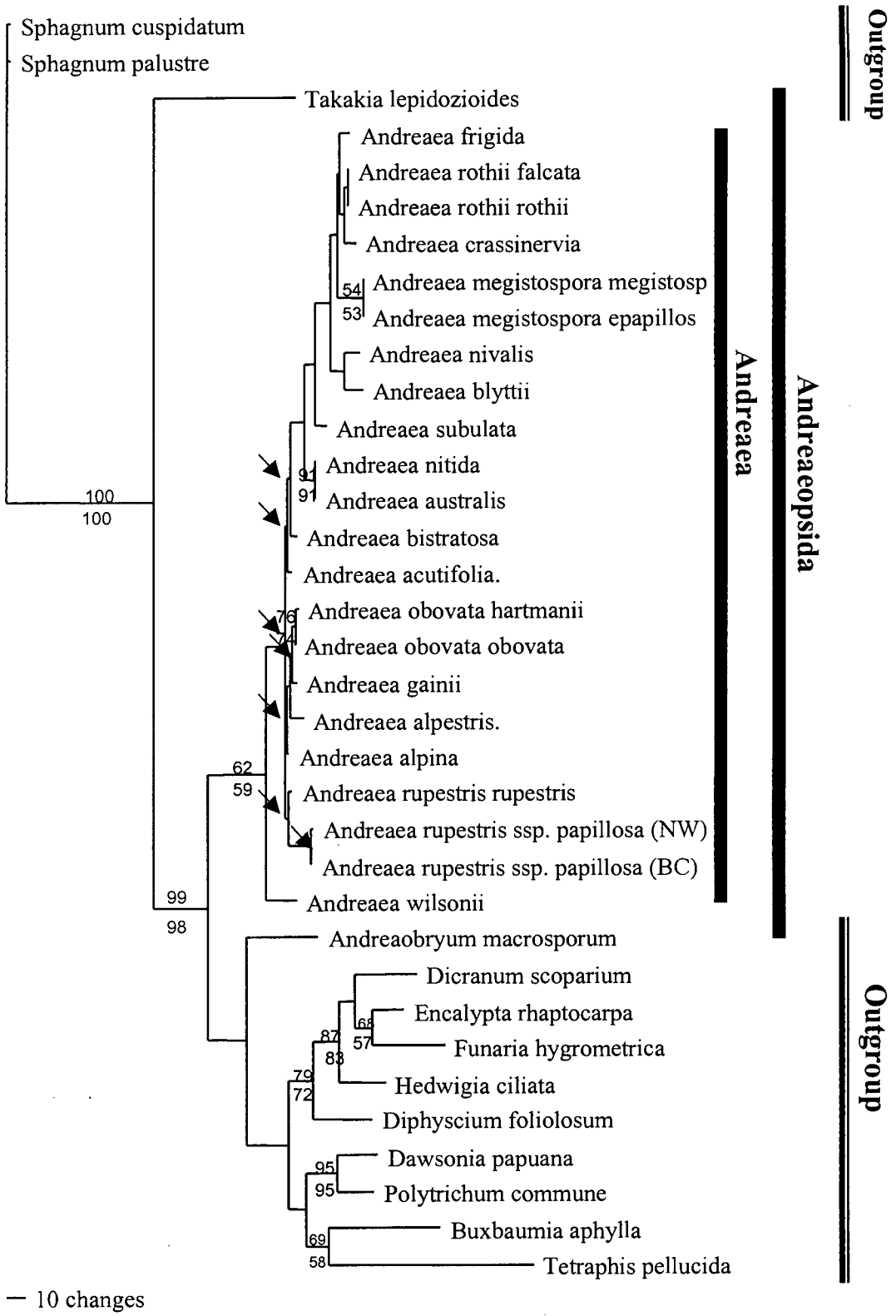


Figure 8: A most parsimonious tree chosen arbitrarily from the 269 MPTs obtained from an analysis of the "complete" datasets of *trnL* and *rps4* regions, and morphology (i.e. with all the available in group and outgroup taxa included). The numbers above branches are Jackknife values of interest whereas those below are bootstrap values. The classification is according to Vitt (1984) and Smith and Davidson (1993).

resolved, but not strongly supported (bootstrap = 59% and jackknife = 62%), as monophyletic. *Andreaeobryum* was resolved as sister to the peristomate mosses but without support. Successive re-weighting resulted in a reduction in number of trees recovered but not in any interesting changes in resolution (Figure 9). Only some of the relationships within the genus *Andreaea* were resolved differently. The relationship of *Andreaea* with other genera remained the same. Support for the monophyly of the genus *Andreaea* and the *Andreaeobryum*/persitomate mosses clade was only improved very slightly (bootstrap = 58%, jackknife = 63%).

Analysis of the "reduced" *trnL–F/rps4*/morphology dataset

The distribution of tree lengths for 10,000 randomly generated trees was significantly left skewed ($g_1 = -0.45$; $p < 0.05$), suggesting the presence of a phylogenetic signal in the *trnL–F/rps4* dataset (Heulsenbeck, 1991). Of the 1070 characters included, 238 were variable and only 84 were parsimony informative. Maximum parsimony analysis of the reduced *trnL–F/rps4*/Morphology dataset recovered 25 equally most parsimonious trees (length = 334, CI = 0.767, CI excluding uninformative characters = 0.4407, RI = 0.731 and RC = 0.5603). These trees were not very congruent as indicated by the nodes that collapse (see figure 10) in the strict consensus tree. However a few clades are well supported in the analysis of the un-weighted combined data. The fairly supported clades included the *A. obovata* var. *obovata* / *A. obovata* var. *hartmannii* clade (node 10; bootstrap = 72%, Jackknife = 76%) and the *A. bistratosa* / *A. sinuosa* clade (node 18; bootstrap = 67%, Jackknife = 73%), whereas the section *Nerviae* / *A. subulata* clade was very strongly supported (bootstrap = 86%, Jackknife = 94%). However the relationships of these clades to the other taxa in the genus were ambiguously resolved.

The highly ambiguous resolution from the analysis of un-weighted data was probably due to the high homoplasy as indicated by the low CI excluding uninformative characters (0.44). This fact was corroborated by the effect of successive re-weighting. After one round of successively re-weighting, topologies converged, resulting in 1

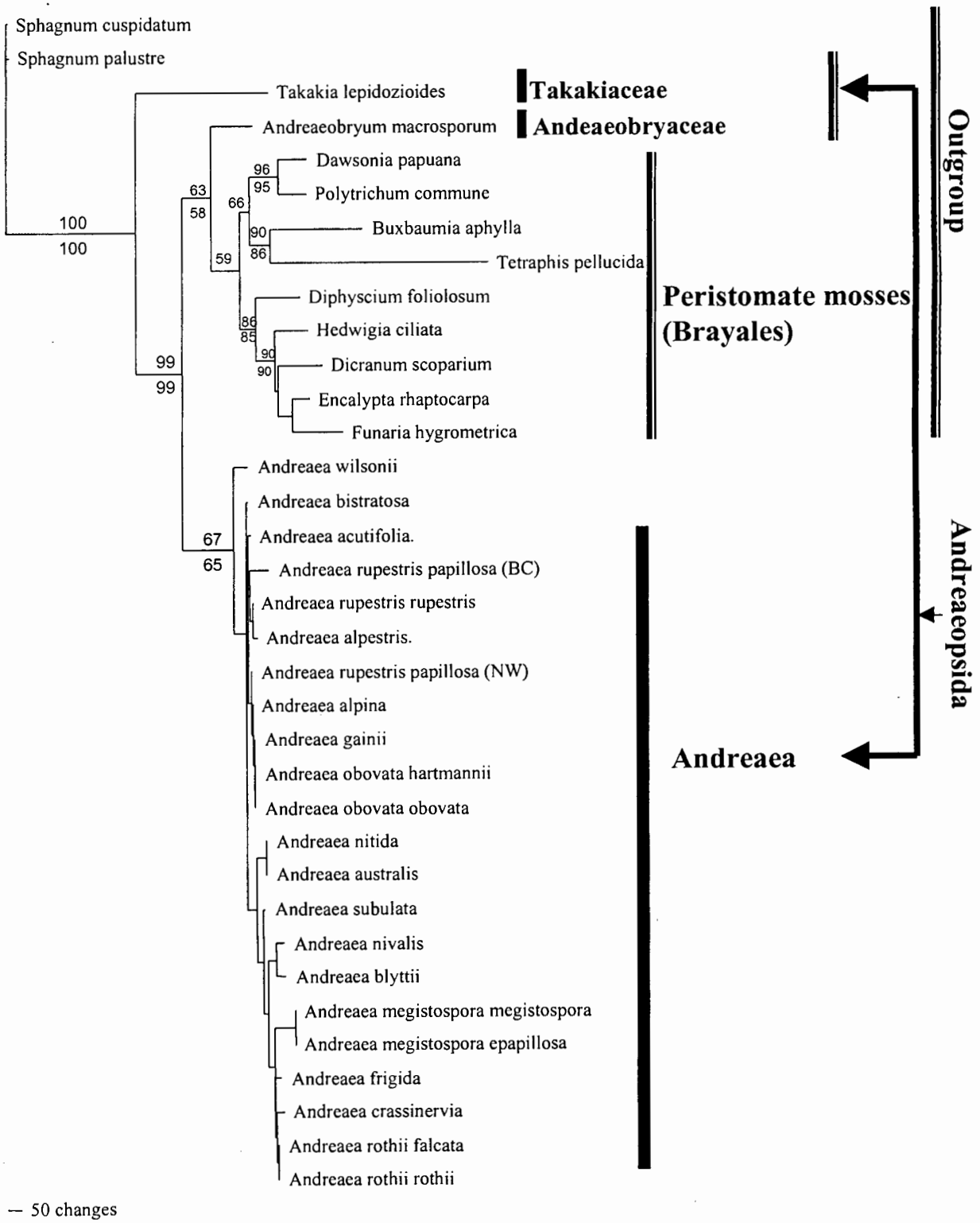


Figure 9: One of the most parsimonious trees obtained from an analysis of the successively re-weighted "complete" dataset of *trnL* and *rps4* regions, and morphology i.e. with all the available in group and outgroup taxa included. The numbers above branches are Jackknife values of interest whereas those below are bootstrap values. The classification is according to Vitt (1984) and Smith and Davidson (1993).

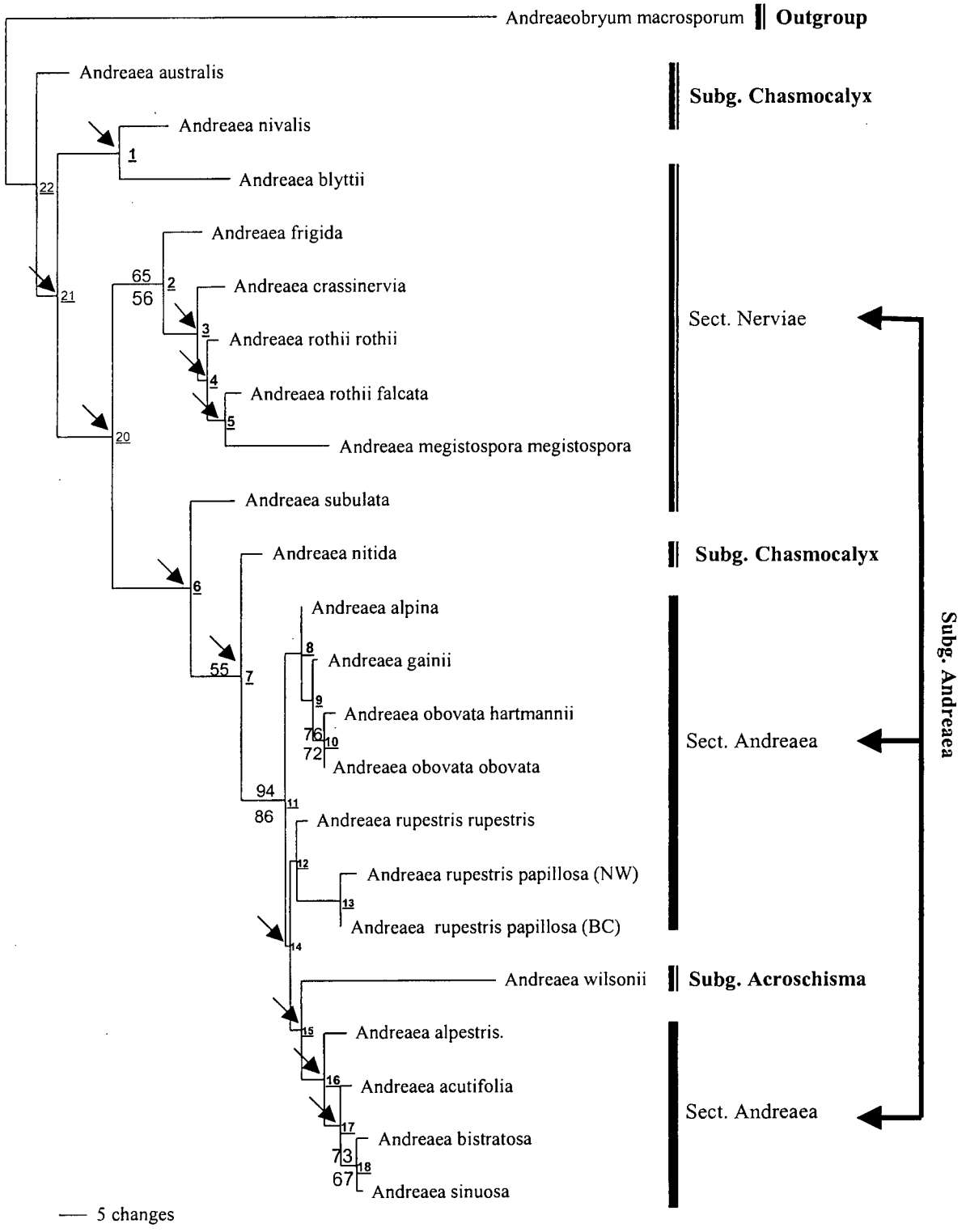


Figure 10: One of the most parsimonious trees recovered from an analysis of a "reduced" *trnL/rps4*/morphology datasets (i.e. with only *Andreaeobryum* as the outgroup). Numbers above the branches represent jackknife value whereas those below represent Bootstrap values. Arrows represent nodes that collapse in the strict consensus tree. Node numbers are indicated as bold type inside the nodes. The classification was derived from Murray (1988), Braithwaite (1887) and Matteri and Farias (1999).

fully resolved tree (Figure 11; length = 2224, CI = 0.7665, CI excluding uninformative characters = 0.559, RI = 0.731 and RC = 0.56). The resolution and support for the tree also improved greatly. In this re-weighted tree, *A. nivalis* was resolved as coordinate to the strongly supported (bootstrap = 85%, Jackknife = 88%) clade containing the rest of the species of *Andreaea* (node 20). *A. nivalis* and *A. blyttii* form a grade with *A. blyttii* coordinate to a moderately supported (bootstrap = 67%, Jackknife = 74%) clade of the remaining species (node 18). Within the latter clade (node 18), two other lineaged emerge (nodes 2 and 5). The well-supported (bootstrap = 74%, Jackknife = 77%) smaller clade (node 2), contained *A. frigida* / *A. megistospora* ssp. *megistospora*, *A. rothii* ssp. *rothii*, *A. rothii* ssp. *falcata* and *A. crassinervia*. The relationships of these species were however, not supported by bootstrap or Jackknife values.

The other moderately supported and larger clade (node 5) contained a basal grade of *A. subulata*, *A. australis*, *A. nitida*, *A. bistratosa* and *A. sinuosa* (nodes 5, 6, 7, 8 and 9 respectively). The most terminal species of this grade, *A. sinuosa*, was sister to the unsupported clade of the remaining ecostate species (node 13). Within this clade moderately supported relationships included the clade of *A. obovata* varieties (node 12: bootstrap = 75%, Jackknife = 78%) and the *A. alpestris*/*A. acutifolia* clade (node 17): bootstrap = 69%, Jackknife = 74%). All other relationships were unsupported. The monotypic subgenus *Acroschisma* represented by *A. wilsonii* was resolved within the ecostate species group (traditional section *Andreaea*) (see node16).

DISCUSSION OF THE COMBINED MORPHOLOGICAL AND MOLECULAR DATA ANALYSES

The "complete" *trnL–F*/*rps4*/morphology dataset analysis

The analysis of the "complete" combined *trnL–F*, *rps4* and morphological data was essential in providing an understanding of the placement of the genus *Andreaea* among the other genera of mosses. However though there was significant phylogenetic signal in the data (g_1 = -0.63; $p < 0.05$), the placement of *Andreaeobryum* in relation to *Andreaea* could not be firmly established on the basis of the "complete" *trnL*/*rps4*/morphology analysis. The monophyly of *Andreaea* was only moderately supported by the analysis of the successively re-weighted dataset.

Andreaeobryum though resolved outside the *Andreaea* clade is weakly supported (Figures 8 and 9). However since it was not resolved within the *Andreaea* clade, and also based on results of the "reduced" *trnL/rps4* analysis in chapter 3, *Andreaeobryum* is a potential outgroup taxon for polarising relationships within the *Andreaea* clade. The weakly supported placement of *Andreaeobryum* outside the *Andreaea* clade is probably only due to lack of enough information at the appropriate level of variation. Other previous molecular studies (e.g. Newton et al., 2000) have also only weakly resolved *Andreaeobryum* as sister to *Andreaea*.

Takakia was strongly resolved as sister to the *Sphagnum* species and not as closely related to *Andreaea*, even with the current addition of more species of *Andreaea* in the analysis. This is in agreement with previous studies in which it has also been resolved as sister to *Sphagnum* (Hedderson et al., 1998 and Newton et al., 2000). *Takakia* and *Sphagnum* however, share no easily discernible morphological similarities.

The analyses of the reduced *trnL–F/rps4*/morphology dataset

Adding morphological characters to the maximum parsimony analysis of the "reduced" *rps4/trnL–F* dataset resulted in lower topological congruence among the recovered most parsimonious trees (See nodes that collapse in strict consensus, Figure 10), than among those retrieved by MP analysis of the *rps4/trnL–F* dataset. This was probably due to conflicting signals of some morphological and molecular characters (CI excluding uninformative characters = 0.44), resulting in noise in the dataset masking the effect of non-homoplasious characters. However, with lowered noise level upon successive re-weighting, there was enough signal to establish a fairly reliable phylogeny (figure 11). Also unlike in many cases of combined molecular and morphological analysis, the recovered topology was not appreciably more dependent on molecular characters than morphological ones, since there were only 53 informative molecular (*trnL/rps4*) characters and 31 for morphology.

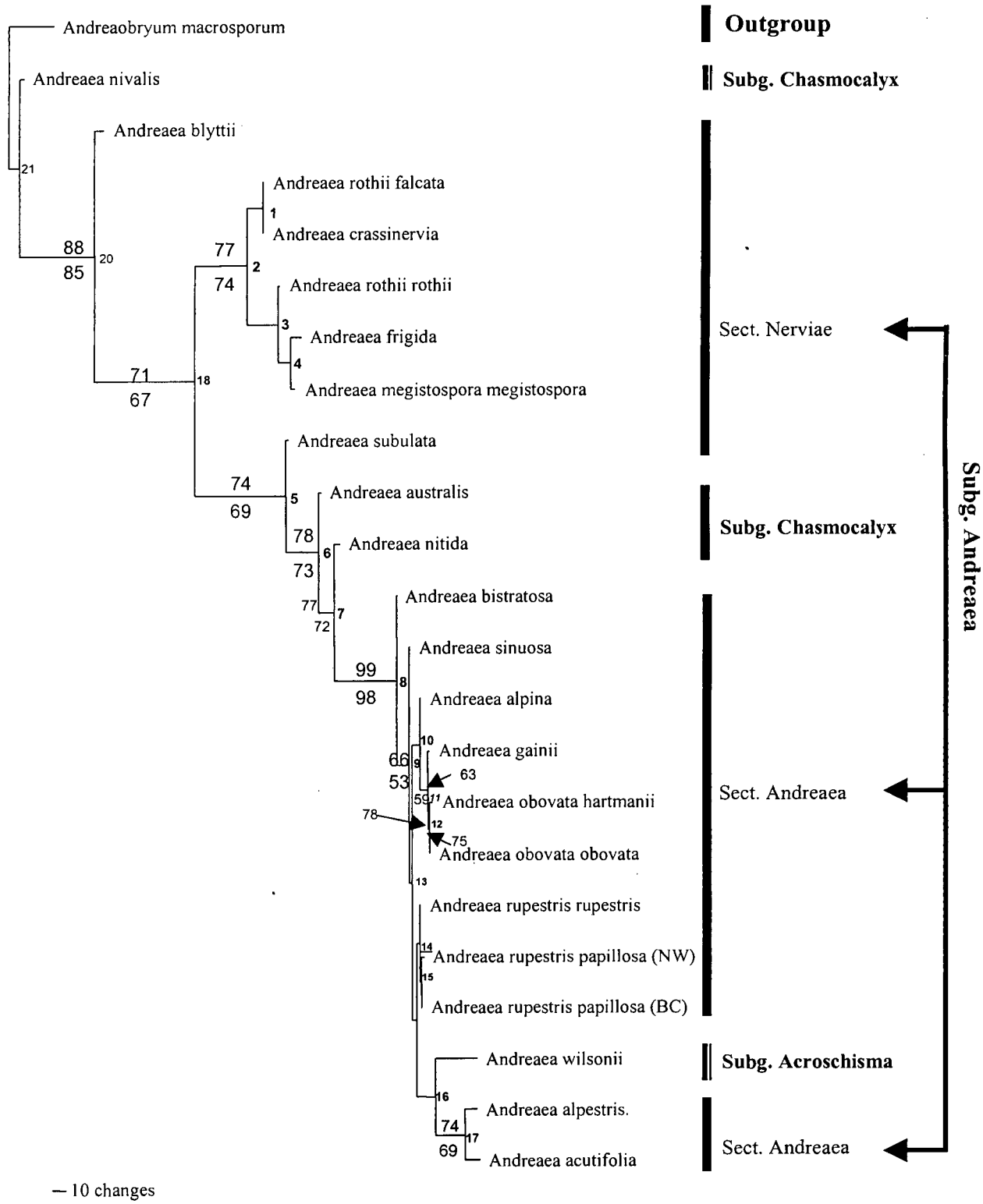


Figure 11: The single tree recovered from an analysis of successively weighted "reduced" *trnL/rps4*/morphology dataset (i.e. with only *Andreaeobryum* as the outgroup). Numbers above the branches are jackknife values, whereas those below are bootstrap values. Node numbers are indicated inside the nodes. Classification was derived from Murray (1988b), Braithwaite (1887) and Matteri and Farias (1999).

The analysis of the "reduced" *trnL–F/rps4*/morphology dataset indicated that some putative infra-generic groups are not monophyletic (Figures 11). The monophyly of subgenera *Chasmocalyx* and *Andreaea* was not upheld. Further, within the subgenus *Andreaea*, the presumed section *Nerviae* was also not supported. This section was instead resolved as polyphyletic. The section *Andreaea* of the same subgenus *Andreaea*, was resolved as paraphyletic by the *trnL–F/rps4/morphology* maximum parsimony analysis (nodes 8 of figure 11).

The resolution of the subgenus *Chasmocalyx*, (represented by *A. nivalis*) was interesting. In this analysis, other species of subgenus *Chasmocalyx* (which however contained only morphological characters) were included. *A. nivalis* was resolved as sister to the rest of the species of the genus and not in a clade with *A. blyttii* (as in the "reduced" *trnL–F/rps4* analysis), or with the these other putative members of subgenus *Chasmocalyx* (i.e. *A. australis* and *A. nitida*). These other species were instead resolved as a grade (nodes 6 and 7 of figure 11) that was coordinate to the clade of the nerveless species and *A. wilsonii*. The position of the monotypic subgenus *Acroschisma* (represented by *A. wilsonii*), like many other relationships within this ecostate species clade, was however not supported by bootstrap or jackknife values. Therefore, the exact position of *A. wilsonii* could not be confidently established from the results of the "reduced" combined *trnL/rps4*/morphology analyses alone. *A. subulata* was resolved in a grade with *A. nitida* and *A. australis*. Its well-supported resolution as basal to the section *Andreaea* is consistent with the results of the molecular analyses alone. This placement suggests that it may probably represent a separate lineage coordinate to a lineage including the ecostate species of the genus *Andreaea*.

It is worth noting here that the species with the molecular data more or less provided a framework on which the other species with only morphological data could be resolved, i.e. the positions of *A. bistratosa, A. nitida, A. australis, A. alpestris* and *A. gainii* were only based on the morphological evidence. It would therefore be interesting to see their placement in the presence of their *trnL–F* and *rps4* data.

CONCLUSIONS

The genus *Andreaea* is resolved as monophyletic by the analysis of the "complete" *trnL/rps4*/morphology dataset, with the position of *Andreaeobryum* ambiguously resolved, but outside the *Andreaea* clade. The "reduced" *trnL–F/rps4*/morphology analysis did not firmly resolve any of the current infra-generic groups as monophyletic. The subgenus *Chasmocalyx* was resolved as paraphyletic whereas the subgenus *Acroschisma* (*A. wilsonii*) could not firmly be established as separate from other lineages based on the current analyses alone. Though the exact position of *A. wilsonii* (subgenus *Acroschisma*) is not established, its closer relationships to the ecostate species of section *Andreaea*, than to other lineages within the genus, is well supported. Similarly, the monophyly of section *Andreaea* of subgenus *Andreaea* has neither been confirmed nor contradicted, whereas the section *Nerviae* (subgenus *Andreaea*) has been firmly resolved as polyphyletic. Other reliable outcomes of the "reduced" *trnL–F/rps4*/morphology analysis include the placement of *A. nivalis* and *A. blyttii* in a grade coordinate to the rest of the species; the well supported clade of some costate species (node 2 of figure 11), the clade containing *A. subulata*, *A. wilsonii* and species of section *Andreaea* (node 5 of figure 11) and the *A. acutifolia* / *A. alpestris* clade (node 17 of figure 11). All other relationships are unsupported.

CHAPTER 5

GENERAL DISCUSSION

This study presents the first cladistically derived hypotheses of relationships and infra-generic groups for the genus *Andreaea*. Firstly analysis of complete data for both the molecular and morphological datasets separately and in various combinations has provided an opportunity to determine the position of the genus *Andreaea* in relation to the major orders of the Bryales. Based on these analyses, the present study supports the monophyly of the genus *Andreaea* as presently circumscribed. *Andreaeobryum* the putative sister taxon to *Andreaea* has been strongly resolved by the combined molecular analysis outside the *Andreaea* clade and in a clade with the peristomate mosses. However, its exact position within this clade is not certain being weakly resolved as sister to the peristomate mosses clade. Combined molecular and morphological analysis in a previous study by Newton et al. (2000), only weakly resolved *Andreaeobryum* as sister to *Andreaea rothii*. The placement in their study was similar to the resolution by the separate morphological data analysis in the current study, although here *Andreaeobryum* is sister to a clade of both *Takakia* and *Andreaea*. However, Neither Newton *et al.*'s (2000) analysis nor the current analyses strongly support this placement. The strongly supported resolution of *Andreaea* and *Andreaeobryum* in a clade with the peristomate mosses by the "complete" *trnL/rps4* data analysis in the current study is therefore the favoured position.

Even with the addition of more species of *Andreaea* in this study, the combined analysis of molecular sequences and morphology together placed the genus *Takakia* as sister to *Sphagnum* and not to *Andreaea* or *Andreaeobryum* (figure 11). These results are also in agreement with the findings of recent higher-level molecular studies by Hedderson et al. (1998) and Newton et al. (2000). The relationship is interesting in view of the high morphological divergence between the *Sphagnum* and *Takakia*. In both this and Newton et al.'s (2000) study, the morphological analysis alone only weakly placed *Takakia* as more closely related to *Andreaea* and *Andreaeobryum* than to *Sphagnum*. Therefore, the view that *Takakia* is more closely related to

Andreaeobryum and *Andreaea* than any other mosses (Murray, 1988a, Smith and Davidson, 1993) has been strongly contradicted by the current and other studies (e.g. Hedderson et al., 1998; Newton et al., 2000). Accurate morphology-based analysis of relationships of *Takakia* to other taxa may have been hindered due to the evolutionary gap between the genus and other groups. In the current study, the difference in the position of *Takakia* in the morphological and molecular data may therefore be due to the difficulties in determining morphological character homology. This also resulted in a large number of inapplicable characters in the morphological data that may have further led to erroneous placement (Nixon and Davies 1991, Nixon et al., 1994) of *Takakia*. One other convincing outcome of the analysis of the complete data of combined morphological and molecular datasets is that the genus *Andreaea* is monophyletic and the other genera are indeed possible outgroups for polarising species relationships in this genus.

CONGRUENCE OF TREES FROM DIFFERENT ANALYSES, AND RELATIONSHIPS WITHIN ANDREAEA

1. Congruence between molecular analyses

Molecular analysis involved MP analysis as well as ML analysis of *trnL–F* sequences and a most parsimonious analysis of the combined *trnL–F/rps4* dataset. The "reduced" *trnL–F/rps4* analysis tree topology (figure 9) was mostly congruent with the tree obtained from the MP analysis of the *trnL–F* data alone (Figure 6). The only difference in resolution was that the clade consisting of *A. obovata* and *A. rupestris* var. *rupestris* (node 2 of figure 6) was collapsed in the combined analysis. A number of the clades were also better supported in the "reduced" *trnL–F/rps4* than the *trnL–F* analysis. The *trnL–F* maximum likelihood analysis tree though incongruent with other analyses in certain parts, was mostly unsupported.

2. Congruence between morphological, molecular and combined morphological analyses

Separate analysis of morphological data produced a topology that was incongruent in certain aspects with both the molecular analysis and the combined morphology/molecular data analyses. Most of the incongruence was between the alternative topologies for the positions of species of the section *Nerviae* (subgenus

Andreaea). There was incongruence in the placement of *A. subulata*, *A. nivalis*, *A. blyttii* and *A. wilsonii* (subgenus *Acroschisma*). The section *Nerviae* was resolved as monophyletic but with no support, by the analysis of morphological data alone, whereas it was strongly resolved as polyphyletic by the analyses of the *trnL–F*, the "reduced" *trnL–F*/rps4 and the "reduced" *trnL/rps4*/morphology datasets. The morphological analysis (figure 2) resolved *A. nivalis* strongly inside a clade with species of section *Nerviae*, and weakly as coordinate to this section. However, in both the *trnL–F* analysis (figure 6) and the combined analyses of the "reduced" *trnL–F/rps4*; (figure 5) and "reduced" *trnL–F/rps4*/morphology (figure 11) datasets, *A. nivalis* was resolved in a separate lineage coordinate to the rest of the species. The "reduced" *trnL–F/rps4* analysis resolved *A. nivalis* as sister to *A. blyttii*, with the *A. nivalis* / *A. blyttii* clade being coordinate to the rest of the species of *Andreaea*, whereas in the "reduced" *trnL–F/rps4*/morphology analysis, *A. nivalis* and *A. blyttii* were resolved as a grade with *A. blyttii* being coordinate to the rest of the species of *Andreaea*. *A. nivalis* was therefore resolved as coordinate to all the other species of *Andreaea*. However, the resolution in the "reduced" *trnL–F/rps4*/morphology analysis (figure 11) is the best-supported placement. This placement is exactly the same as that of the poorly supported tree (figure 7) recovered from the maximum likelihood analysis of the *trnL–F* dataset.

Other areas of incongruence between the morphology, the molecular and the combined analyses were the positions of *A. subulata* and *A. wilsonii*. *A. subulata* was resolved, without support, by the analysis of the morphological data as coordinate to a clade of the rest of the species of section *Nerviae*, whereas in the "reduced" *trnL–F/rps4* data and the *trnL–F* data maximum parsimony analyses (figures 5 and 6 respectively), it was strongly resolved in a clade with species of section *Andreaea* and weakly as coordinate to this section. In the same analyses *A. wilsonii* was resolved as coordinate to the clade of *A. subulata* and the section *Andreaea* species (node 8 of figures 5 and 6 respectively). In the "reduced" *trnL–F/rps4*/Morphology data analysis (figure 11) however, *A. subulata* was well-supported as coordinate to a clade including the section *Andreaea* species, *A. australis* (subgenus *Chasmocalyx*), *A. nitida* (subgenus *Chasmocalyx*) and *A. wilsonii* (subgenus *Acroschisma*). The exact positions of *A. subulata* and *A. wilsonii* were therefore uncertain due to this ambiguous resolution. However the two species were evidently more closely related

to the section *Andreaea* (ecostate) species than to the costate species of section *Nerviae*. It must be borne in mind that for the combined morphology and molecular analysis, molecular characters were not in such high numbers as to have their usual predominant influence on the topology. Much of the resolution and support is based on morphological characters; (some of the taxa did not have molecular evidence).

IMPLICATIONS FOR MORPHOLOGY AND CLASSIFICATION

Though the evidence is not overwhelming at this stage for new circumscriptions of the infra-generic groups, it is sufficient to suggest that most of the putative groups are not natural. This therefore also raises questions about the putative synapomorphies for these infra-generic groups and prompts a call for reassessment of their limits or even their taxonomic value. Some of the subgenera and sections may need to be circumscribed more broadly or more narrowly than previously envisaged. From this study, several lineages are apparent within the genus *Andreaea*.

Based on the available evidence and with reference to figure 11, subgenus *Andreaea* may include a morphologically heterogeneous clade containing all the species (node 18), except *A. nivalis* and *A. blyttii,* whereas *A. nivalis* and *A. blyttii* would each form separate monotypic subgenera. The *A. nivalis* lineage would be distinguished from other lineages by its non-differentiated, non-sheathing, perichaetial leaves, and by its auriculate leaf bases. *A. blyttii* differs from the other members of section *Nerviae* (where it is currently placed) by its dioicous condition and small sized spores. Except for *A. australis* and *A. nitida*, there is consistency within the possible large subgenus *Andreaea* in having differentiated sheathing and convolute perichaetial leaves. Within this possible large subgenus *Andreaea* group, two sections would be possible. The smaller section (i.e. node 2 of figure 11) would contain only, but not all costate species. These species include *A. rothii*, *A. crassinervia, A. frigida,* and *A. megistospora*. Possible exclusive morphological synapomorphies defining this group are not clear at this stage. The groups however share features such as presence of leaf costae, and falcata-secund, ovate lanceolate and subulate leaves. The suggestion by Murray (1988b) that *A. frigida* may be one of the primitive *Andreaea* species closely related to *A. nivalis* is also therefore not supported by this study.

The larger section of subgenus *Andreaea* would comprise *A. subulata*, *A. nitida A. australis* and all the ecostate species. Apparently this group consists of several distinct lines and therefore further sub-sections would be possible. *A. subulata* is one exception to the general morphological trends of the clade. It possesses a number of features that are characteristic of the species of the smaller of the two sections e.g. falcate leaves with reduced and elongated apical parts.

The subgenus *Chasmocalyx*, as presently circumscribed, is polyphyletic. However it appears that it should be monotypic including only *A. nivalis*. The other species of this subgenus (*A. nitida* and *A. australis*) would probably be placed in the possible large section of subgenus *Andreaea*. However, as suggested below, they would probably form separate lineages (subsections) coordinate to the ecostate species (see figure 11). The *A. australis* lineage would be distinguished by its large oblong leaves and a decurrent costa, whereas the *A. nitida* lineage would be distinguished by cymbiform leaves with umbonate apices, and a broad costa that fans out apically.

A. wilsonii (Subgenus *Acroschisma*) was consistently resolved by the *trnL–F* MP analysis and all combined analyses in a clade including *A. subulata* and all the species of section *Andreaea*. However its position within the clade was ambiguously resolved. Therefore, whether *A. wilsonii*, sometimes afforded generic status (e.g. Reimers, 1954), forms a separate lineage that is coordinate to other species of the section *Andreaea* or is simply embedded among the other species is indeterminable at this stage. However, the fact that it is strongly supported by the "reduced" *trnL–F/rps4*/morphology analysis (figure 11) as part of the clade containing ecostate (node 8 of figure 11) suggests that *A. wilsonii* definitely has closer relationships to the ecostate species than to the costate species. The ambiguous resolution of *A. subulata* and *A. wilsonii*, when their placement by the "reduced" *trnL–F/rps4* data and the "reduced" *trnL–F/rps4*/morphology data are considered, further justifies the sinking of both species into one group (section *Andreaea* of subgenus *Andreaea*). The circumscription of possible lower level groups within the presumed section *Andreaea* (see figure 11) therefore also remains uncertain.

However, if *A. wilsonii* is resolved as coordinate to the ecostate species clade by further evidence, then the putative section *Andreaea* may form a monophyletic group

within the possible large section *Andreaea*. It is expected that the majority of the section *Andreaea* species that are not included in the analysis would resolve as members of this clade. Then the group with only ecostate species possessing differentiated perichaetial leaves, may include two sections: the large section with all ecostate species that possess capsules with four valves splitting beyond the upper third of the capsules and a smaller monotypic section comprising *A. wilsonii*, which possesses capsule with up to 8 valves confined only to the upper third of the capsule. However, if *A. wilsonii* is not coordinate to all the other species within the group, the current section *Andreaea* may need to be more broadly circumscribed. The final decision awaits further studies.

According to Vitt (1984), numbers and amount of separation of species contained in individual genera may indicate the age of the genera. Great numbers of structurally related species in some groups may suggest recent active evolution (Vitt, 1984). This may be the observed phenomenon in the group of section *Andreaea* (ecostate species) with some closely related species that are difficult to delimit e.g. the rupestris group (*A. rupestris A. obovata* and *A. alpestris*). Thus it is possible that these species are difficult to delimit because they belong to a fairly novel and actively-evolving lineage (Vitt, 1984). Relatively few and isolated species may suggest old stable genera that underwent active evolution some time back in the geological past (Vitt, 1984). Such a situation also seems to be found in the case of the *A. nivalis* lineage and possible separate lineages for *A. blyttii*, *A. australis*, and *A. nitida*. Due to low support and ambiguous resolution of some groups, it is better to refrain at the moment from making formal conclusions about infra-generic groups until more data and more species are included.

CHARACTER EVOLUTION.

Combined analysis of successively weighted "reduced" *trnL–F*/*rps4*/morphology data has provided an estimate of phylogenetic relationships within *Andreaea* that is favoured because it is well resolved and contains higher number of supported nodes than any other analysis of partitioned or combined data. This tree also expresses hypotheses of relationships based on all the available evidence for the species of *Andreaea*. Therefore this topology was the one on which character state optimisations

and their subsequent discussions were based. The character state optimisation onto the "reduced" *trnL–F*/*rps4*/morphology tree was used as the means of assessing homology statements and suggesting hypotheses for morphological character evolution. A few hypotheses of morphological character evolution are presented below.

Sexuality

According to Schofield and Hebant (1984), and Watson (1971), more than 50% of mosses are dioecious. However only 33% of the species of *Andreaea* studied were dioecious. Dioecy was optimised as the pleisiotypic condition (figure 12) and was shown to have given rise to the monoecious condition three times within *Andreaea*, once in *A. nivalis* and twice in the broadly circumscribed subgenus *Andreaea*. Such a situation is common among the bryophytes (Watson 1971) and according to Miller (1979), the monoicous condition is advanced over the dioicous condition. However, the apparent reversal to dioecy in *A. alpina* may be explained by a number of possibilities. It may be due to suppression of one sex on plants resulting in different sexes predominating on separate individual plants, or it may simply be due to differentiation of the time of production of the different sexual organs resulting in the presence of only one sexual organ on any plant examined. Another possibility may simply be an erroneous placement within the clade including monoecious species: i.e. considering the tree on which this optimisation was based (figure 11), there was no support for the position of *A. alpina* within this clade (see node 16). If *A. alpina* were placed as coordinate to the rest of the species in this clade there would be no apparent reversal event.

The evolution of monoecy from dioecy has genetic implications. The gametophyte is usually haploid as it arises from haploid spores or from broken fragments of haploid gametophores. However, in certain cases, fragments broken off from sporophytes (diploid) may give rise to protonemata, which ultimately produces diploid (possible progenitors of both sexes), rather than the usual haploid gametophores (thus with progenitors of only one type of sexual organs). Since these diploid gametophores

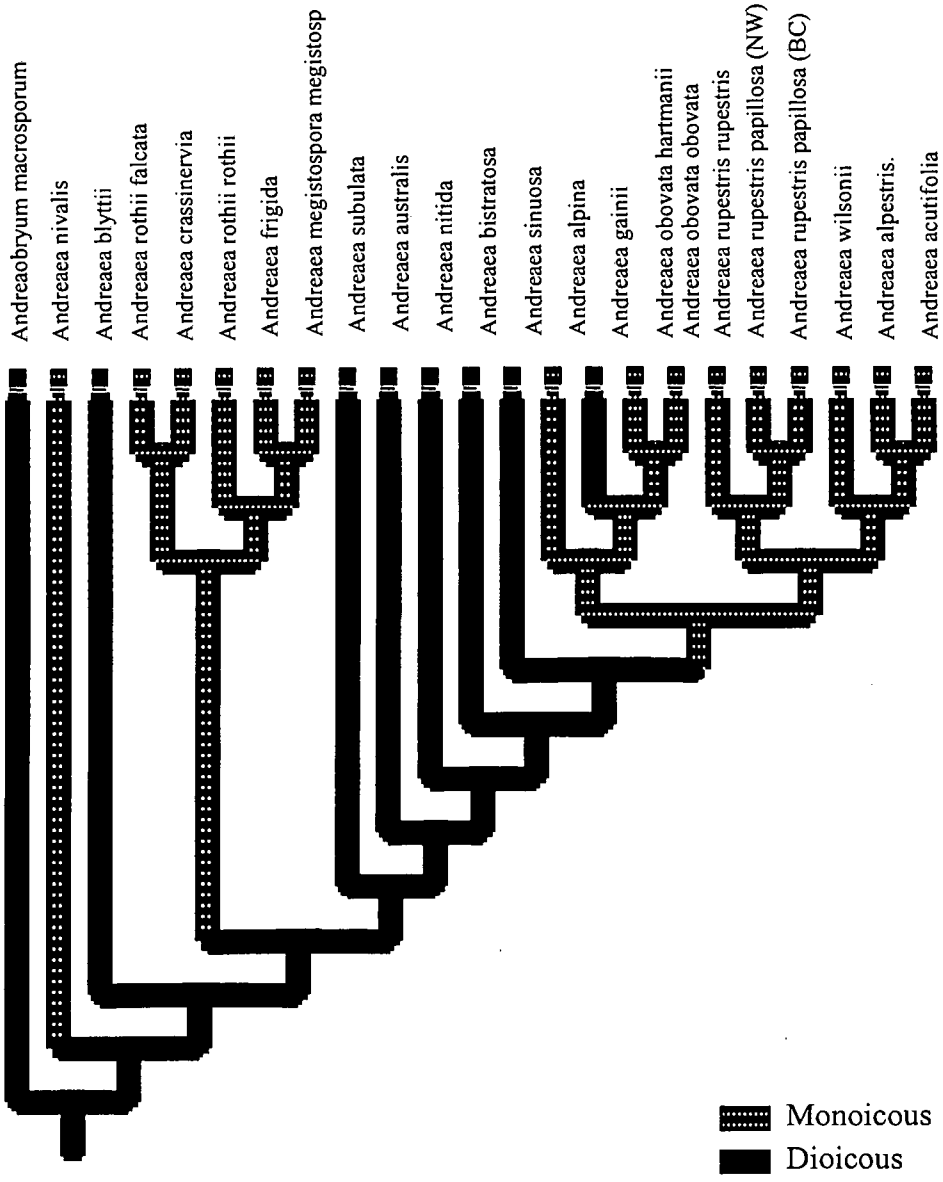


Figure12: Optimization (DELTRAN) of the character, sexuality.

possess both male and female chromosomes they can produce both male and female sex organs on the same plant thus giving rise to the monoecious condition. It is thus easy for the dioecious condition to change to a monoecious condition and hence the multiple evolution of monoecy in many moss genera (Parihar, 1965).

Falcate leaves

Presence of falcate leaves was optimised as the pleisiotypic condition (figure 13). Non-falcate leaves seem to be associated with leaf widening as most of the non-falcate leaves have more apically widened leaves than the falcate ones. The non-falcate leaves are the derived condition that evolved only once in *Andreaea*. Non falacata leaves are a good synapomorphy for the broadly circumscribed subgenus *Andreaea* but excluding *A. subulata* and *A. australis*.

Costa presence

The putative sections *Nerviae* and *Andreaea* are based on the presence or absence of a leaf costa. Presence of a leaf costa was optimised as the pleisiotypic condition (figure 14). A leaf costa has been lost only once and this is a good synapomorphy for the traditional section *Andreaea*, but including *A. wilsonii*. According to Miller (1979), a strong costa is similarly more primitive than a weak one with the ecostate conditions being the most derived state.

Perichaetial leaves

The subgenus *Chasmocalyx* is distinguished from the other two presumed subgenera *Acroschisma* and *Andreaea* by its possession of undifferentiated perichaetial leaves. Differentiated perichaetial leaves are the general condition that was optimised as derived from an undifferentiated condition (figure 15). However there has been at least one reversal from differentiated to undifferentiated perichaetial leaves. However, whether there were 1 or 2 separate reversals to non-convolute perichaetial leaves in *A. nitida* and *A. australis* is equivocally optimised. Though the differentiated perichaetial leaves are of widespread occurrence in the broadly circumscribed subgenus *Andreaea* it is thus contradicted by the reversal in *A. nitida* and *A. australis*.

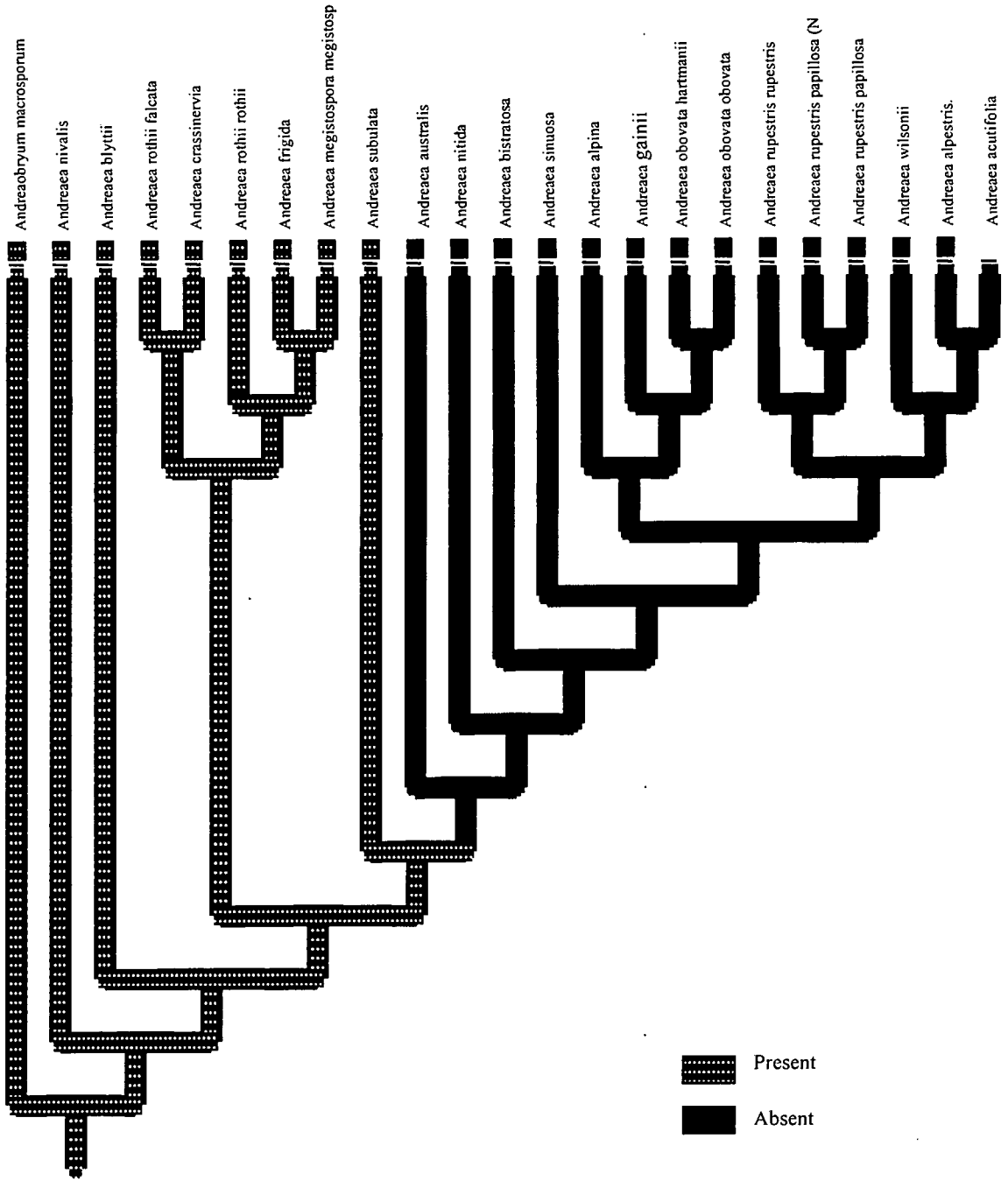


Figure 13: Optimization (DELTRAN) of the character, falcate leaves

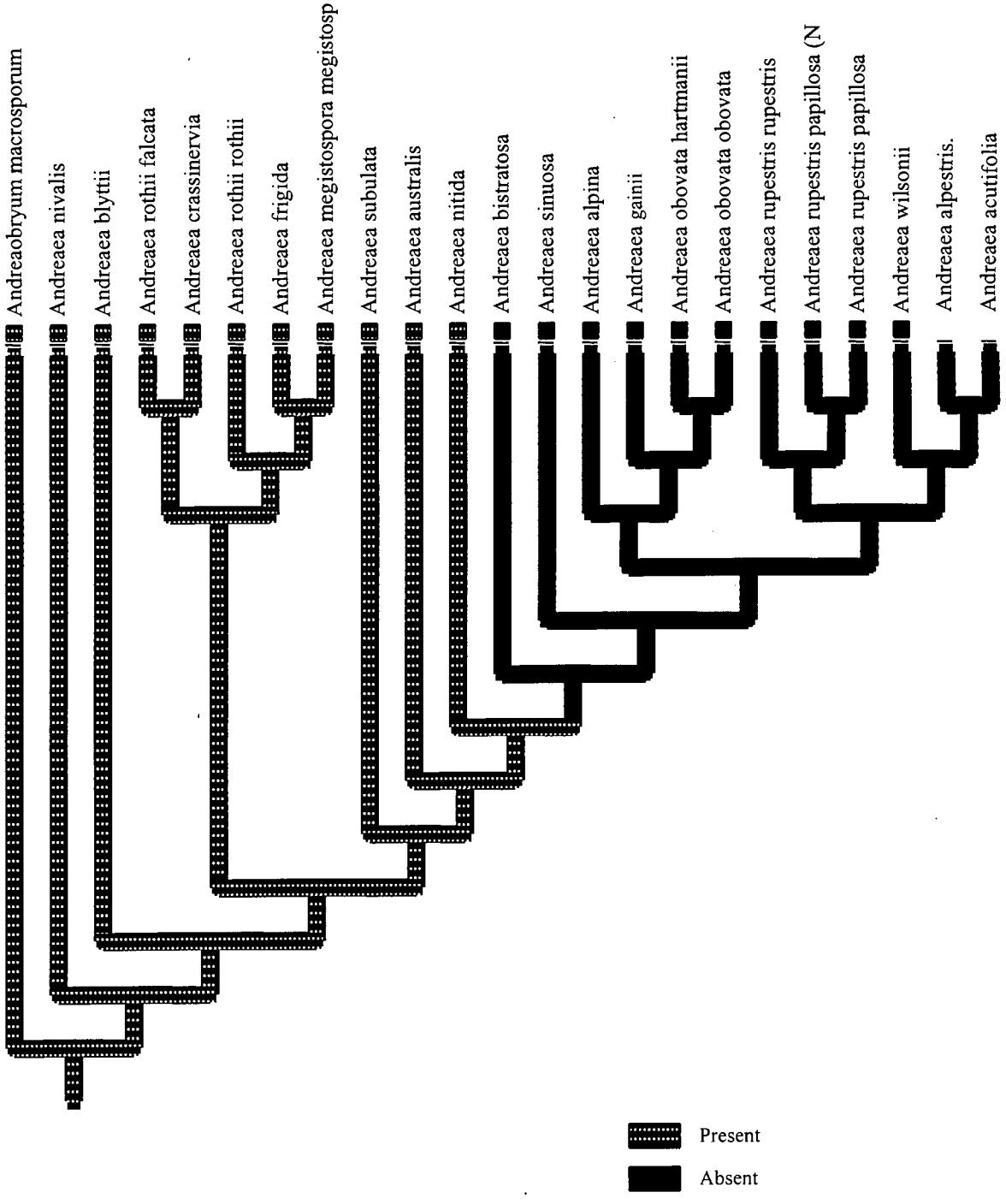


Figure 14: Optimization (DELTRAN) of the character, Leaf Costa

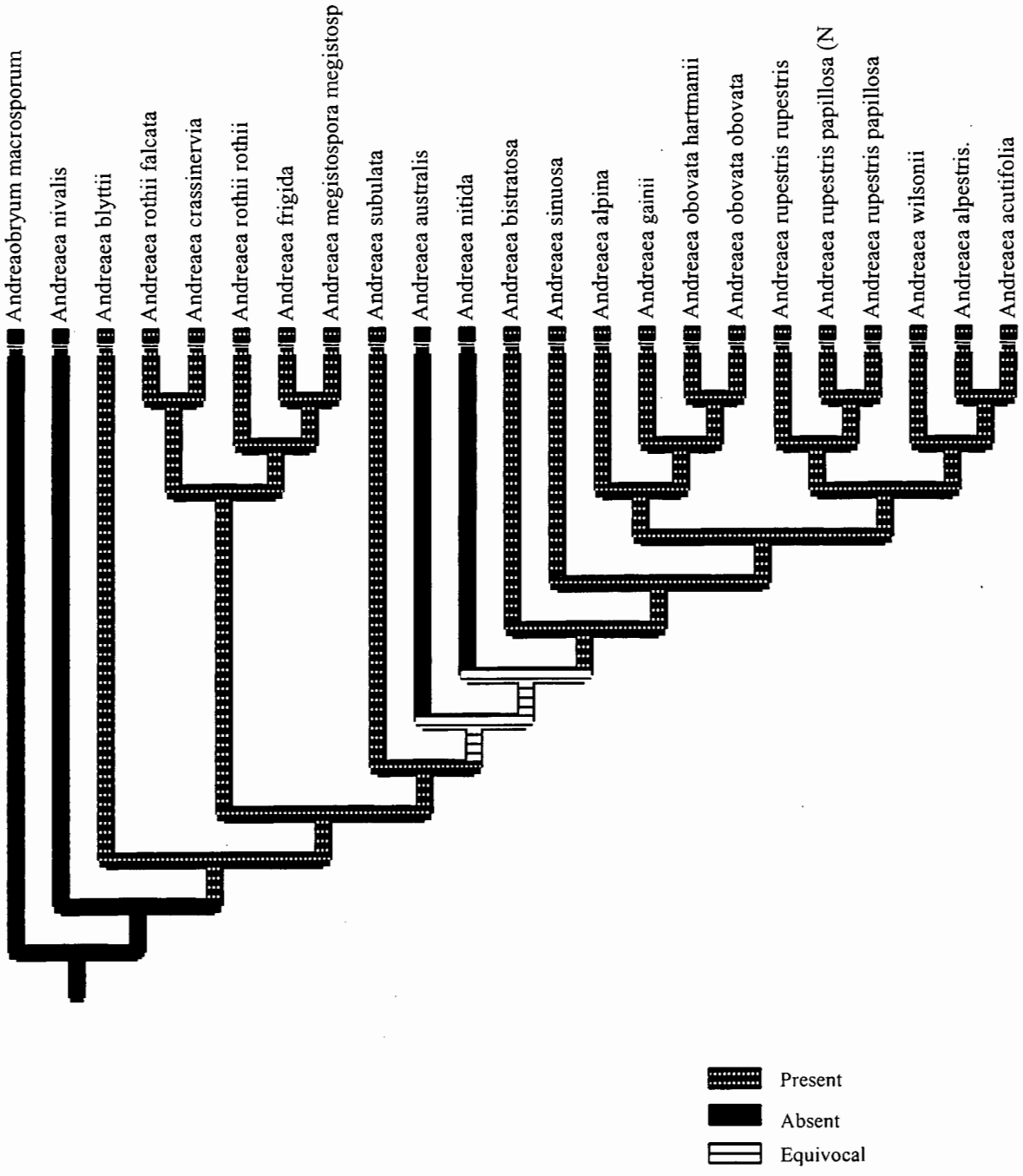


Figure 15: Optimization (DELTRAN) of the character,differentiation of perichaetial leaves

Number of capsule valves

In most cases the capsules of *Andreaea* open by 4 slits, extending neither to the apex nor the base of the capsule. In *A. nivalis*, however, there are up to 6 slits. In *A. wilsonii*, on the other hand, the capsule opens by up to 8 slits that reach the apex and extend only a short distance downwards, so that the whole tip of the capsule splits into up to 8 valves. Whether a four-valved or a six-valved capsule was the ancestral condition was equivocally optimised (figure 16). However, based on the general occurrence of the four-valved capsule, it is likely that this was the ancestral condition at least for the broadly circumscribed subgenus *Andreaea*. A capsule with up to eight valves is an autapomorphic state suggesting specialisation for spore dispersal.

Spore size

Spore size shows complex and homoplasious patterns of evolution within *Andreaea* (figure 17). Intermediate sized spore seems to be the pleisiotypic and general condition, whereas enlarged spores as in *A. rothii* and *A. megistospora* or reduced spores as in *A. blyttii* or *A. sinuosa*, suggest specialisation (Schofield, 1985). The occurrence of the same size range of spores in divergent lineages can be explained in terms of factors effecting convergent evolution. Though spore size may prove to be a valuable taxonomic character for delimiting species, it is of lesser significance for circumscribing infra-generic groups in *Andreaea*. Perhaps for many species, spore characters have been a means of adapting to their particular environmental conditions for survival optimisation.

Mucilage hairs on pseudopodium

Occurrence of structures such as mucilage hairs or even bracteoles on the pseudopodium has been considered a characteristic of the putative subgenus *Chasmocalyx* (e.g. Murray, 1988b). Presence of mucilage hairs shows an equivocal optimisation basally (figure 18), being present in *A. nivalis*. However mucilage hairs on pseudopodia arose again within the broadly circumscribed subgenus *Andreaea* in *A. nitida*.

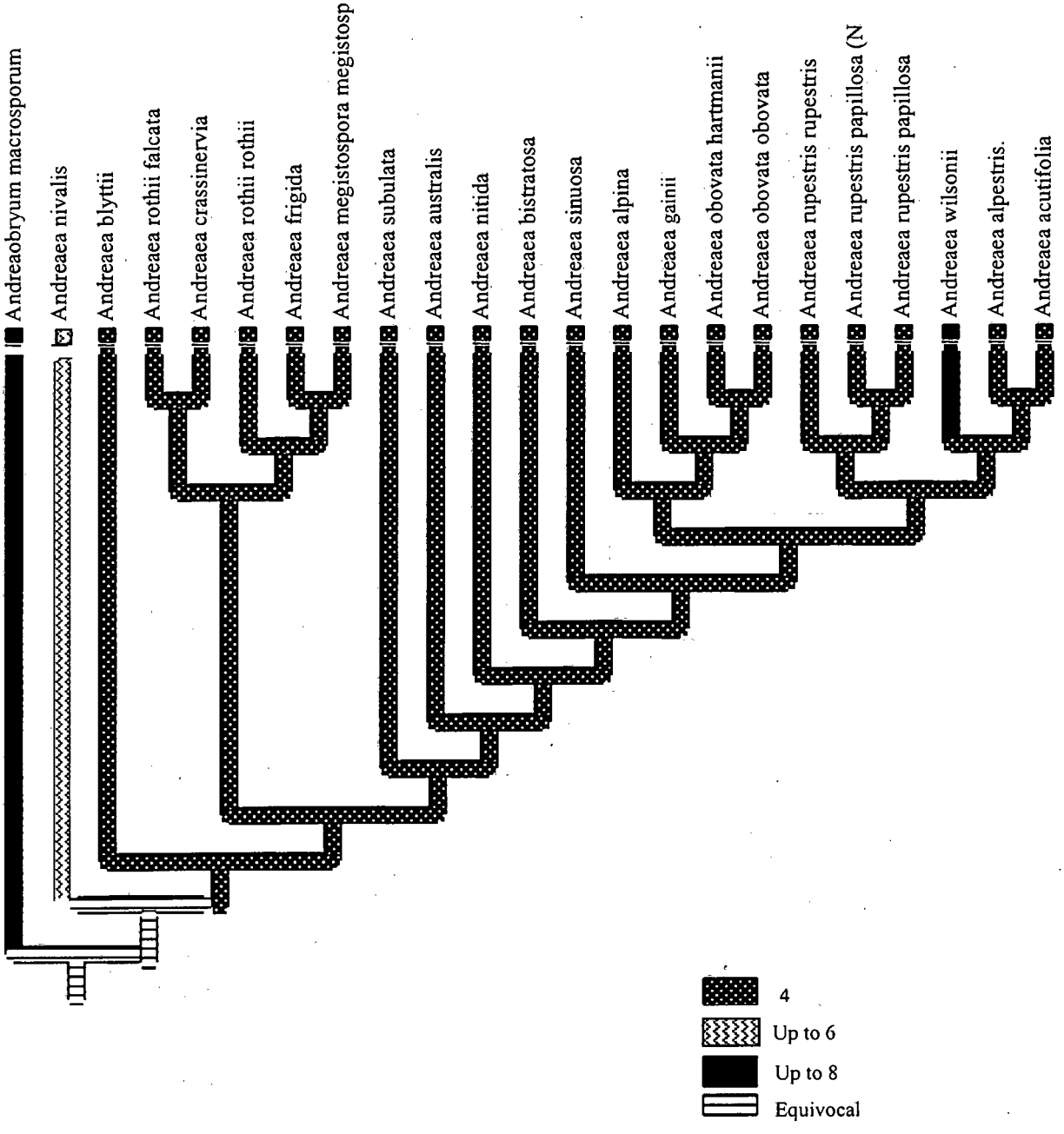


Figure 16: Optimization (DELTRAN) of the character, Number of capsule valves

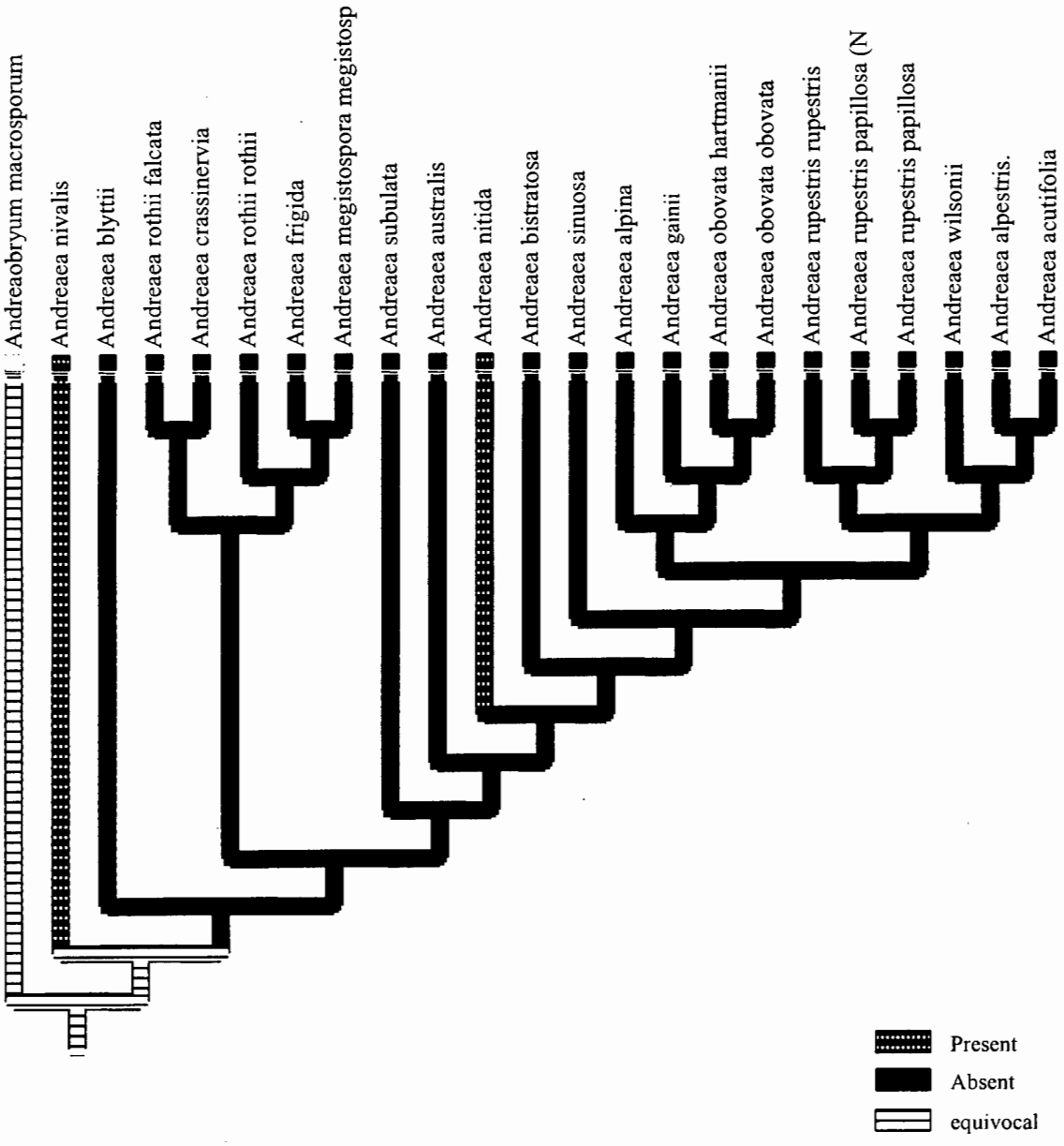


Figure 18: Optimization (DELTRAN) of the character, Mucilage hairs on Pseudopodium.

CONCLUSIONS

A complete phylogenetic treatment of the genus is beyond the scope of this thesis. The current study has shown that the genus *Andreaea* is monophyletic and forms a lineage distinct from other moss lineages. A number of parts of the inferred phylogeny contain components whose relationships are still not clearly understood. However the work done here, though essentially preliminary, and despite the limited number of species used, has been sufficient in revealing trends for sub-generic relationships within the genus. Some systematic conclusions that can be made and possible nomenclatural groups that can be suggested are as follows: Most of the putative infra-generic groups of *Andreaea*, as presently circumscribed are not supported by this study and can therefore not be upheld. Three separate subgenera are possible. Two of these are monotypic, containing *A. nivalis* and *A. blyttii* respectively, and a large subgenus containing all the remaining species. Taxonomic implications are that the subgenus *Andreaea* may have to be more broadly circumscribed, whereas the section *Nerviae* may have to be narrowly circumscribed. Within the large, broadly circumscribed subgenus *Andreaea*, there may be two sections with the larger of the two including all ecostate species and *A. subulata*, *A. nitida* and *A. australis*. Several lines are possible within this group. This will become clearer when more data are available. The ecostate putative subgenus *Acroschisma* is embedded within the large subgenus *Andreaea*. Implications for morphology are that new infra-generic groups should therefore be established based on characters states other than those previously employed. The determination of these characters requires further careful study. Certain character evolution trends within the genus have also been highlighted. Costate leaves are generally the pleisiotypic condition. Though many costate species are related to each other than to ecostate species a few are more related to ecostate species than costate ones. Other pleisiotypic conditions include; dioecy, falcate leaves and medium sized spores. The dioecious condition in *Andreaea* has given rise to the monoecious condition multiple times.

FUTURE DIRECTIONS

The results of this thesis have also given useful pointers for further inquiry. Resolution of the relationships that are currently ambiguous may require the inclusion of more sequences from the non-coding region at the 3'-end of the *rps4* region. Further, the assessment of relationships and character evolution tendencies would be improved by addition of datasets with both faster evolving rates (such as the Internal transcribed spacer (ITS) sequences) and slower evolving rates (such as 18S) than the sequences utilized here. Such molecular studies promise to play a vital role in providing a more complete understanding of species relationships for the genus *Andreaea*, especially in giving better-supported resolution of terminal taxa. Preliminary attempts to obtain ITS sequences during the course of the current study indicated that cloning work might be necessary in order to obtain the targeted sequences and such work would need more time than available for the current study. A morphological character reassessment, perhaps with increased emphasis on characters such as ultra-structural (e.g. detailed spore morphology) and ontogenetic characters would also be useful in shedding more light on, and strengthening the understanding of, relationships within the genus and refining our interpretations of the sub-generic groups and evolutionary polarities of the characters. Once a robust phylogeny is established it would be worth investigating biogeographic relationships for the group. Certain relationships resulting in some characters showing homoplasious patterns would call for further search along the lines of ecological and other factors leading to phenomena such as convergent or parallel evolution.

LIMITATIONS OF THE STUDY

It was difficult to obtain many of the species of *Andreaea* due to the short period in which the study had to be conducted. Many of the species of *Andreaea* did not arrive in time and were therefore left out of the study. These will be included in a later, much broader and more detailed phylogenetic study. This therefore is only a preliminary study, which indeed, still adequately highlights the phylogenetic, and character evolution trends and taxonomic implications for infra-generic groups in the genus *Andreaea*. The highly polymorphic nature of many characters is the other factor that provided credible challenge in generating the morphological data as well as

in the confidence that might be vested in this data. However, molecular data, though with fewer species of *Andreaea* and some incomplete sequences for the *rps4* gene, provided independent evidence that shaded more light on the reliability of the phylogenies found from analyses of morphological data. Obtaining sequences presented a number of problems. Attempts to obtain ITS sequences were abandoned due to the preponderance of multiple bands in the PCR products run on gels. These bands apparently represented different versions of the targeted sequences. There was therefore need for cloning work in order to obtain the same set of target sequences for all the species. However, because of the intensive laboratory work required and the limited time available, this work could not immediately be conducted. Hence the use of only the chloroplast sequences from the *trnL–F* and *rps4* gene regions. As earlier alluded to, it was however not easy to obtain usable *rps4* sequences in the reverse direction. Most of the reverse direction sequences were not satisfactory and sequences for the whole non-coding region of the *rps4* were discarded resulting in exclusion of a lot of potentially informative characters. An appreciable amount of useful literature was in languages unknown to the author such as German and Russian and translation of large amounts of this literature was not possible within the limited time available. However a few of the apparently useful parts of this literature were translated for use, by reference to Dictionaries and the use of Alta Vista's Babel Fish translations (http://babelfish.altavista.com/translate.dyn).

REFERENCES

Anderberg, A. and A. Tehler. 1990. Consensus trees a necessity in taxonomic practice, *Cladistics*, 6: 399–402.

Bakker, F. T., A.Culham, C. E. Pankhurst, M. Gibby. 2000. Mitochondrial and chloroplast DNA-based phylogeny of Pelargonium (Geraniaceae). *American Journal of Botany* 87(5): 727-734.

Barrett, M., M. J. Donoghue, and E. Sober. 1991. Against consensus. *Systematic Zoology*, 40: 486-493.

Braithwaite, R. 1887. *The British Moss Flora. Vol. 1 Acrocarpi I. Andreaeacea, Buxbaumiaceae, Georgiaceae, Polytrichchaceae, Fissidentaceae, Leucoberaceae, Dicranaceae, Tortulaceae, Weberaceae.* pp. 4-18. L. Reeve & Company, London.

Brouat, C., L. Gielly and D. McKey. 2001. Phylogenetic relationships in the genus *Leonardoxa* (Leguminosae: Caesalpinioideae) inferred from chloroplast *trnL* intron and *trnL-trnF* intergenic spacer sequences. *American Journal of Botany* 88(1): 143-149.

Buck, W. R., B. Goffinet and A. J. Shaw. 2000. Testing morphological concepts of orders of pleurocarpous mosses (Bryophyta) using phylogenetic reconstructions based on TRNL-TRNF and RPS4 sequences. *Molecular Phylogenetic Evolution* 16(2): 180-98.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford and P. J. Waddell, 1993. Partitioning and combining of data in phylogenetic analysis. *Systematic Biology* 42: 384-397.

Cao, T. and C. Gao. 1995. A revised taxonomic account of the genus *Andreaea* (Andreaeaceae, Musci) in China, Contributions to the Bryoflora of China 6. *Harvard Papers in Botany*, 7: 11—24.

Carpenter, J. M. 1988. Choosing among equally parsimonious cladograms. *Cladistics* 4: 291-296.

Cavers, F. 1911. *The Inter-relationships of the Bryophyta*. New Phytologist Reprint No. 4. pp. 173-176. The Botany School, Cambridge.

Chiang, T. Y. 1998. Taxonomic revision of *Andreaea* (Mosses, Andreaeaceae) of Taiwan. *Botanical Bulletin of Academia Sinica* 39: 57-68.

Cox, J. C. and T. A. J. Hedderson. 1999. Phylogenetic relationships among ciliate arthrodontous mosses evidence from chloroplast and nuclear DNA sequences. *Plant Systematics and Evolution* 215: 119-139.

Crandall-Stotler, B. 1984. Musci Hepaticae and Anthocerotes— an essay on analogies, pp. 1093-1129, *In* R. M. Schuster. *New Manual of Bryology* Vol.2 The Hattori Botanical Laboratory. Nichinan, Miyazaki.

Davis, J. I. 1993. Character removal as a means of assessing stability of clades. *Cladistics* 9: 201-210.

de Queiroz, A. 1993. For consensus (sometimes). *Systematic Biology* 42: 368-372.

Dillenius, J. J. 1741. *Historia Muscorum* in qua circiter sexcentae species veteres et novae ad sua genera relatae describuntur, et iconibus genuinis illustratur: cum appendice et indece synonimorum. pp. 482-532. Oxford.

Dixon, H. N. 1929. *Studies in the Bryology of New Zealand with special reference to the Herbarium of Robert Brown*. Part VI. pp. 346-355. Ferguson and Osborn Ltd. Lambton, Quay.

Dixon, H. N. 1932. Classification of Mosses. *In* F. Verdoorn (ed.) *Manual of Bryology*. pp. 397-412. Martius Nijhoff, The Hague.

DNASTAR. 1994. *Lasergene, biocomputing software for the Macintosh,* Madison, Wisconsin.

Donoghue, M. J. and M. J. Sanderson. 1992. The suitability of Molecular and Morphological Evidence in Reconstructing Plant Phylogeny, *In* Soltis, P. S., Soltis, D.

E. and J. J. Doyle (eds.): *Molecular Systematics of Plants*. pp. 340-368. Chapman & Hall, London.

Ehrhart. 1778. *Andreaea*, eine neue Pflanzengttung *Hannover Magaz*. 16: 1601-1604.

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

Farris, J. S. 1970. Methods for computing Wagner Trees. *Systematic Zoology* 19: 83-92.

Farris, J. S., 1989, The retention index and the rescaled consistency index, *Cladistics*, **5**: 417-419.

Felsenstein, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Systematic Zoology* 27: 401-410.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368-376.
Felsenstein, J. 1983. Inferring evolutionary trees from DNA sequences. *In* B. S. Weir. (ed.) Statistical Analysis of DNA Sequence Data. Marcel Dekker. New York:

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

Fitch, W. M. 1971. Towards defining the course of evolution, Minimal change for a specific tree topology. *Systematic Zoology* 20: 406-416.

Garbary, D. J. and K. S. Renzaglia. 1998. Byophyte phylogeny and the evolution of land plants: Evidence from development and ultrastructure, pp 45 – 63. *In* J. W. Bates N. W. Ashton and J. G. Duckette (eds.), *Byology for the Twentieth-Century*. Maney Publishing and British Bryological Society, Leeds.

Garbary, D. J., K. S. Renzaglia and J. G. Duckett. 1993. The phylogeny of land plants: a cladistic analysis based on male gametogenesis. *Plant Systematics and evolution* 188: 237-269.

Gawel, N. J. and R. L. Jarret. 1991. A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Plant Molecular Biology Reports* 9: 262-266.

Hartigan, J. A. 1973. Minimum mutation fits to a given tree. *Biometrics* 29: 53-65.

Hattori, S. and M. Mizutani. 1958. What is *Takakia lepidozioides*? *Journal of the Hattori Botanical Laboratory*. 20: 295-303.

Hedderson T. A., R. Chapman and C. J. Cox. 1998. Bryophytes and the origins and diversification of land plants: new evidence from molecules. pp. 65 – 77. *In* J. W. Bates N. W. Ashton and J. G. Duckett (eds.), *Bryology for the Twentieth-Century*. Maney Publishing and British Bryological Society, Leeds.

Hedderson, H. A., R. L. Chapman, W. L. Rootes. 1996. Phylogenetic relationships of bryophytes inferred from nuclear encoded rRNA gene sequences. *Plant Systematics and Evolution* 200: 213-224.

Hedwig, J. 1801. Species Muscorum frondosorum desciptae et tabulis aeneis LXXVII Coloratis illustratae, lepzig, Paris. pp. 47-49.

Hendy, M. D., D. Penny. 1989. A framework for the quantitative study of evolutionary trees. *Systematic Zoology* 38: 297-309

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

Hillis, D. M. and J. P. Huelsenbeck. 1992. Signal, noise, and reliability in molecular phylogenetic analysis. *Journal of Heredity* 83: 189-195.

Huelsenbeck, J. P., 1991, Tree length distribution skewness: an indicator of phylogenetic information. *Systematic Zoology* 40(3): 257-270.

Janzen, P. 1917. Die Haube der Laubmoose. *Hedwigia* 58: 156-280.

Jefferies, R. P. S. 1979. The origin of chordates – a methodological essay. pp 443 – 477. *In* R. M. House (ed.) "The Origin of major invertebrate groups". Academic Press. London, New York.

Kitching, J. I., P. L. Forey. C. J. Humphries. D. M. Williams. 1998. Cladistics; Theory and Practice of parsimony analysis, second edition. Oxford, New York, Tokyo, *The Systematics Association Publication*, No. 11.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology*. 38: 7-25.

Kohjyouma M., I. J. Lee, O. Iida. K. Kurihara. K. Yamada. Y. Makino. S. Sekita. M. Satake. 2000. Intraspecific variation in *Cannabis sativa* L. based on intergenic spacer region of chloroplast DNA. *Biological & Pharmaceutical Bulletin.* 23(6):727-30.

La Farge-England, C. 1996. Growth form branch pattern, and perichaetial position in mosses: Cladocarpy and pleurocarpy redefined. *The Bryologist* 99(2): 170-186.

Lewis, P. O., 1998, Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data, pp132–163, *In* D. E., Soltis, P. S., Soltis and J. J. Doyle, *Molecular Systematics of Plants II: DNA sequencing*, Kluwer Academic Publishers, Massachusetts, New York.

Linnaeus C. 1753. Species Plantarum, Tomus pp. 1131 – 1136. Holmiae, Impensis. Salvii.

Maddison, W. P., and D. R., Madisson, 1992. MacClade: analysis of phylogeny and character evolution, Version 3.0, Sinauer Associates, Sunderland, Massachusetts.

Magill, R. E. 1981, Andreaeaceae. *In* O. A. Leistner. Bryophyta, Part 1 Mosses Fascicle 1 Sphagnaceae — Grimmiaceae. *Flora of Southern Africa.* Botanical Reasearch Institute, Pretoria.

Marshall, C. R. 1992. Character analysis and integration of molecular and morphological data in an understanding of sand dollar phylogeny. *Molecular Biology and Evolution* 9: 309-322.

Matteri, C. M. and Farias, R. M. 1999. Leptotypification of *Andreaea pachyphylla* (musci, Andreaeopsida). *Lindbergia* 24: 11-14.

Mickevich, M. F. 1978. Taxonomic congruence. *Systematic Zoology* 25: 143-158.

Miller, H. A. 1979. The Phylogeny and Distribution of Musci. pp. 11-39. *In* G. C. S. Clarke and J. G. Duckett. "*Bryophyte Systematics*". Systematics Association Special volume No. 14. Academic Press London, New York.

Mishler, B. D., P. H. Thrall, J. S. Hopple Jr., E. De Luna and V. Rytas. 1992. A molecular approach to the phylogeny of bryophytes: Cladistic analysis of chloroplast encoded 16S and 32S Ribosomal RNA genes. *Bryologist* 95: 172 –180.

Missouri Botanical Garden, 05 Jul 2000. *TROPICOS Nomenclatural Data Base http://mobot.mobot.org/Pick/Search/pick.html.*

Miyamoto, M. M. 1985. Consensus cladograms and general classifications. *Cladistics* 1: 186-189.

Murray, B. M. 1987. *Andreaea schofieldiana* and *A. megistospora* species novae and taxonomic criteria for sect. *Nerviae* (Andreaeopsida). *Bryologist* 90(1): 15-26.

Murray, B. M. 1988a. Systematics of the Andreaeopsida (Bryophyta): Two orders with links to *Takakia. Beiheft zur Nova Hedwigia* 90**:** 289-336.

Murray, B. M. 1988b. The genus *Andreaea* in Britain and Ireland. *Journal of Bryology*. 15: 17—82.

Nadot, S., R. Bajon and B. Lejeune, 1994, The Chloroplast gene *rps4* as a tool for the study of Poaceae phylogeny, *Plant Systematics and Evolution.* 191:27-38.

Newton, A. E., C.Cox, J. G. Duckett, J. A. Wheeler, B. Goffinet, T. A. J. Hedderson, and B. D. Mishler. 2000. Evolution of the major moss lineages: phylogenetic analyses based on multiple gene sequences and morphology, The *Bryologist* 103 (3): 187-211.

Nixon, K. C. and J. I. Davies. 1991. Polymorphic taxa, missing values and cladistic analysis, *Cladistic* **7:** 233-241.

Nixon, K. C., and J. M. Carpenter. 1993. On outgroups. *Cladistics* 9**:** 413- 426.

Nixon, K. C., W. L., Crepet, D. Stevenson, E. M., Friis. 1994. A re-evaluation of seed plant phylogeny. *Annals of Missouri Botanical Garden* 81**:** 484-533.

Page R. D. M. 2000, Nexus data editor, *http//taxonomy.zoology.gla.ac.uk/rod.html.*

Parihar, N. S. 1965. *Introduction to Embryophyta.* Volume 1: Bryophyta. Central Book Depot. Allahabad.

Pimetel, R. A. and R Riggins. 1987. The nature of cladistic data. *Cladistics* 3: 201-209.

Platnick, N. I., C. E. Griswold and J. A. Coddington. 1991. On missing entries in cladistic analysis. *Cladistics* 7: 233-243.

Pleijel, F. 1995. On character coding for phylogeny reconstruction. *Cladistics* 11: 309-315.

Reimers, H. 1954. Bryophyta. pp. 218 – 268. *In* A. Engler, H. Melchior and E. Werder-mann, Syllabus der pflanzenfamilien, 12, völlig neugestaltete Aufl. Bd. I, Berlin-Nikolassee.

Saiki R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, N. Arnheim. 1985. Enzymatic Amplification of β-Globin Genome sequences and restriction site Analysis of sickle cell Anemia. *Science* 230**:** 1350-1354.

Sainsbury, G. O. K. 1955. A handbook of the New Zealand Mosses. Royal Society of New Zealand Bulletin. No. 5. pp. 18-23. The royal Society of New Zealand. Wellington.

Schofield W. B. and C. Hebant, 1984. The Morphology and anatomy of the moss gametophore, *In* R. M. Schuster. *New Manual of Bryology* Vol.2 The Hattori Botanical Laboratory, Nichinan, Miyazaki.

Schofield, W. B. 1985. *Introduction to Bryology.* The Lantern Mosses Subclass Andreaeidae pp. 23-31. Macmillan. New York, London.

Schultze-motel, W. 1970. Monographie der Laubmoosegattung *Andreaea* 1 Die costaten Arten. Willdenowia 6: 25-110.

Scott, A. M. G., I. G. Stone, C. Rosser. 1976. *The mosses of Southern Australia.* pp. 61 — 65. Academic press, London, New York, San Francisco.

Sim, T. R., 1926; *The Bryophyta of South Africa*, Transactions of the Royal Society of South Africa, Vol. XV. pp 118, 133-134

Smith, A. J. E. 1986. Bryophyte phylogeny; fact or fiction? *Journal of Bryology* 14: 83-89.

Smith, D. K. and P. G. Davidson. 1993. Antheridia and sporophytes of *Takakia ceratophyllla* (Mitt.) Grolle: Evidence for a reclassification among the mosses, *Journal of the Hattori botanical laboratory* 73:263-271.

Soltis, D. E. and P. S. Soltis. 1998. Choosing an approach and an appropriate Gene for phylogenetic analysis, *In* D. E. Soltis, P. S. Soltis and J. J. Doyle, *Molecular Systematics of Plants II: DNA sequencing*. pp.1-42, Kluwer Academic Publishers, Massachusetts New York.

Swofford, D. L. 1991. When are phylogeny estimates from Molecular and Morphological data incongruent? pp. 295-333 *In* Phylogenetic analysis of DNA sequences (M. M. Miyamoto and J. Cracraft eds.) Oxford University Press, New York.

Steere, W. C. and B. M. Murray. 1976. *Andreaeobryum macrosporum*, a new genus and species of Musci from Alaska and Canada. *Phytologia* 33: 407-410.

Swofford, D. L. 1998. *"PAUP*, Phylogenetic analysis using parsimony (*and other methods)," version 4.0b4a*, Sinauer Associates, Sunderland, MA.

Terry, R. G., R. S. Nowak. R. J. Tausch. 2000. Genetic variation in chloroplast and nuclear ribosomal DNA in Utah juniper (Juniperus osteosperma, Cupressaceae): evidence for interspecific gene flow. *American Journal of Botany* 87(2): 250-258.

van Rooy, J. 1997. Introduction to bryology in southern Africa; 2. The structure and life of mosses. *Plant Life* 17: 27-29.

Vitt, D. H. 1984. Classification of Bryophyta. In R. M. Schuster. *New Manual of Bryology* Volume 2: 969-759. Hattori Botanical Laboratory. Nichinan.

Watson, E. V. 1971. *The structure and life of Bryophytes*. The Anchor press Limited. Essex.

Wilkinson, M. 1995. A comparison of two methods of character construction, *Cladistics* 11: 297-308.

Yang, Z. 1994. Maximum likelihood phylogeny estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* 39: 306-314.

Yang, Z., N. Goldman. and A. Friday. 1995. Maximum likelihood trees from sequences: A peculiar statistical estimation problem. *Systematic Biology* 44(3): 384-399.

Appendix 1A

List of Herbarium specimens examined for gametophytic and sporophytic characters. The classification is derived from Braithwaite (1887), Murray (1988b) and Matteri & Farias (1999). BG = Herbarium of University of Bergen, Norway, BOL = Bolus Herbarium, University of Cape Town. HIB = Herbarium instituti botanici, UBC = University of British Columbia Herbarium. T. H. = Dr. Terry Hedderson's Herbarium.

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Andreaea	*Andreaea*	*A. rupestris* Hedw. var. rupestris	Hedderson T. A 10439	Newfoundland, Canada.	T. H.
Andreaea	*Andreaea*	*A. rupestris* var. *papillosa* (Lindb.) Podp.	Einar Heegaard, 249	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Andreaea*	*A. rupestris* var. *papillosa* (Lindb.) Podp.	Einar Heegaard, 262	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Andreaea*	*A. rupestris* var. *papillosa* (Lindb.) Podp.	Hedderson T.A 6640	Newfoundland, Canada	T. H.
Andreaea	*Andreaea*	A. *rupestris* Hedw. var. *rupestris*	Hedderson T. A 1845	Newfoundland, Canada.	T. H.
Andreaea	*Andreaea*	*A. rupestris* Hedw. var. *rupestris*	Einar Heegaard, 6	Frafjord, Rogaland, Norway.	BG
Andreaea	*Andreaea*	*A. rupestris* Hedw. var. *rupestris*	Einar Heegaard, 261	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. acutifolia* Hook et Wils.	Schofield 108339 With Talbot	British Columbia, Canada.	UBC

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Andreaea	*Andreaea*	*A. alpestris* (Thed.) Schimp.	Einar Heegaard, 208	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Andreaea*	*A. alpestris* (Thed.) Schimp.	Einar Heegaard, 277	Haukelidseter, Telemark, Norway, Mapnr.	BG
Andreaea	*Andreaea*	*A. alpestris* (Thed.) Schimp.	Hedderson T. A 4755	Newfoundland, Canada.	T. H
Andreaea	*Andreaea*	*A. alpina* Hedw.	Einar Heegaard, 3	Miganfjell, Rogaland, Norway.	BG
Andreaea	*Andreaea*	*A. alpina* Hedw.	Hedderson T. A 12916	Newfoundland, Canada.	T. H
Chasmocalyx		*A. australis* F. Muell. ex. Mitt.	Smith R. 1145	Southern Georgia, USA	BOL
Andreaea	*Andreaea*	*A. obovata* var. *obovata* Thed.	Einar Heegaard, 268	Haukelidseter, Telemark Norway.	BG
Andreaea	*Andreaea*	*A. obovata* var. *obovata* Thed.	Einar Heegaard, 263	Haukelidseter, Telemark Norway.	BG
Andreaea	*Andreaea*	*A. obovata* var. *hartmannii*	Einar Heegaard, 247	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Andreaea*	*A. alpina* Hedw.	Einar Heegaard, 31	Rage, Gjesdal, Rogaland, Norway.	BG

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Andreaea	*Andreaea*	*A. sinuosa* B. Murr.	R. L. Halbert 6802	British Columbia, Canada.	UBC
Andreaea	*Andreaea*	*A. sinuosa* B. Murr.	W. B. Schofield, J. Spence 84073	British Columbia, Canada.	UBC
Andreaea	*Andreaea*	*A. sinuosa* B. Murr.	W. B. Schofield, I. A. Worley, 37892	British Columbia, Canada.	UBC
Andreaea	*Andreaea*	*A. sinuosa* B. Murr.	W. B. Schofield, J. Spence 74208	British Columbia, Canada.	UBC
Andreaea	*Andreaea*	*A. sinuosa* B.M.Murray	W. B. Schofield, J. 14561	British Columbia, Canada.	
Andreaea	*Nerviae* Card. ex. Broth.	*A. blyttii Schimp.*	Einar Heegaard, 256	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. blyttii* Schimp.	Einar Heegaard, 279	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. blyttii* Schimp.	Hedderson T. A 4976	Newfoundland, Canada	T. H
Andreaea	*Nerviae* Card. ex. Broth.	*A. crassinervia*	Einar Heegaard, 239	Haukelidseter, Telemark, Norway	BG
Andreaea	*Andreaea*	*A. frigida* Huebener	Einar Heegaard, 245	Haukelidseter, Telemark, Norway.	BG

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Andreaea	*Andreaea*	*A. frigida* Huebener	Einar Heegaard, 252	Haukelidseter, Telemark, Norway.	BG
Andreaea	*Andreaea*	*A. gainii* Card.	B. Jablonski M167	POLONAE- CRACOVIAE	HIB
Andreaea	*Andreaea*	*A. gainii* Card.	Green S. R. 1470	Southern Georgia, USA.	**BOL**
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *epapillosa* B. Murr.	Schofield 86799	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr ssp. *epapillosa* B. Murr.	Schofield 31437	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *megistospora*	Schofield 83690	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *megistospora*	Schofield 74851	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *megistospora*	Schofield 77730	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *megistospora*	Einar Heegaard, 87	Toftøy, Sotra, Hordaland, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. megistospora* B. Murr. ssp. *megistospora*	Einar Heegaard, 93	Kåravika, Sotra, Hordaland, Norway.	BG

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Andreaea	*Nerviae* Card. ex. Broth.	*A. rothii* F.Weber & D. Mohr. ssp. *rothii*	Hedderson T. A 11770	Norway	T. H.
Andreaea	*Nerviae* Card. ex. Broth.	*A. rothii* ssp. *falcata* (Schimp.) Lindb.	Einar Heegaard, 5	Frafjord, Rogaland, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. rothii* ssp. *falcata* (Schimp.) Lindb.	Einar Heegaard, 49	Mulen, Bergen, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. rothii* ssp. *rothii*	Einar Heegaard, 46	Mulen, Bergen, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	*A. rothii* Web & Mohr ssp. *rothii*	Einar Heegaard, 54	Storekvit, Os, Hordaland, Norway.	BG
Andreaea	*Nerviae* Card. ex. Broth.	A. *schofieldiana* B. Murr.	R. L. Halbert 6507	British Columbia, Canada.	UBC
Andreaea	*Nerviae* Card. ex. Broth.	*A. schofieldiana* B. Murr.	W. B. Schofield 12555	British Columbia, Canada.	UBC
Chasmocalyx Lindb. ex. Braithw.	*Nitidae* W.	*A. nitida* Hook. & Wils.	E. Esterhuysen 18475.	Cape Town, R.S.A.	BOL
Chasmocalyx Lindb. ex. Braithw.	*Nitidae* W.	*A. nitida* Hook. & Wils.	E. Esterhuysen 25110.	Cape Town, R.S.A.	BOL

Subgenus	Section	Species and infra-specific taxa	Collector	Locality	Source
Chasmocalyx Lindb. ex. Braithw.	*Nitidae* W.	*A. nitida* Hook. & Wils.	E. Esterhuysen 18475.	Wuppertal Division, Cape, R.S.A.	BOL
Chasmocalyx Lindb. ex. Braithw.	*Nitidae* W.	*A. nitida* Hook. & Wils.	E. Esterhuysen 15347.	Clanwilliam Division, Cape, R.S.A.	BOL
Chasmocalyx Lindb. ex. Braithw.		*A. nivalis* Hook.	Einar Heegaard, 250	Haukelidseter, Telemark, Norway.	BG
Chasmocalyx Lindb. ex. Braithw.		*A. nivalis* Hook.	Einar Heegaard, 253	Haukelidseter, Telemark, Norway.	BG
Chasmocalyx Lindb. ex. Braithw.		*A. nivalis* Hook	Hedderson T. A 5016	Newfoundland, Canada	
Andreaea	*Nerviae* Card. ex. Broth.	*A. subulata* Harv. ex Hook.	Hedderson T. A 13229	Cape Town	T. H.
Acroschisma		*A. wilsonii Harv. ex Hook.*	C. J. Cox and B. Goffinet No. 668/00	Rosales	

Appendix 1B

List of Herbarium specimens for the non-*Andreaea* species examined for gametophytic and sporophytic characters. UBC = University of British Columbia Herbarium. T. H. = Dr. Terry Hedderson's Herbarium.

SPECIMEN	COLLECTOR	LOCALITY	HERBARIUM
Sphagnales			
***Section Acutifolia* Wils.**			
Sphagnum girgensohnii Russ.	W. B. Schofield, 82257.	British Columbia, Canada.	UBC
Subsecunda (Lindb.) Schimp.			
Sphagnum pylaesii Brid.	Hedderson T.A 3755.	Newfoundland, Canada.	T. H.
Takakiales			
Takakia Lepidozioides S. Hatt. & Inoue.	W. B. Schofield, 26962.	British Columbia, Canada.	UBC
Takakia Lepidozioides S. Hatt. & Inoue.	W. B. Schofield, BE Lemmon & R. C. Brown. s.n.	British Columbia, Canada.	UBC
Takakia Lepidozioides S. Hatt. & Inoue.	W. B. Schofield, 38319.	British Columbia, Canada.	UBC
Tetraphidales			
Tetraphis geniculata Girg. Ex. Milde c. fr. Hedw.	Hedderson T.A 3787.	Newfoundland, Canada.	T. H.
Tetraphis pellucida Hedw.	Hedderson T.A 5699.	Newfoundland, Canada.	T. H.