The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

PHYLOGEOGRAPHIC PATTERNS IN THREE SOUTH AFRICAN FOREST MOSSES

by

Elizabeth Mwafongo

A thesis submitted in partial fulfilment of the requirements for the degree Masters of Systematics and Biodiversity Science, in the Department of Botany, Faculty of Science, University of Cape Town, Rondebosch, South Africa.

University of Cape Town

Supervisor: Dr Terry A.J. Hedderson

June, 2003

UNIVERSITY OF CAPE TOWN

TABLE OF CONTENTS

Page

TITLE PAGE I

TABLE OF CONTENTS II

ACKNOWLEDGEMENTS VI

ABSTRACT VII

LIST OF FIGURES VIII

LIST OF TABLES X

CHAPTER 1. INTRODUCTION

1.1 South African forests 2

1.2 Phylogeography: Inferring historical processes 3

1.3 Molecular markers 6

1.3.1 Some assumptions and potential problems 7

1.3.2 Application of non-coding chloroplast DNA in plant systematics 9

1.3.3 Nuclear DNA markers 12

1.3.4 Practical concerns of using nuclear DNAs in phylogeography 13

1.3.5 Status of phylogeographic studies in plants 14

1.4 A brief introduction to bryophytes 15

1.5 Bryophytes and phylogeography 17

1.6 Objectives 19

University of Cape Town

CHAPTER 2. MATERIALS AND METHODS 20

2.1 Study species 21

2.1.1 *Wardia hygrometrica*... 21

2.1.2 *Pterogonium gracile* 24

2.1.3 *Leptodon smithii* 26

2.2 Sampling and laboratory methods 28

Plant material 28

2.2.1 Molecular methods 34

DNA extraction procedure 34

Amplification and Sequencing of the ITS and trnL-F gene regions 36

2.3. Data analyses 37

Character treatment 37

2.3.1 Neighbour- joining analyses.......................................38

Separate analysis...38

Combined analysis...39

2.3.2 Population demography...39

2.3.3 Phylogegeographic analysis 40

2.3.4 Population divergence estimates...............................41

University of Cape Town

CHAPTER 3. RESULTS 42

3.1 *Leptodon smithii* 43

3.1.1 Levels of divergence and Neighbour-joining analyses...........46

Separate analysis of ITS gene region...................................46

Separate analysis of TrnL-F data 45

Combined data (ITS+trnL-F) 47

3.1.2 Population demography..50

3.1.3 Haplotype distribution and phylogeography 52

Haplotype distrbution in *L. smithii* 52

ITS data 52

trnL-F data 55

3.2 Wardia hygrometrica 57

3.2.1 Neighbour-joining analyses 57

Separate analysis of ITS data 57

Separate analysis of TrnL-F data.......................... 60

Combined (ITS+trnL-F) 63

3.2.2 Population demography 66

3.2.3 Haplotype distribution in *Wardia hygrometrica*....... 68

ITS haplotype distribution 68

TrnL-F haplotype distribution 72

CHAPTER 4 DISCUSSION AND CONCLUSIONS 74

4.1 Major lineages in *L. smithii* 76

4.2 Phylogeographic pattern in *L. smithii* 77

4.3 Major lineages in *Wardia hygrometrica* 78

4.4 Lack of variation in *P. gracile* 80

4.5 Phylogeographic hypotheses 80

4.6 History and age of the forests in southern Africa 81

University of Cape Town

4.7 Sequence variation in the ITS and trnL-F regions and their suitability for phylogeographic studies 82

4.8 Conclusions 83

4.9 Limitations of the study and future prospects 83

Sampling considerations 83

The future of phylogeography in southern Africa 84

REFERENCES 86

APPENDICES 98

University of Cape Town

ACKNOWLEDGEMENTS

I wish to thank Southern African Botanical Diversity Network for their financial assistance. My sincere thanks to Dr Terry Hedderson for his continuous supervision and encouragement through the entire period of the project. I would also like to thank Tracey Nowell for her assistance with all the lab work, Di James (University of Cape Town) and Carel (University of Stellenbosch) for their help in sequencing, Gonzalo Aguilar for all the technical help, Terry Trinder-Smith and Cornelia Clarke for helping me in the herbarium. I would also like to thank the Botany Department for their support. Many thanks goes to my Husband Baleke and son Austin for their moral support and for enduring my long absence from home. Finally, I would like to thank God for carrying me through the entire period of my study.

University of Cape Town

ABSTRACT

Leptodon smithii (Leptodontaceae) and *Pterogonium gracile* (Leucodontaceae) are widespread and disjunctly distributed moss species, that in South Africa are confined to forest patches believed to be relics of vast forests that existed before the Pleistocene glaciation period. These two species exhibit similar distribution and ecologies and frequently co-occur. *Wardia hygrometrica* is a southwestern Cape endemic that is restricted to streams within the kloofs that the forests generally occupy. In an attempt to trace the species histories and their subsequent colonization of forest patches, two molecular markers, internal transcribed spacer (ITS) nuclear rDNA and trnL-F cpDNA were employed. The *ITS* and *trn* regions in *W. hygrometrica* and *L. smithii* revealed the occurrence of multiple haplotypes and significant geographical structuring of the populations. Within *L. smithii*, the South African haplotypes are highly divergent from other sampled regions in Europe, Asia, and Africa. However, in *P. gracile*, there was no geographic structuring in either the *ITS* or *trn* regions suggesting that the species has attained its current distribution only recently. These results suggest that even bryophyte floras from a single biome may be a composite of species with different histories.

University of Cape Town

Keywords: Phylogeography, molecular markers, haplotypes, *Wardia hygrometrica*, *Leptodon smithii*, *Pterogonium gracile*.

LIST OF FIGURES

FIGURE 1	Geographic distribution of *W. hygrometrica* in southern Africa
FIGURE 2	Geographic distribution of *P. gracile* in southern Africa
FIGURE 3	Geographic distribution of *L. smithii* in southern Africa
FIGURE 4	ITS Neighbour-joining tree for *L. smithii*
FIGURE 5	TrnL-F neighbour-joining tree for *L. smithii*
FIGURE 6	Combined ITS+trnL-F tree for *L. smithii*
FIGURE 7	Geographic distribution of ITS+trnL-F clades in *L. smithii*
FIGURE 8	Mismatch distribution for *L. smithii*
FIGURE 9	ITS haplotype network for *L. smithii*
FIGURE 10	TrnL-F haplotype network for *L. smithii*
FIGURE 11	ITS neighbour-joining tree for *W. hygrometrica*
FIGURE 12	Geographical distribution of the ITS lineages in *W. hygrometrica*
FIGURE 13	TrnL-F neighbour-joining tree for *W. hygrometrica*
FIGURE 14	Geographical distribution of the trnL-F lineages in *W. hygrometrica*
FIGURE 15	Combined ITS+trnL-F neighbour-joining tree using maximum likelihood distances of *W. hygrometrica*
FIGURE 16	Geographical distribution of the combined ITS+trnL-F lineages in *W. hygrometrica*
FIGURE 17	Mismatch distribution per population of *W. hygrometrica*
FIGURE 18	TCS haplotype network for *W. hygrometrica* based on ITS sequence data

University of Cape Town

FIGURE 19 TCS haplotype network for *W. hygrometrica* based on trnL-F sequence data

University of Cape Town

LIST OF TABLES

TABLE 1. List of accessions used

TABLE 2. Primers used for amplification and sequencing

University of Cape Town

CHAPTER 1. INTRODUCTION

University of Cape Town

1.1 South African forests

In southern Africa, natural forests exist as large or small remnants scattered from Table Mountain in the southwestern Cape, across coastal regions of the southern and eastern Cape, and link up with eastern Africa (Magill 1981). The forest remnants vary in size from the larger Knysna forest in the southern Cape, to the tiny kloof forests of the west.

The patchiness of forests in southern Africa led Acocks (1953) and others to hypothesize that the forests are the climatic climax over much of the eastern seaboard and southern coast of South Africa, and that they are relictual (Midgley et al. 1997). A record of forest history in southern Africa provides evidence for a much wider distribution of forest patches during the middle and late Pleistocene (Scott et al. 1997) that has been suggested by palaentological data. Climatic changes during the Last Glacial Maximum (180,000 yrs) led to fluctuation in the size of forests and the only patches that survived this ordeal were those in protected mountain valleys (Scott et al. 1997). In addition, continental movements are believed to have influenced global climate and the establishment of the Antarctic Circumpolar Current had an effect on Africa's vegetation (Shackleton & Kenent 1975). Alternative hypotheses are that forest patches reflect recent expansion (Meadows & Linder 1993) or that forests have always been patchily distributed (Midgley et al. 1997).

University of Cape Town

A broad range of organisms is more or less restricted to these forest patches, and these seem unable to occur in the sea of fynbos or other vegetation types in which the forests occur. Such species with comparable natural histories or habitat

requirements can be tested for phylogeographic pattern. Shared patterns should reflect shared historical elements in the evolutionary or ecological factors that shape the intraspecific phylogeographic architectures within these biotas (Rosen 1975). Among these is an assemblage of mosses that are entirely restricted to forest patches in southern Africa. In the present study we use phylogeographic approaches to examine population structure in these mosses in order to test some of the current hypotheses regarding forest history.

1.2 Phylogeography: Inferring historical processes

Studies in the history of a biota are multifaceted (Myres & Giller 1988). The most common is the historical geography of the biota, which investigates the way in which patterns of distributions may have changed in the past (Linder et al. 1992). Methods in historical biogeography ranges from narrative, fitted to sets of observations to rigorous hypothesis testing (i.e cladistic approaches) (Ball 1976). While the narrative approach may produce a story, rich in testable hypotheses, the relationship between it and the physical geographic history is not always clear (Linder et al. 1992). However, with the advent of vicariance biogeography, the relationship between extant patterns of distributions and historical changes in the geography can be tested (reviewed in Linder et al. 1992).

University of Cape Town

The field of "phylogeography" has recently evolved (Avise et al. 1987). Phylogeography shares a conceptual framework with historical biogeography (Arbogast & Kenagy 2001), but also permits investigation of biogeographic questions on spatial and temporal scales that are smaller than those typically addressed with other approaches (Avise 2000). It is a field of study concerned with

the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species (Avise 2000). The field deals with historical, phylogenetic components of the spatial distributions of gene lineages (Avise 2000).

The analysis and interpretation of lineage distributions usually requires extensive input from molecular genetics, population genetics, ethology, demography, phylogenetic biology, paleontology, geology, and historical geography (Avise 2000). Thus, phylogeography is an integrative endeavour that lies at important crossroads of diverse microevolutionary and macroevolutionary disciplines (Avise 1998).

As a sub-discipline of biogeography, phylogeography serves to place into a broader temporal context traditional ecogeographic perspectives that emphasize the role of contemporary ecological pressures in shaping the spatial distributions of organismal traits (Avise 2000). Phylogeography also serves as a useful conceptual umbrella covering alternative historical scenarios to account for the spatial arrangements of organisms and their features, with vicariance and dispersal as two often-competing possibilities. These two historical phenomena are often suggested to account for the origins of spatially disjunct taxa.

University of Cape Town

Avise et al (1987) first coined the term "phylogeography" to refer to the relationship between mitochondrial DNA (mtDNA) phylogeny and geographic distribution in animals. Since that time, numerous studies have used variation in the mitochondrial genome to analyze phylogeographic patterns in a diverse array of

animal species (Mitton 1994). This cytoplasmically-housed genome is usually targeted in animals because of its phylogenetically favourable properties of maternal transmission, and thus an effective population size one-fourth that of the nuclear genome (Avise et al. 1987). Secondly, the mtDNA genome has a relatively rapid rate of molecular evolution (Arbogast & Kenagy 2001). These features of the mtDNA lead to non-reticulating (bifurcating) gene trees and rapid geographic sorting and genetic divergence of populations in the absence of gene flow (Arbogast & Kenagy 2001). As a result, phylogeographic studies using mtDNA usually provide more resolution of intraspecific patterns of geographical variation than non-molecular methods (Avise 2000). Phylogeographic studies (sensu Avise et al. 1987) involving plants and cpDNA are less common than those involving animals and mitochondrial DNA. Presumably this is because plant mtDNA, which also dwells in the cell cytoplasm and encodes similar suites of genes, does not exhibit similar evolutionary genomic patterns . Instead, plant mtDNA evolves rapidly with respect to gene order yet slowly in nucleotide sequence (Palmer 1991).

Since phylogeography was first introduced (Avise et al. 1987), it has been used primarily to examine geographical structuring of gene lineages within single species (Arbogast & Kenagy 2001). Typically, individuals are sampled from throughout the geographical range of species, and the target genome is characterized for each individual, through either restriction fragment analysis or direct sequencing. The resulting haplotypes are then used to infer a phylogeny, or gene tree, which reflects the evolutionary relationships of the individuals and populations sampled (Arbogast & Kenagy 2001). By combining the resulting gene trees with geographical locations from which each individual was sampled, one can

University of Cape Town

elucidate the geographical distributions of major gene lineages (monophyletic clades) that comprise the gene tree (Arbogast & Kenagy 2001). The application of coalescent and likelihood methods and approaches such as nested clade analysis allows for the interpretation of phylogeographic patterns within the context of evolutionary and biogeographic models. Thus, phylogeography is a powerful approach for investigating a wide range of issues related to biogeography, including the relative roles of gene flow, bottlenecks, population expansion, and vicariant events in shaping geographical patterns of genetic variation.

1.3 Molecular markers

The spatial genetic architecture of any species is likely to be a complex outcome of contemporary demographic and ecological forces acting upon a pre-existing population structure that was moulded by biogeographic factors operative throughout the evolutionary history of a species (Walker & Avise 1998). Molecular methods are well suited for (a) describing current population genetic structures and (b) recovering historical components of those structures (Walker & Avise 1998). Molecular markers have contributed much to the study of plant evolutionary ecology by providing methods for detecting genetic differences among individuals, especially at population level when morphological variation contains little information (Hess et al. 2000).

University of Cape Town

The application of molecular markers to population biology began with the benchmark studies of Harris, Hubb, and Lewontin (Arnold & Emms 1998). These studies used isozyme variation to test two models of natural selection: the classical (purification selection) model and the balancing selection model. Although these

models make different predictions about the amount and pattern of standing genetic variations in populations, Lewontin (1991) argued that isozyme studies have not adequately distinguished between them (Arnold & Emms 1998). Despite this setback, isozymes have paved the way for application of DNA analyses. This technique has also allowed a remarkable degree of understanding of population-level processes and continues today to be a powerful diagnostic tool for the study of plant and animal populations.

After isozymes, the next major technological advance came in the form of DNA re-association (Britten & Koehn 1968). This technique was used to address macroevolutionary questions such as the origin of evolutionary novelties and the effect of genomic evolution, which later on extended to the use of DNA restriction data and DNA sequencing. The application of molecular markers, especially of nucleotide sequences, is a recent development, but due to their high resolution and relevance for systematic surveys, they are the methods of choice at present (Cruzan 1998). Several of these methods are capable of detecting single nucleotide mutations. For example, highly variable regions of DNA can sometimes provide a unique 'fingerprint' for each individual and access to such fine scale genetic variation is one of the compelling reasons for using DNA. DNA offers advantages over allozymes in that it is found in nearly all cells of all organisms and it can be recovered from both living and dead tissue (Cruzan 1998).

University of Cape Town

1.3.1 Some assumptions and potential problems

Many studies provided evidence that molecular variation is high in most natural populations (Kimura & Ohta 1972). The neutral theory proposed by Kimura (1968a

& b) asserts that a "great majority of evolutionary changes at the molecular levels are caused not by Darwinian selection but by random drift of selectively neutral mutants" (Kvist 2000). The theory, originally applied to protein variation, does not deny the role of natural selection in determining the course of adaptive evolution. Rather, it assumes that only a minute fraction of DNA changes are adaptive in nature, while the great majority of phenotypically silent molecular substitutions exert no significant influence on survival and reproduction and drift randomly through species (Kvist 2000). Without the equivalent of dated fossils for the ancestral sequences in gene trees, application of the neutral model provides the only solution available to the problem of time scaling (Harding 1996). Time-scaling on all gene trees is based on the assumption that the turnover of variation at the DNA sequence locus of interest ticks with the regularity of a molecular clock (Kimura 1983). That is, there is a strict functional neutrality among the DNA sequence variants and the nucleotide mutation rate, which is the time-scaling factor, should be constant (Harding 1996).

Compared with gene trees that relate different taxa, there is an added complication for interpreting individual-level DNA variation and this is due to their assembly from DNA polymorphism rather than from fixed nucleotide differences (Harding 1996). Whereas fixation is simply proportional to mutation rate, transient polymorphism results from a balance between the input of new variants by mutation and their loss by genetic drift (Kimura 1983). Since genetic drift occurs as stochastic variation in reproductive success when each successive generation is sampled from a previous one, it is therefore a tracer of population size (Harding 1996). In order to appreciate intraspecific gene trees and demographic history, it is necessary to

University of Cape Town

consider the interaction between time and population size. It is therefore worth investing in these new phylogenetic studies because this is the only possible way in which details of the demographic history of a species can be inferred (Harding 1996). Neutral molecular variation together with population genetics theory offers powerful tools for ecological research. A variety of neutral molecular markers, including mini and micro-satellites, different kinds of restriction fragments, PCR markers, and DNA sequences are available today for studying ecology and evolution.

The distribution of genetic diversity within and among populations is affected both by recurrent factors and by historical events (Actander et al. 1999). In this respect, the use of fast evolving DNA regions is critical because these genes have proven useful in inferring species recent history and evolution despite some inherent drawbacks (Actander et al. 1999). For instance, historical fragmentation causes departures from some assumptions of population genetics theories and leads to overestimated levels of population differentiation (Actander et al. 1999). Conversely, true ancestry may be obscured by current gene flow. One way to assess the occurrence and magnitude of a historical event is to compare the patterns observed in several co-distributed species and determine the extent of their biogeographic congruence (Zink 1996).

University of Cape Town

1.3.2 Application of non-coding Chloroplast DNA in plant systematics

There is a growing interest in comparative analysis of non-coding chloroplast sequences for plant systematic studies at low taxonomic levels (Kelchner 2000). Recognition of the limitations of coding (genic) DNA for resolving relationships at

these levels inspired the probing of chloroplast introns and intergenic spacers for phylogenetic utility (Kelchner 2000). Underlying this effort was the reasonable premise that non-coding regions experience limited or no selective pressure and are more likely to evolve at rates surpassing those of genic regions (Wolfe et al. 1987; Palmer, 1987 &1991). There was also an expectation that non-coding regions should experience random and independent mutations, both in mode and distribution (Kelchner 2000).

Variation patterns in intraspecific chloroplast DNA (cpDNA) have been the subject of recent studies examining numerous aspects of evolutionary genetics including migration patterns and rates, drift, and population structure (reviewed in Golden & Bain 2000). In addition, such variation has been used to infer hybridisation and introgression and to trace the ancestry of polyploids (Rieseberg & Brunsfeld 1992; Liston & Kadereit 1995). Although the rate of evolution of the chloroplast genome is more conservative than that of the mitochondrial genome of animals, intraspecific cpDNA variation in angiosperms is much more extensive than initially thought (reviewed in Soltis et al. 1992). Furthermore, cpDNA variation is geographically structured in some plant species (Soltis & Novak 1997). Thus, sufficient cpDNA variation exists within many plant species to permit the elucidation of evolutionary processes and the study of intraspecific phylogeography. Over the years, a series of studies employing cpDNA restriction site variation has revealed a similar phylogeographic pattern in a diverse array of plant species from different geographic areas (Soltis et al. 1989a, 1991, 1992). Over 50 examples of intraspecific cpDNA variation have been reported in plants (reviewed in Soltis & Soltis 2000). When patterns of variation in the chloroplast genome of a species

University of Cape Town

have geographic basis, inferences can be made concerning its biogeographic history (Liston et al. 1992). In a similar manner, the degree of molecular divergence between vicariant species can provide an important source of evidence for the biogeographic processes responsible for their current distribution (Kvist 2000).

A considerable amount of work already published has demonstrated the potential phylogenetic utility of discrete non-coding regions in the chloroplast genome (reviewed in Kelchner 2000). Such regions are easy to amplify in a wide range of plants and are known to provide short segments of DNA that are easily sequenced (Soltis & Soltis 2000). One such region is the *trnL (UAA)-trnF*(GAA) intergenic spacer. This region is small (120-350bp), and has proved very useful in phylogenetic reconstruction at lower taxonomic levels (Soltis & Soltis 2000). As a molecular marker, *trnL-trnF* has many advantages. It is known to evolve three times faster than *rbcL* (Soltis & Soltis 2000). For example, in the fern family Ophioglossaceae, this region is about one-third the size of *rbcL*, with sequence divergences for the spacer three to five times higher than sequence divergence for *rbcL* for the same taxa (Hauk et al. 1996). This gene region has also been used widely in intraspecific studies. For example, Fugi et al. (1995), report on the intraspecific sequence variation of *Primula cuneifolia* (Primulaceae), a perennial herb occurring in the subalpine and alpine areas of Japan. In their study they found six distinct haplotypes, each of which was found to be geographically restricted. Results from the above study and literature from elsewhere reveal extensive use of this gene region. Recognition of the potential use of non-coding cpDNA regions is critical for the understanding of bryophyte phylogeography attempted in this study.

University of Cape Town

1.3.3 Nuclear DNA markers

The most widely used nuclear DNA markers for measuring population genetic variation in plants have been obtained from ribosomal loci (Soltis & Soltis 2000). These markers have the advantage of being easily detected and the degree of variability typically observed makes them ideal for measuring such processes of gene flow within and between taxa (Soltis & Soltis 2000).

The discovery of variation among sequences of the internal transcribed spacer (ITS) region of 18S-26S ribosomal RNA genes has led to new opportunities for systematic studies at lower taxonomic levels. The ITS region has shown its greatest utility in phylogeny reconstruction at the interspecific and the intergeneric levels (reviewed by Baldwin et al. 1995), providing a data source comparable to variation in chloroplast DNA (cpDNA) restriction sites (e.g., Baldwin 1992; Soltis and Kuzoff 1995; Manos 1997; Sang et al. 1997). One of the most attractive aspects of using the ITS region for plant molecular phylogenetic studies is the ease with which sequences of both spacers can be obtained and interpreted (Baldwin et al. 1995). The small size, highly conserved flanks, high copy number, rapid concerted evolution, and long conservation of the ITS sequences greatly aid their PCR amplification, sequencing, alignment, and phylogenetic analysis (Baldwin 1992). A minor sequencing effort can yield complete sequences of ITS-1 and ITS-2 from many taxa (Baldwin et al. 1995). In addition, ITS sequences are believed to have few evolutionary constraints, despite the fact that secondary structural elements in both ITS regions are known to play an important role in the processing of the pre-rRNA molecule (Mattaj et al. 1994; van Nues et al. 1994, cited in Bakker et al. 1995b). Schlötterer et al. (1994) showed that in alignable portions of

University of Cape Town

Drosophilla ITS, approximately 60% of the sequence is free to diverge, the rest being constrained by secondary structure (Bakker et al. 1995b). The unconstrained portion was found to evolve at a highly divergent rate (2.4% per Ma), which is similar to divergence rate estimates in alignable portions of Cladophora ITS sequences (0.8-2.0% per Ma). According to Bakker et al. (1995b), these divergence rates provide the right level of resolution for phylogeographic studies in the range of 1-25 Ma, which fits perfectly with major Miocene and Pleistocene paleoclimatic events that have shaped most terrestrial and marine provinces. Although the results of several studies have suggested this potential (Sytsma & Schaal 1990; Baldwin 1993; Walker & Paris 1997), analyses of ITS variation among conspecific populations are still uncommon.

1.3.4 Practical concerns of using nuclear DNAs in phylogeography

The marker of first choice for phylogeography will continue to be plasmid DNA for initial characterisation of population structure, testing population monophyly, or inferring gene flow. However, studies of this nature often use nuclear markers to corroborate initial results based on cytoplasmic loci (Hare 2001). In theory, phylogeographic structure is expected to be less pronounced at diploid nuclear loci than at cytoplasmic loci because autosomal nuclear loci have an effective population size (N_e) four times larger than that of uniparentally inherited cytoplasmic markers. Thus under a neutral model of evolution, genetic drift will cause divergence between isolated populations to occur four times more slowly at nuclear than at cytoplasmic loci, all else being equal (Hare 2001).

University of Cape Town

Two practical concerns have been raised about using nuclear haplotypes for phylogeography (Avise 2000). The first concern is that they are impeded by recombination. Secondly, poor resolution could result from low mutation rates, leading to too few informative polymorphisms (Hare 2001).

The question of whether nuclear DNA can be considered a strictly neutral marker has been controversial (Soltis & Soltis 2000). Studies have shown that rDNA data suffer from the same potential major weakness as do isozyme data – a lack of sufficient variation for some purposes (Soltis & Soltis 2000). However, the high value of using nuclear DNA in population studies is not severely diminished by the uncertainty of nuclear neutrality.

1.3.5 Status of phylogeographic studies in plants

Phylogeographic studies in plants have lagged behind those of animal studies, primarily because of difficulties in identifying DNA sequences with appropriate levels of ordered variation within chloroplast, mitochondria, or nuclear genomes. In addition, chloroplast spacer regions that have been informative for some species, have been problematic because of the difficulty in finding regions that have sufficient levels of neutral variation. High-resolution nuclear markers such as random amplified length polymorphisms (RFLPs) and amplified fragment length polymorphisms (AFLPs) are known to be unordered, and variants cannot be related easily in a genealogical manner (Schaal & Olsen 2000). Because of the difficulty in finding genealogical informative markers in plants, many studies have been phylogeographic only in broad sense, meaning that they detect an association between patterns of genetic variation and geography but fail to

University of Cape Town

incorporate a genealogical perspective (Schaal & Olsen 2000). However, in spite of these problems, chloroplast and nuclear DNA variation is present and known to be structured geographically in several plant species (reviewed in Avise 2000). Some of the most elegant studies in plants have examined the postglacial migration of species from Pleistocene refugia. It has been shown that most organisms presently distributed across Europe were in refugia in the south at the height of glaciation 18000BP, many in the peninsulas of Iberia, Italy, and the Balkans and some possibly near the Caucasus and Caspian Sea (Hewitt 1999). When the ice retreated (from about 16000BP), species expanded their ranges out of these refugia northwards. There is evidence of early spread of pine, oaks (*Quercus spp*), beech (*Fagus sylvatica*), black alder (*Alnus glutinosa*), and elm (*Ulanus spp*) up the western fringe to Brittany, Ireland and Scotland, believed to have been transported by water currents or animals (Hewitt 1999). Investigators interpret this cline as a result of postglacial migration from the same refugia, leading to the concordance of variation patterns among species (Schaal & Olsen 2000).

1.4 A brief introduction to bryophytes

The bryophytes are a diverse group of "lower" land plants, with some 23,000 described species worldwide, making it the largest group of land plants excepting flowering plants (Mishler 1997). The group consists of three quite distinct lineages (i.e., mosses, hornworts, and liverworts) of species encountered in a broad range of terrestrial habitats including mesic forests, streams, tropical rain forests, arctic tundra, and desert boulders.

University of Cape Town

Their physical structure and physiological attributes (for example, desiccation resistance/drought tolerance, nutrient-capturing abilities) allow them to play major roles in many of the world's ecosystems, affecting water and mineral fluxes, controlling surface run-off and erosion, and providing food and habitat for a wide range of organisms (Newton et al 2000). The bryophytes are generally considered as a "key" group in our understanding of how the modern land plants (the embryophytes) are related to each other phylogenetically and how they came to conquer the hostile land environment from their ancestral home in fresh water. This is because of their apparent basal phylogenetic position among the embryophytes; the remnant lineages present today are from a spectacular radiation of the land plants in the Devonian period, some 400 million years ago (Mishler 1997).

Bryophytes, at all taxonomic levels, tend to have distributional ranges that correspond to historical biogeographic patterns of tracheophytes, but most species of bryophytes are more widely distributed (Watson 1971). Many families are cosmopolitan (Sim 1918). Even at species level, the list of cosmopolitan species is quite large (reviewed in Watson 1971).

University of Cape Town

The fact that bryophytes, like other spore producing plants, generally tend to have broad ranges suggests that long distance dispersal of spores or other propagules is an important factor explaining their distributions (Shaw 1995). Correlations between spore longevity and breadth of geographic distribution have been reported (Zanten & Pócs 1983). While disjunct populations are usually thought of as morphologically uniform, the extent to which such patterns reflect underlying genetic structure is not known (Shaw 1995).

1.5 Bryophytes and phylogeography

Plants offer unique challenges and opportunities for the study of ecological genetics. With the exception of a few species, plants are sedentary over most of their life history, so the genetic composition of mature individuals may reflect selection on early stages by local microenvironments (Trewick & Wallis 1999). From a population-genetic point of view, this lack of mobility combined with prevalence of asexual reproduction is expected to contribute to a higher degree of genetic structure, resulting in a population that consists of groups of closely related individuals. This process can be intensified by varying levels of self-compatibility in hermaphroditic species, contributing to local inbreeding and the generation of highly homozygous offsprings. Hence, understanding both genetic and ecological processes is critical as these are most likely to carry molecular signals that are informative of phylogenetic events in space and time (Trewick & Walls 1999).

It is important to know the extent to which genetic variation of a species distributed within and among populations for a sound understanding of bryophyte evolution (Akiyama 1999). There has been much discussion in recent years of the "evolutionary potential" of bryophytes (Wyatt & Odrzykoski 1998). The view of these plants as evolutionary failures, lacking genetic variability, is based on the argument that dominance of a gametophytic generation precludes heterozygosity and permits selection to act directly on the haploid genotype (Wyatt & Odrzykoski 1998)

The prevalence of asexual reproduction among bryophytes has been cited as an additional factor reducing genetic variation within species (Wyatt & Odrzykoski 1998). However, the prevailing view that bryophytes maintain little genetic variation

University of Cape Town

has been seriously challenged in the past decade by evidence from electrophoretic studies. A number of previous studies on bryophytes have shown that bryophytes exhibit genetic variation comparable to diploid-dominant plants (reviewed in Akiyama 1999). For species depending totally or heavily on asexual reproduction, there are at least two additional factors limiting genetic variation within and among populations. If there is no migration among populations, all members of a particular population are of clonal origin and belong to the same genet. In this case, genetic variation accumulated within a population can be attributed to mutation, which, according to Akiyama (1999) is a very rare event compared to genetic shuffling through meiosis in sexual reproduction, and hence a much lower genetic variation is expected. The second factor is that asexual species lack spores, which are supposed to be the most effective type of propagules for dispersal. Thus, it is expected that there will be fewer genetic exchanges among asexual populations compared to sexual species. According to Akiyama (1999), such populations would be expected to accumulate unique mutations, which cause genetic diversification especially among distant populations. This pattern has been documented in *Pteridium aquilinum* and similar situations have been recorded in bryophytes and liverworts (reviewed in Longton & Hedderson 2000)

Bryophytes present unique opportunities for the study of phylogeography combining simplicity with technical convenience. DNA and RNAs are also easy to extract in bryophytes (Cove 2000). The haploid nature of the gametophytes from which genomic DNAs are usually extracted, makes it possible for researchers to work with nuclear genes.

University of Cape Town

1.6 Objectives

The aim of the work was to examine the extent of genetic variation within and diversity among populations of three mosses, *Leptodon smitthii, Pterogonium gracile, and Wardia hygrometrica* using the chloroplast *trnL-trnF* spacer region, and the nuclear rRNA *ITS* region. The main reasons for choosing these species was their association with riverine forest patches which could help elucidate the history of forest patches in southern Africa.

What is revealed will reflect events that took place in the remote past. At the same time it will reveal the dispersal mechanisms of southern African forest bryophytes. Data gathered are relevant to the following major questions:

1. Are forest moss populations structured geographically in southern Africa?
2. What is the amount of gene flow between the sampled populations?
3. How did past historical events affect the present population structures?
4. How is genetic variation partitioned within versus among populations of the three moss species?
5. How old are the forest patches in which the bryophytes occur?

University of Cape Town

CHAPTER 2. MATERIALS AND METHODS

University of Cape Town

2.1 Study species

Due to the short time in which the study had to be conducted (6 months), only three species were considered. The three species show ecological affinity with forest patches. Their distributions are therefore very discontinuous, with populations often separated from each other by large distances (1km or more). This kind of structure provides an excellent opportunity to examine within and among population variation and the history of the forests in which the bryophytes occur.

2.1.1 *Wardia hygrometrica*

Wardia hygrometrica Harv. & Hook belongs to the monotypic family Wardiaceae. It is endemic to the Southwest Cape (Fig. 1), where it is abundant on rocks in running streams (Sim 1918, Hedderson et al. 1999). *W. hygrometrica* is associated with forests because it grows in streams within the kloofs that the forests generally occupy.

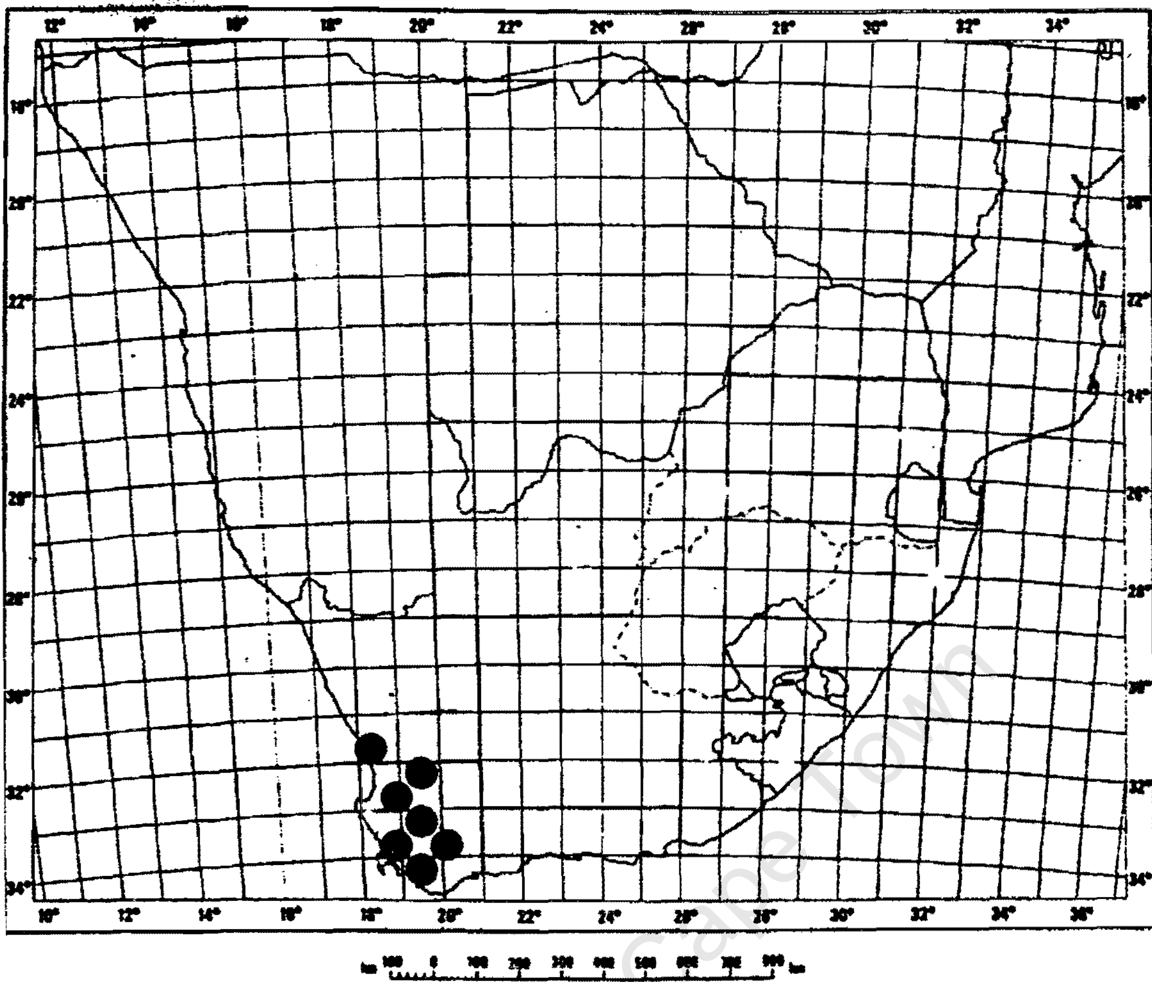
Wardia is only one among a plethora of moss taxa whose phylogenetic relationships have been difficult to determine because they possess unusual morphologies (Hedderson et al. 1999). However, molecular data Hedderson et al. (in press) were unambiguous in placing *Wardia* as sister to a clade comprising the Dicranaceae and Dicnemonaceae, suggesting that it represents an ancient divergence.

Endemics, whether at family, genus or species level, have always held a special interest for biogeographers. There are of two principle kinds of endemics (Watson

University of Cape Town

1971). Firstly, there are forms so recently evolved that they have lacked the time to achieve a wider distribution (neoendemics). Secondly, there are those so ancient that, having vanished from their former stations elsewhere, they have become restricted to a single area (paleoendemics). Still others may have evolved quite far back in time but because of geographical barriers remained endemic. Phylogenetic studies have shown that *Wardia hygrometrica* represents an ancient divergence, but the age of the species remains unclear.

University of Cape Town



University of Cape Town

Figure 1. Map showing the geographic distribution of *W. hygrometrica* in South Africa

2.1.2 *Pterogonium gracile*

Pterogonium gracile (Hedw.) Sm. belongs to the family Leucodontaceae, a family restricted to woodlands and forests throughout America, Europe, Africa, and parts of Asia (Magill & van Rooy 1998). *P. gracile* occurs in eastern Africa, the Eastern African islands, Macaronesia, Europe, the Middle East, and western North America. In southern Africa, the species occurs in the northern and eastern Transvaal areas, KwaZulu-Natal, and eastern, southern, and southwestern Cape (Fig. 2).

P. gracile is known only as female plants and sporophytes have not been reported in the past ten years (Magill & van Rooy 1998). However, recently (September 2001), sporophytes have been found in populations growing in ravines on Table Mountains, Cape Town, South Africa by Dr Terry Hedderson (Pers.comm). Nevertheless, *P. gracile* seems to depend totally or heavily on asexual reproduction for its propagation.

University of Cape Town

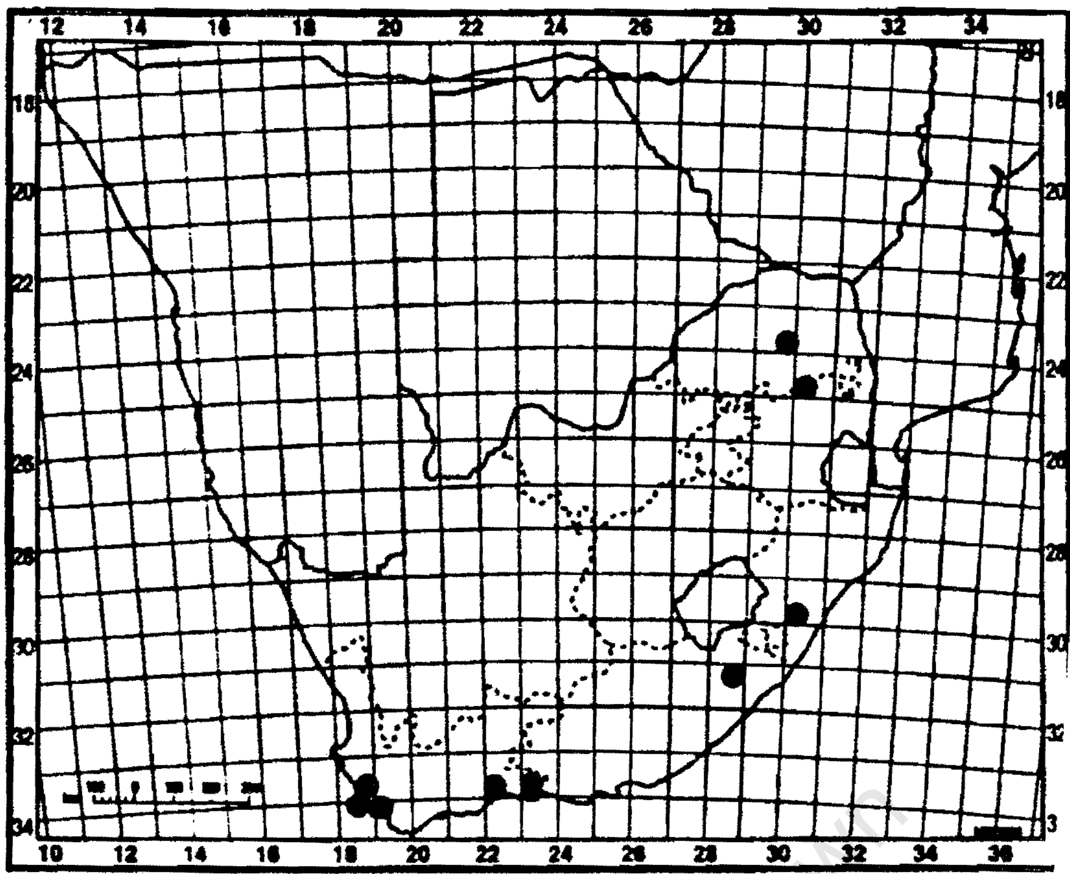


Figure. 2. Map showing the geographical distribution of *P. gracile* in southern Africa (After Magill & van Rooy 1998)

University of Cape Town

2.1.3 *Leptodon smithii*

Leptodon smithii (Hedw.) Web. & Mohr. belongs to the family Leptodontaceae. The genus is widely, although sporadically, distributed throughout the world and according to Magill & van Rooy (1998) appears to be of southern hemisphere origin. The genus is represented in Africa by a single species (Magill & van Rooy 1998). It is known to occur in Europe, southwestern Macaronesia Asia, Australia, South America, Juan Fernández islands, and New Zealand (Magill & van Rooy 1998). In southern Africa, the species is found in the northern, central and eastern Transvaal areas, Zululand, KwaZulu-Natal, Eastern Free State, and northwestern Cape regions (Fig. 3). *L. smithii* is found growing on bark and rock in woodlands and forests.

University of Cape Town

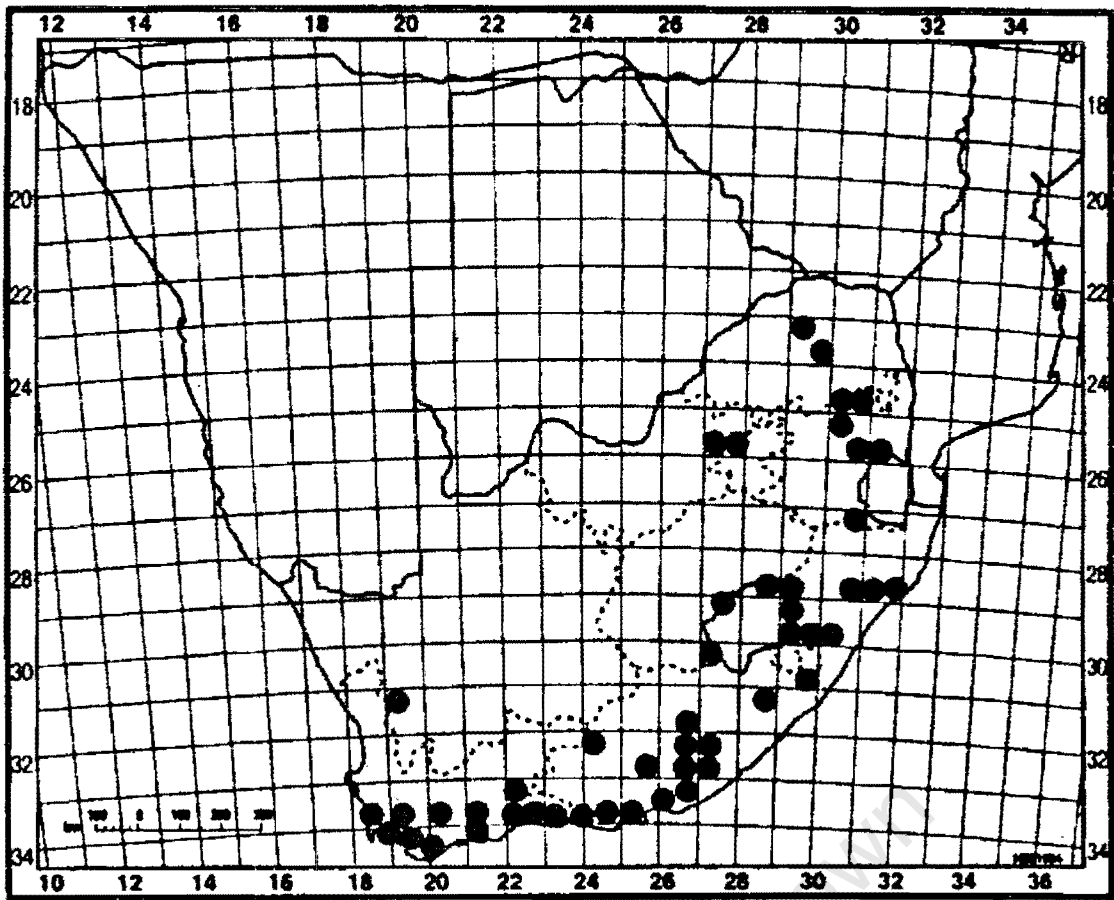


Figure. 3. Geographical distribution of *L. smithii* in southern Africa (after Magill & van Rooy 1998)

University of Cape Town

2.2 Sampling and laboratory methods

Plant material

Populations of *P. gracile*, and *L. smithii*, were sampled from Africa, Madeira, and Europe whereas those of *W. hygrometrica* were collected throughout the species distributional range in the western Cape region. Twenty-six individuals in 10 populations of *W. hygrometrica*, 32 individuals from 11 populations of. *L.smithii*, and 27 individuals from 10 populations of *P. gracile* were sampled (Tab. 1). Populations were chosen to represent the geographic range of these species, although inevitable logistic constraints limited the numbers and distributions that could be sampled. Dried herbarium specimens were used in this study. Since amplification of both regions failed for some herbarium material, field collection for fresh material was carried out wherever possible.

University of Cape Town

Table1. List of accessions used for the sequencing of *ITS* and *trn* of *Wardia hygrometrica*, *Leptodon smithii,* and *Pterogonium gracile*.

Abbreviations: BOL, Bolus Herbarium, South Africa; S, Stokholm Herbarium, Sweden; PRE, National Herbarium, Pretoria, South Afric

Lab. No	Species Name	Origin	Locality	Coll/No.	Date	ITS	TRN	Herb.
1W	*Wardia hygrometrica*	Cape, South Africa	Mcgregor, De hook. Houtbais Rivier	Hedderson 13864	2001	X	X	BOL
2W	*Wardia hygrometrica*	Cape, South Africa	Riviersonderend Mts. Boesmanskloof	Hedderson 13920	2001	X	X	BOL
3W	*Wardia hygrometrica*	Cape, South Africa	Hottentot's Holland Mts.	Hedderson 13123	2001	X	X	BOL
4W	*Wardia hygrometrica*	Cape, South Africa	Mcgregor,De hook kloof, Houtbais rivier	Hedderson 13864 1/3	2001	X	X	BOL
5W	*Wardia hygrometrica*	Cape, South Africa	Riviersonderend mts. Boesmans kloof	Hedderson 13920	2001	X	X	BOL
6W	*Wardia hygrometrica*	Cape, South Africa	Orangekloof, Table Mts.	Hedderson 13484	2001		X	BOL
7W	*Wardia hygrometrica*	Cape, South Africa	Baineskloof	Hedderson 13235	2001	X	X	BOL
8W	*Wardia hygrometrica*	Cape, South Africa	Baineskloof	Hedderson 13236	2001	X	X	BOL
9W	*Wardia hygrometrica*	Cape, South Africa	Baineskloof	Hedderson 13237	2001	X	X	BOL
10W	*Wardia hygrometrica*	Cape, South Africa	Orangekloof, Table Mts.	Hedderson 13567	2001		X	BOL
11W	*Wardia hygrometrica*	Cape, South Africa	Sunbird nature reserve, Silvermine	Hedderson 13155	2001	X	X	BOL
13W	*Wardia hygrometrica*	Cape, South Africa	Dehook Kloof, Houtbais Rivier	Hedderson 13864 2/3	2001		X	BOL
14W	*Wardia hygrometrica*	Cape, South Africa	Riviersonderend mts. Boesmans Kloof	Hedderson 13920	2001	X		BOL

University of Cape Town

15W	*Wardia hygrometrica*	Cape, South Africa	Cederberge Mts.	Hedderson 13054	2001	X		BOL
16W	*Wardia hygrometrica*	Cape, South Africa	Hottentot's Holland Mts.	Hedderson 13123	2001	X	X	BOL
17W	*Wardia hygrometrica*	Cape, South Africa	Boesmans Kloof	Hedderson 13920	2001	X	X	BOL
19W	*Wardia hygrometrica*	Cape, South Africa	Ceres district Oliphant's River	Hedderson 14000	2001	X	X	BOL
20W	*Wardia hygrometrica*	Cape, South Africa	Skeleton gorge, Table Mts.	Hedderson 12820	2001	X	X	BOL
21W	*Wardia hygrometrica*	Cape, South Africa	Silvermine Nature Reserve	Hedderson 13285	2001		X	BOL
22W	*Wardia hygrometrica*	Cape, South Africa	Jonkershoek, on rock/waterfall	EM1	2001	X	X	BOL
23W	*Wardia hygrometrica*	Cape, South Africa	Jonkershoek, on rock/waterfall	EM2	2001	X	X	BOL
24W	*Wardia hygrometrica*	Cape, South Africa	Jonkershoek, on rock/waterfall	EM3	2001	X		BOL
25W	*Wardia hygrometrica*	Cape, South Africa	Table mts. Kasteel's Poort	Hedderson 14213	2001		X	BOL
26W	*Wardia hygrometrica*	Cape, South Africa	Groot Winterhoek area, Tulbagh	Hedderson 14302 1/3	2001	X	X	BOL
27W	*Wardia hygrometrica*	Cape, South Africa	Cederberge wilderness area	Hedderson 14371	2001	X	X	BOL
28W	*Wardia hygrometrica*	Cape, South Africa	Hottentots Holland Mts. Landroskop	Hedderson 14408	2001	X	X	BOL
1L	*Leptodon smithii*	Madeira	Ribeira da ponta do solo	Fontinha, Hedenäs, Nobrega B9718	1991	X		S
2L	*Leptodon smithii*	Madeira	Ribeira da Ponta do Solo	Fontinha, Hedenäas, Nobrega B9718	1991	X	X	S
3L	*Leptodon smithii*	Madeira	Cidrao	Hedenäas B9717	1991	X	X	S
4L	*Leptodon smithii*	Madeira	Siera Margarita	Een B57829	1993	X	X	S
5L	*Leptodon smithii*	Portugal	Coimbra	Hedenäas B44518	2000	X		S

University of Cape Town

6L	*Leptodon smithii*	Natal, South Africa		Magill 5130	1998		X	X	PRE
7L	*Leptodon smithii*	Cape, South Africa	Table mountains	Hylander	2000		X		BOL
9L	*Leptodon smithii*	Cape, South Africa	Buffells forest station				X		PRE
10L	*Leptodon smithii*	Cape, South Africa	Buffells forest station,	Magill 5985	1979		X		PRE
14L	*Leptodon smithii*	Traansvaal, South Africa	Graskop mts.,	Stirton 9928	1981		X	X	PRE
15L	*Leptodon smithii*	Transkei, South Africa	Kwa Matiwana Mts.	Van Rooy 2271	1985		X		PRE
16L	*Leptodon smithii*	Cape, South Africa	Caledon	Taylor 10397	1982		X		PRE
18L	*Leptodon smithii*	Natal, South Africa	Uitvlugt farm	Glen 2287	1990			X	PRE
20L	*Leptodon smithii*	Orange Free State, South Africa	Zastron	Van Rooy 2561	1986			X	PRE
22L	*Leptodon smithii*	Eastern Cape, South Africa	Somerset East	Smook 3957	1981			X	PRE
30L	*Leptodon smithii*	Kenya	Masai district	Pócs 89028/A	1989			X	PRE
37L	*Leptodon smithii*	Natal, South Africa	South Africa	Smook 1742	1980			X	PRE
41L	*Leptodon smithii*	Natal, South Africa	Drakensberg	Magill 6754	1984			X	PRE
42L	*Leptodon smithii*	South Africa	Natal, South Africa	Magil 6749	1984		X		PRE
47L	*Leptodon smithii*	Western Cape, South Africa	Orange Gully	Hylander 10823	2000		X		BOL
49L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001		X		BOL
51L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001			X	BOL
52L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001		X	X	BOL
53L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001		X		BOL
54L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001		X	X	BOL
56L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek	EM	2001		X	X	BOL
57L	*Leptodon smithii*	Western Cape, South Africa	Jonkershoek, stream bank	EM	2001	X			BOL

University of Cape Town

58L	*Leptodon smithii*	Western Cape, South Africa	Nursery Ravine, Table Mts.	EM	2001	X	X	BOL
59L	*Leptodon smithii*	Western Cape, South Africa	Nursery Ravine, Table Mts.	EM	2001	X		BOL
60L	*Leptodon smithii*	Western Cape, South Africa	Nursery Ravine, Table Mts.	EM	2001	X		BOL
61L	*Leptodon smithii*	Western Cape, South Africa	Nursery Ravine, Table Mts.	EM	2001	X		BOL
62L	*Leptodon smithii*	Western Cape, South Africa	Nursery Ravine, Table Mts.	EM	2001	X		BOL
67L	*Leptodon smithii*	Drakensberge, South Africa	Olivershoek,	Magill 6759		X		PRE
1P	*Pterogonium gracile*	Uganda	Mt. Elgon	Miehe U31-1074813	1997	X		S
2P	*Pterogonium gracile*	Uganda	Mt. Elgon	Miehe U31-1074901	1997	X		S
3P	*Pterogonium gracile*	Madeira		Hedenäs & Bisang B22467	1999	X		S
4P	*Pterogonium gracile*	Madeira		Hedenäs B9631	1990	X		S
6P	*Pterogonium gracile*	Uganda	Mt. Elgon,	Miehe U31-1075007	1997	X		S
7P	*Pterogonium gracile*	Tanzania	Mt Kilimajaro,	Hedenäs B28308	1998	X		S
8P	*Pterogonium gracile*	Madeira		Hedenäs B4587	1998	X		S
9P	*Pterogonium gracile*	Madeira		Hedenäs B9627	1991	X		S
10P	*Pterogonium gracile*	Spain	Cardiz	Gillis B7832	1993	X		S
11P	*Pterogonium gracile*	Portugal		B44519	2000	X		S
16P	*Pterogonium gracile*	Portugal	Baixo Alentego	Een B57833	1994		X	S
17P	*Pterogonium gracile*	Madeira	Algarve Caldas	Een B57835	1994		X	S
18P	*Pterogonium gracile*	Madeira	Alto Alentejo	Een B57834	1994		X	S
19P	*Pterogonium gracile*	Madeira	Sao Vicente,	Hedenäs B9671	1991		X	S
20P	*Pterogonium gracile*	Madeira		Hedenäs B9672	1991		X	S
21P	*Pterogonium gracile*	Portugal	Ribeiro	Hedenäs B9638	1991		X	S
23P	*Pterogonium gracile*	Madeira	Sao Vicente	Hedenäs B9670	1991		X	S
24P	*Pterogonium gracile*	Portugal	Cidrao	Hedenäs B9632	1991		X	S
29P	*Pterogonium gracile*	Syria	Tartous province,	Nowell s.n.	1999		X	BOL
30P	*Pterogonium gracile*	Western Cape,South	Kirstenbosch Botanic Garden	Hedderson 13816			X	BOL

University of Cape Town

		Africa					
34P	*Pterogonium gracile*	Greece	Lefkada	Hedderson 12101		X	BOL
35P	*Pterogonium gracile*	Western Cape, South Africa	Bosboukloof	Hylander 10918	2001	X	BOL
36P	*Pterogonium gracile*	Western Cape, South Africa	Orangekllof, Disa ravine	Hedderson 13501		X	BOL
37P	*Pterogonium gracile*	Western Cape South Africa	Verlorekloof	Hylander 10917	2000	X	BOL
38P	*Pterogonium gracile*	Stellenbosch, South Africa	Swartboskloof,	Hylander 10919	2000	X	BOL

University of Cape Town

2.2.1 Molecular methods

DNA extraction procedure

About 25mg of plant material was placed in separate 1.5ml microcentrifuge tubes together with a small quantity of sterile beach sand to facilitate the grinding process. These samples were then snap frozen in liquid nitrogen and ground using a plastic pestle. Genomic DNAs were extracted from the powdered tissue using a modified CTAB extraction protocol from Gawel & Jarret (1991). 700µl of 2xCTAB buffer (Doyle & Doyle 1987) and 1µl of β-mercaptoethanol were mixed with each sample, vortexed, and incubated for 1h at 65^{0}C in a water bath. After adding 600µl of chloroform:isoamyl alcohol (24:1,v/v) to each sample, the tissue mixture was centrifuged at 12,000rpm for 5 minutes. The supernatant was transferred to a 1.5-ml microfuge tube followed by addition of an equal volume of ice-cold isopropanol. Thereafter, samples were incubated overnight or for several days at -20^{0}C, and DNA was recovered by centrifuging the mixture at 12,000rpm for 5 minutes. The pellet was then rinsed in 90% ethanol and centrifuged for 3 minutes at 12,000rpm. The sample was dried overnight in a dessicator. The DNA pellet was re-suspended in 50µl of elution buffer and stored in the fridge.

DNA was visualised by running it on a 1% agarose gel stained with ethidium bromide to determine the efficacy of the extraction procedure. From each sample, 10µl of genomic DNA was transferred to 0.5ml microtubes and diluted to either 10^{-1} or 10^{-2} depending on the strength of the extractions as determined by the visualisation process. These dilutions were used in the PCR reactions.

University of Cape Town

Table 2. Primers used for amplification and sequencing the ITS and trnL-F regions

Region	Genome	Primer	Sequence	Direction	Source
ITS	Nuclear	ITS2	GCT GCG TTC TTC ATC GAT GC	Forward	White et al. 1990
		18KRC	GCA CGC GCG CTA CAC TGA	Reverse	Hamby et al. 1988
TrnL-F	Chloroplast	trn-C	CGA AAT CGG TAG ACG CTA CG	Forward	Taberlet et al. 1991
		trn-F	ATT TGA ACT GGT GAC ACG AG	Reverse	Taberlet et al. 1991

University of Cape Town

Amplification and sequencing of the ITS and trnL-F gene regions

Polymerase chain reaction (PCR) was used to amplify the ITS and trnL-F regions. Amplifications were carried out in 50µl reaction volumes consisting of;

5µl of genomic DNA,

5.0µl $MgCl_2$, (50mM)

1.5µl (10µm) of each primer

2.0µl (2.5mM) of dNTPs

0.3µl of *Taq* polymerase (1.5 units)

5µl (10x) of NH_4 buffer (Qiagen)

Thermocycling was carried out in a PCR Sprint Temperature Cycling System (Hybaid Ashford, UK) using the following profile: 97^0C for 2 min. to denature the DNA, followed by 30 cycles of denaturing (1min at 97^0C), annealing (1min. at 52^0C), and extension (2 min at 72^0 C) with a final extension step of 7 min at 72^0 C. Amplified products were resolved on 1% ethidium bromide-stained agarose gels and photographed under UV light. Double-stranded PCR products were purified using QIAquick spin-columns (Qiagen) following protocols provided by the manufacturer and eluted in 30µl of elution buffer (Qiagen).

University of Cape Town

The purified PCR products were cycle sequenced using the ABI Prism Big Dye Terminator kit (P.E. Applied Biosystems) in conjunction with the amplification primers. Cycle sequencing reactions were performed in a total reaction volume of 10µl containing: 3µl of double-stranded PCR product, 4µl Terminator Ready Reaction mix, 0.16µl sequencing primer, and 2.84µl of distilled water. The

sequencing reaction amplification was done in a Hybaid PCR Sprint Thermal Cycler. The cycling protocol consisted of 25 cycles of denaturation (30sec. at 96^{0}C), annealing (15sec. at 50^{0}C), and extension (4 min. at 60^{0}C), with a final soak at 4^{0}C. In order to remove unincorporated dye terminators final products were cleaned using spin-columns before being run on an ABI 3900 autosequencer. Sequences were assembled using SeqMan II (Lasergene System Software, DNAstar Inc.). For each taxon consensus sequences from the two regions were saved as a single file using SeqEdit (Lasergene System Software, DNAstar Inc.) and aligned manually in MegAlign (LaserGene). Following the alignment and editing, all sites that showed any variation were checked to verify that such variation was unambiguous by referring to the raw datas. The alignment was exported as a nexus file for use in subsequent analyses. The alignments for each gene region have been provided on a diskette at the back of the thesis.

2.3. Data analyses

Character treatment

All characters were equally weighted. Where alignment gaps (indels) occurred, these were treated as missing data.

University of Cape Town

2.3.1 Neighbour-joining analyses

Separate analysis

To infer the relationships among populations of *W. hygrometrica* and *L. smithii*, three data sets were used (i) ITS, (ii) trnL-F, and (iii) combined ITS + trnL-F. Each data set was used to construct a neighbour-joining (NJ) tree using PAUP version 4.0b4a (Swofford 2000). The NJ method works by identifying the closest pairs of taxonomic units by the distances between them. A pair of these neighbours is defined to be two units connected through a single node in a tree, where two branches join at each node. This method continues by successive clustering of the lineages, setting branch lengths as the lineages join (Kvist 2000). The neighbour-joining analyses were done using Maximum Likelihood distances under a substitution model estimated from the data. The estimated Ti/Tv ration in this case was equivalent to the HYK85 model (Hasegawa, Kishino & Yano 1985). This model allows transitions (Ti) and transversions (Tv) to occur at different rates, and allows base frequencies to vary as well. The distances obtained were used to construct neighbour-joining trees (Saitou & Nei 1987). The resultant trees (printed as phylograms) were rooted using mid-point rooting. This is because midpoint rooting, as opposed to additive methods, depends on an assumption of rate uniformity that is somewhat weaker than assuming a molecular clock across the entire tree [i.e. if two most divergent lineages have evolved at the same rate, then the appropriate root is at the mid-point of the path connecting these taxa (Swofford et al. 1996)]. In addition, other methods require

University of Cape Town

an outgroup and, as discussed below, defining an outgroup is often not straightforward for population analyses.

Combined analyses

Whether or not to combine different datasets for phylogenetic analysis has been a contentious subject in the field of systematics. Systematists have expressed different views on whether data sets from the same group should be (a) analysed separately (Swofford & Olsen 1991; de Queiroz 1993, (b) combined as total evidence (Miyamoto 1985; Donoghue & Sanderson 1992), or (c) analysed separately before combining (Swofford 1991, de Queiroz 1993). Advocates of these principles give their own reasons as to how their principle is better than the others.

In the current study, ITS and trnL-F data sets were first analysed separately and later combined, to increase the resolving power of the data. Separate analyses were done because we expect different genomes to tell different stories since they do not abide by similar evolutionary rules. The combined data sets included only those taxa from which both of the sequences were available. Tree construction for the combined data sets was performed in the same manner as conducted in the separate analyses.

University of Cape Town

2.3.2 Population demography

Population historical dynamics were inferred from the frequency distributions of pair wise distances between sequences (mismatch distribution; Arctander et al.

1999). Such estimates accumulated for many individuals yield pairwise frequency distributions of coancestry. The advantage of this approach is that in some cases (e.g. under rapid population expansion), a gene tree might be nearly devoid of information on clade structure because many nodes are close, yet a frequency histogram of coancestry could be enlightening with regard to historical population demography (Rogers & Harpending 1992). Chi-square tests provide a means of evaluating the discrepancy of the observed data from respective theoretical distributions (STATISTICA ver. 5.5 1999-2000). Multimodal distributions are consistent with demographic stability, while sudden expansion would generate unimodal patterns. Departures of the observed distribution from that expected under the expansion hypothesis were tested with a chi-squared test of goodness of fit using the Non-Parametric distribution command in the Descriptive Statistic Module (STATISTICA ver. 5.5 Statsoft inc,1999-2000) .

2.3.3 Phylogeographic analysis

The evolutionary relationships of all haplotypes were examined using TCS version 1.13 (Clement et al. 2000). This programme implements the estimation of gene genealogies from DNA sequences. The resultant network is known as a Statistical Parsimony Network. These network methods are better suited to analysing closely related haplotypes than classical phylogenetic methods (Posada & Crandal 2001). This "statistical parsimony" procedure determines the probability of genealogical connections in haplotype networks being parsimonious (Crandal 1994) by providing a measure of the probability that

University of Cape Town

connections between two haplotypes differing by some number of steps are nonhomoplasious (Brower et al. 1996).

No outgroups were chosen in this analysis because in populations, as mutations occur to create new haplotypes, they rarely result in the extinction of the ancestral haplotypes, ie. they are expected to persist in the population (Crandal & Templeton 1996). Gaps were treated as a fifth character.

2.3.2 Population divergence estimates

In most cases, genetic variability analyses assume populations to be in equilibrium with respect to mutations, gene flow, and population size (Kvist 2000). However, this assumption hardly ever holds. For instance, measures of genetic variability may increase in a spatially structured population (Amos & Harwood 1998). On the contrary, the measures may show quite a uniform structure if the population has rapidly expanded in size following a bottleneck, even if the population has some structuring (Kvist 2000). Thus, the knowledge of sequence divergences within and among populations will help in understanding of some of the demographic parameters of the populations under study. In this study, genetic divergence between sequences was estimated as the Maximum Likelihood distances calculated from total character distances under a substitution model estimated from the actual data.

University of Cape Town

CHAPTER 3. RESULTS

University of Cape Town

The ITS and trnL-F regions were variable in *W. hygrometrica* and *L. smithii* but showed no variation among *P. gracile* samples over a wide geographic sampling. Because of this lack of variation in *P. gracile*, no further amplification and sequencing was done for this species. The results presented here are from the analyses of *W. hygrometrica* and *L. smithii* data.

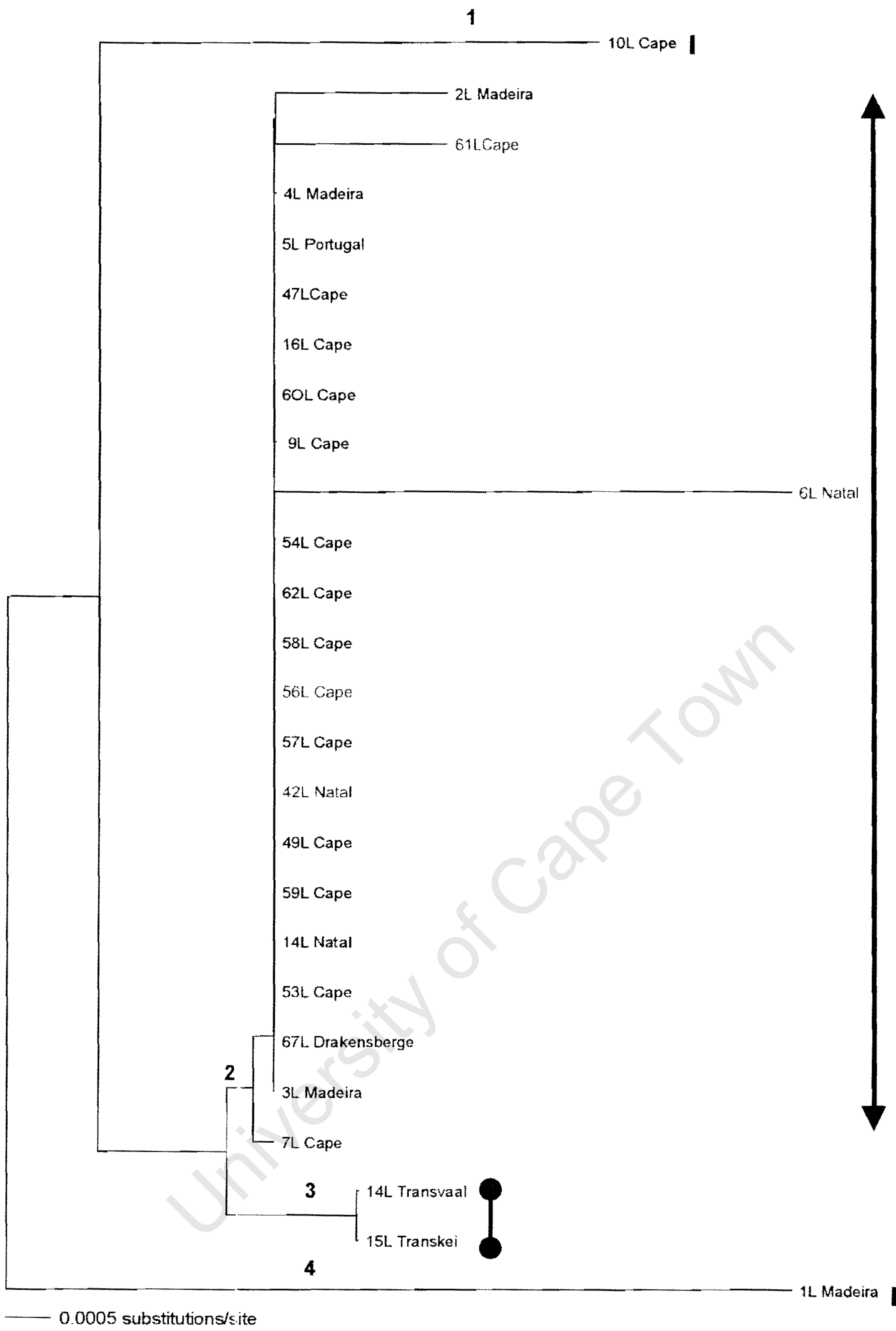
3.1 *Leptodon smithii*

3.1.1 Levels of divergence and neighbour-joining analysis

Separate analysis of ITS gene region

The neighbour joining analysis of the ITS data shows four major clusters (Figure 4). Clusters 1 and 4, contain an individual from the Cape region and Madeira respectively. Cluster 2 contains a mixture of individuals from Africa as well as Europe. The third cluster contains only accessions from the African continent (Eastern Cape and Transvaal). The split up into these four groups does not match the geography since European specimens (2L, 3L, 4L and 5L) are seen clustered within a group containing mostly African accessions (cluster 2). Within the African accessions, there is also very high genetic structuring evidenced from the high structure within the larger cluster 2. However, the differences among them were mostly single nucleotide substitutions. Groups1 and 4 were very divergent from the rest of the accessions (d= 0.004). The most divergent lineage group 4. The genetic distance (d) of this particular accession from the rest of the samples was approximately 0.02.

University of Cape Town



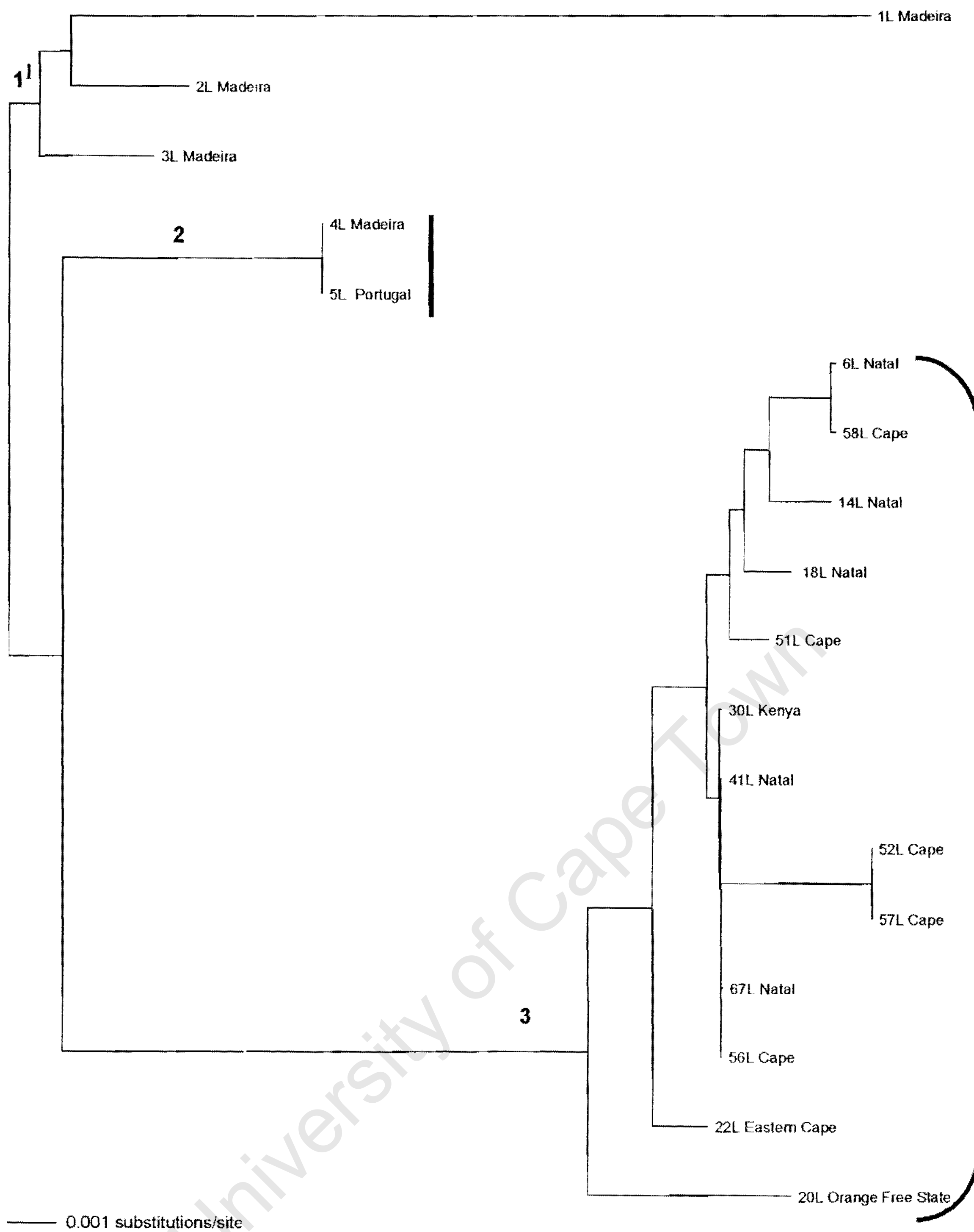
University of Cape Town

Fig. 4. ITS Neighbour-Joining tree using maximum likelihood distances of *L. smithii*. Numbers correspond to sample numbers in Table 1.

Separate analysis of the trnL-F region

The neighbour-joining analysis of trnL-F data is depicted in Figure. 5. At least three major lineages are present within the sampled populations (1, 2, & 3), one of which encompasses all the specimens from Africa (3). Two lineages (1 & 2) are identified within samples from Europe, with strong divergences between the African and European accessions (d=0.02). Of particular interest is the high genetic structure even within each region. For example, the African group (3) is further divided into subclusters indicating genetic distinction among these populations evident from the number of nucleotide substitutions.

University of Cape Town



University of Cape Town

Fig. 5. trnL-F Neighbour-joining tree using maximum likelihood distances of *L. smithii*. Numbers correspond to sample numbers in Table 1. Clades **1** & **2** are European; Clade **3** is African.

Combined data (ITS+trn L-F)

Figure. 6 shows a neighbour-joining tree for the combined ITS and trnL-F data set. Although the sample size was relatively small, the NJ tree shows a split of the accessions into four major groups (1,2,3,4), which perfectly match the geography of the samples. Clusters1, 2, & 4 are solely European whereas 3 is African. These clusters can thus be regarded as phylogeographic groups because of their high specificity to different geographic regions (Figs, 6 & 7). Whereas phylogroups in both Europe and Africa are restricted to their respective geographic regions, limited dispersal and gene flow is observed within each region evidenced by the high degree of structure.

Resolution within geographic groups is greatest in the trees derived from the combined and the trnL-F data sets; less resolution is apparent at this level in the ITS data set.

University of Cape Town

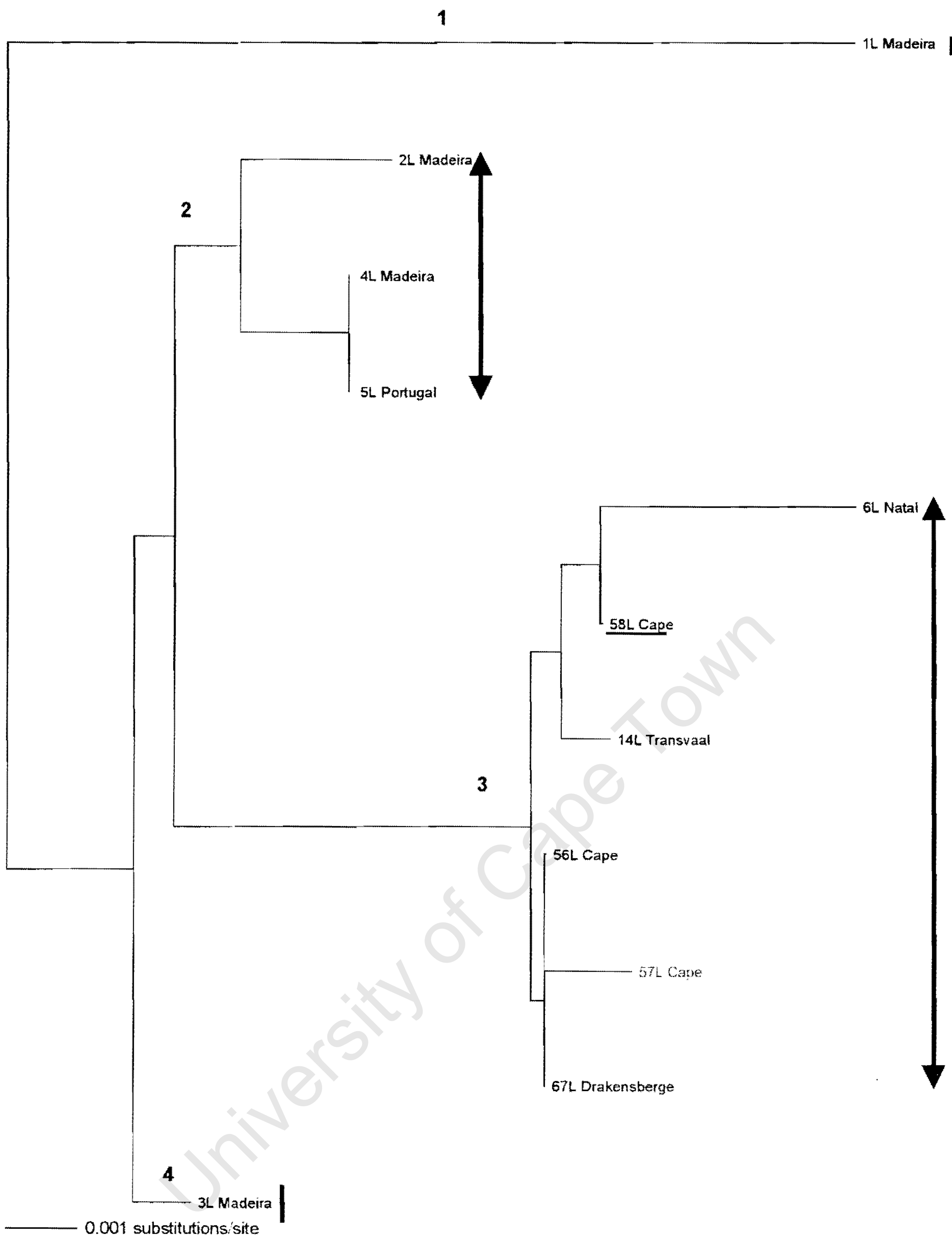
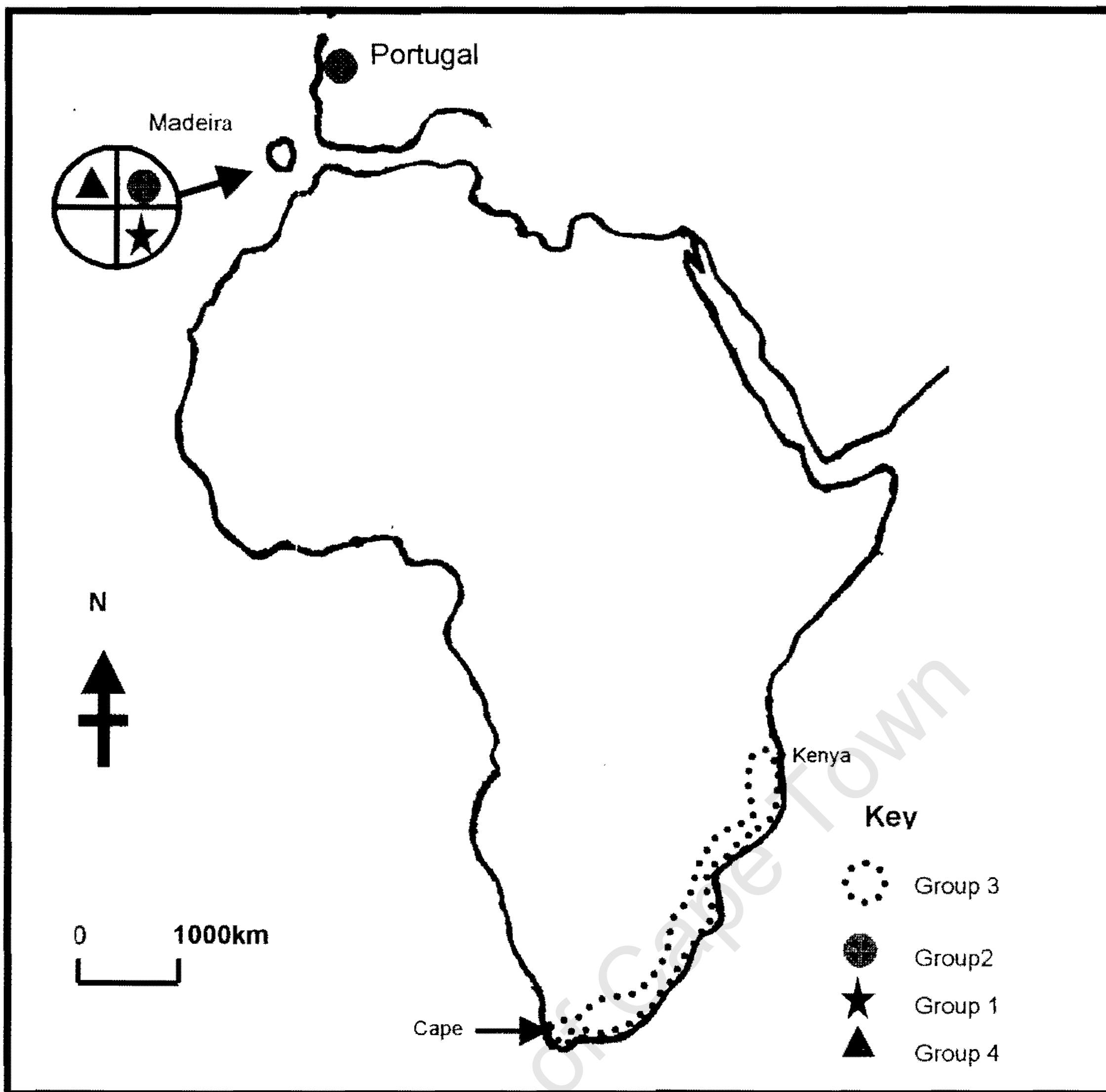


Figure. 6. Combined ITS+trnL-F Neighbour-joining tree using maximum likelihood distances of *L. smithii*. Sample names as in Table 1.

University of Cape Town



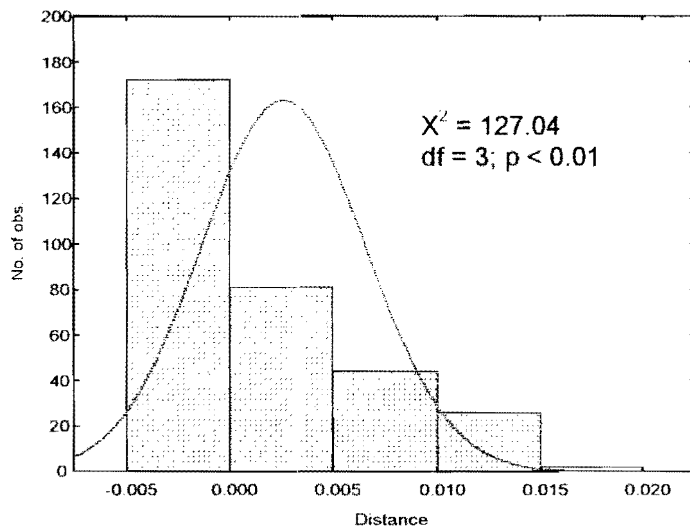
University of Cape Town

Fig.7. Geographic distribution and location of combined ITS+trnL-F lineages in *L. smithii*.

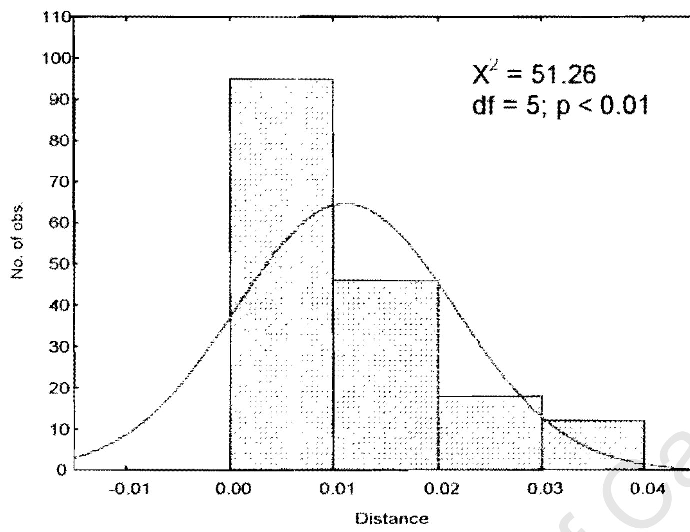
3.1.2 Population demography

Figure. 8 (a, b, & c) shows the mismatch distributions of ITS, trnL-F, and combined data in *Leptodon smithii*. These distributions display one peak and are significantly different from the expectation of either sudden expansion or long term exponential growth ($P < 0.01$; 0.01; 0.01 respectively; Fig. 8a, b, c). Sudden expansion is expected to produce an even distribution of pairwise differences among all geographic groups leading to a smooth and unimodal mismatch distribution. Mismatch distributions from populations of constant size are characteristically ragged and erratic due, in part, to minimal differences among groups, maximal differences among groups, and varying effects of haplotype sorting throughout the gene tree (Matocq et al. 2000). Results from the three distributions are consistent with demographic stability.

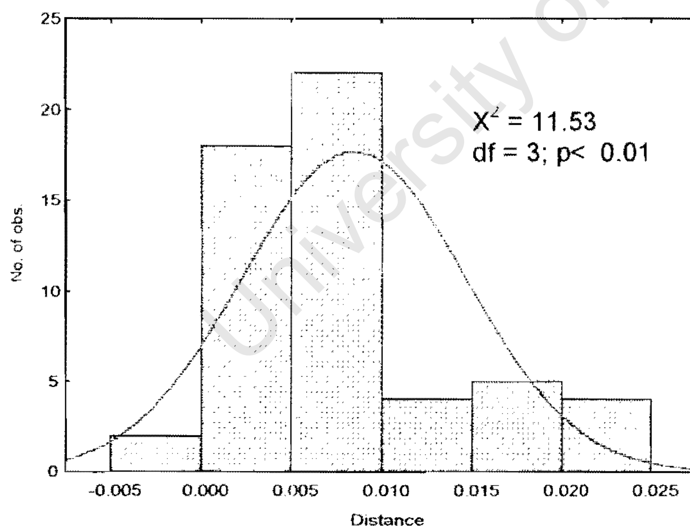
University of Cape Town



(a)



(b)



(c)

University of Cape Town

Figure. 8. Mismatch distribution per population of *L. smithii*. "a" , ITS data, "b", trnL-F "c" , combined ITS+trnL-F. Departures from the expected vaue were tested with X^2 goodness of fit.

3.1.3 Haplotype distribution and phylogeography

Haplotype distribution in *L. smithii*

ITS data

Figure 9 shows the inferred genealogical relationships among the populations observed in *L. smithii*. The ITS sequence data revealed 14 different haplotypes in 25 samples, among which nucleotide sequence divergence (d) was 0.0 - 0.02. This network is more star like compared with the trnL-F network, with the most common haplotype (H5) being observed in 12 individuals and occurring throughout the Cape and Transvaal regions. Haplotype 5 is separated from haplotype H11 (Madeira) by five site changes. Four mutations separate haplotype H5 from haplotype H7 (Cape). Haplotypes H8 and H5 are separated from each other by four site changes; haplotype H14 (Cape) is separated from H11 (Madeira) by five site changes.

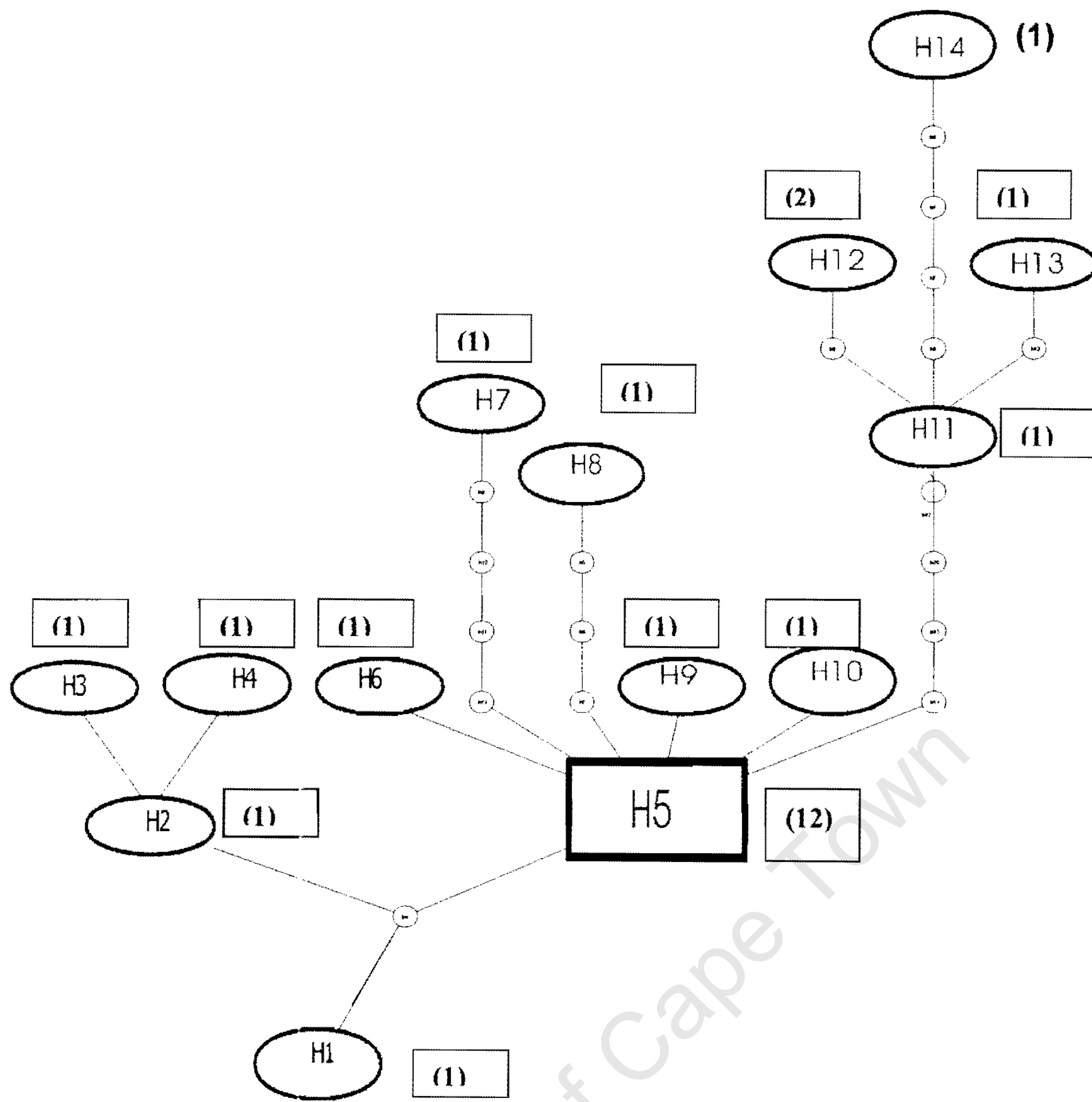
Haplotype sharing among populations is relatively low in *L. smithii*. Only one of the 14 haplotypes is present in more than one population (H5) and this haplotype sharing occurs only within southern Africa. European haplotypes are separated from African ones by an average of 3.1±1.2 mutational steps. Like the patterns depicted in the NJ tree analysis of ITS, trnL-F, and the combined data sets, this network is characterized by regionally unique sets of haplotypes. For example, all haplotypes from the European region (H2, H3, H4, H11, and H13) were unique to Europe and there was no sharing of haplotypes between the two regions. The network also depicts haplotypes 12 and 14 (African) as being closer to European

University of Cape Town

haplotypes. However, there was sharing of haplotypes among populations within each region, which may reflect introgression or, alternatively, ancestral haplotypes.

In Figure 9 an "O" represents an interior node haplotype that was not actually present in the sample. These missing nodes may simply be absent because of insufficient sampling or they may represent extinct ancestral haplotypes. Note also that multifurcations are common throughout this tree and the single most common haplotype in the gene pool (haplotype H5) is at the center of an extreme multifurcation. These topologies, so different from species trees, are exactly what we expect from coalescent theory. The clear-cut genetic differentiation of the populations of *L. smithii* from both Europe and Africa (shown by absence of shared haplotypes), indicates that there is very limited (if any) gene flow among these populations.

University of Cape Town



Geographic locality of haplotypes

H1 – Cape

H2-H4, H11, H13 – Europe

H5- H10; H12-H14 – Cape and Natal

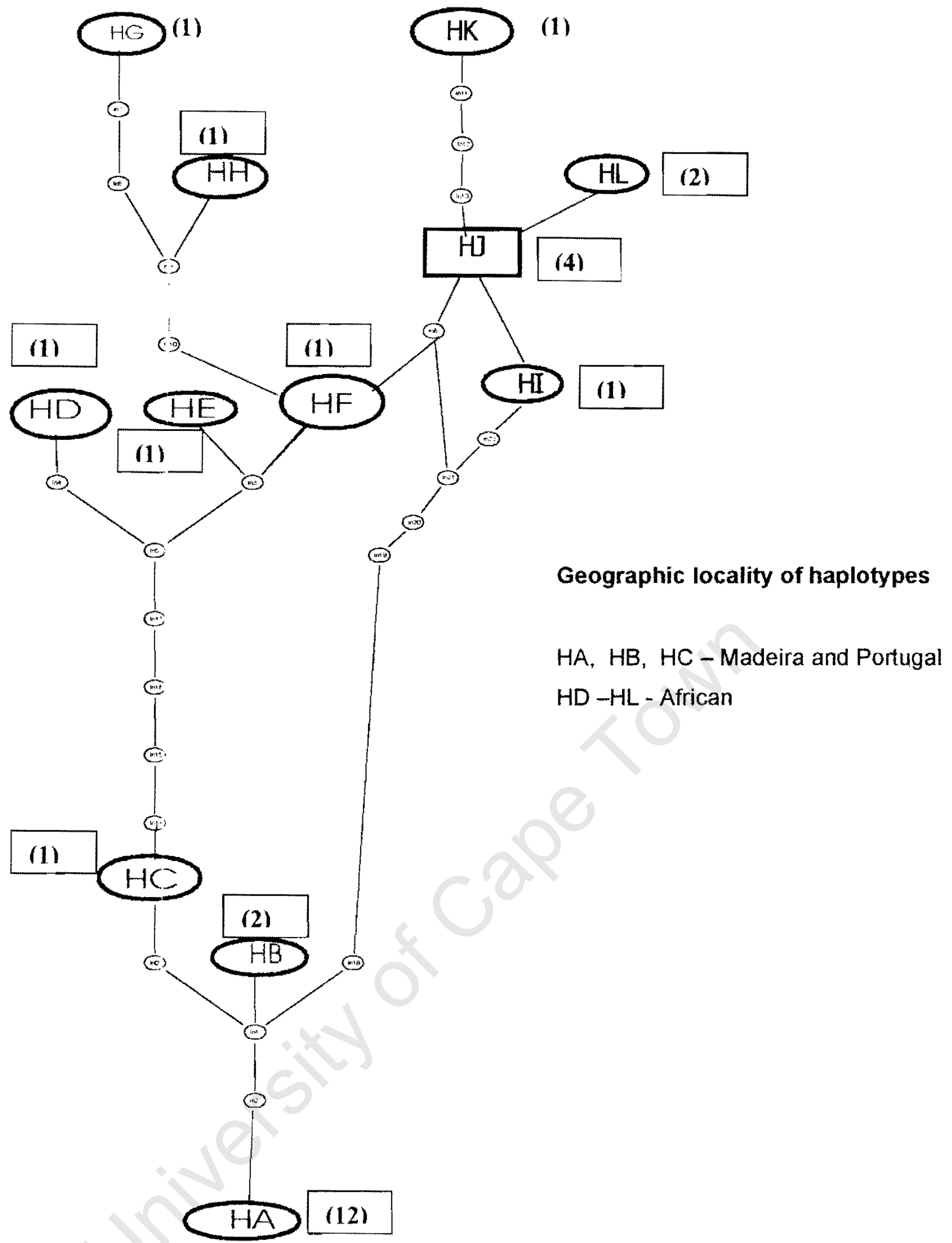
University of Cape Town

Figure 9. TCS haplotype network for *L smithii* based on ITS sequence data. Small circles between haplotypes indicate an intermediate haplotype state that was not found in the sample. Each line between the haplotypes represents a mutational step. The frequency of each haplotype is in parentheses. The square shape depicted in this figure represents the frequency of samples represented.

TrnL-F data

The trn haplotype tree for *L. smithii* was resolved into 12 different haplotypes (Fig. 10), with nucleotide sequence divergence among individuals ranging from 0 - 0.039. Of the 12 haplotypes observed in the study system, three (HA, HB and HC) were detected in populations from the Madeiran and Portuguese regions. Madeira shared haplotype HA with the Portuguese mainland suggesting possible gene flow between the two regions or that the Madeiran population is derived from Europe. The remaining nine were unique to the Cape region, Eastern Cape, Orange Free State, Natal and Kenya. The most common haplotype (HJ) was widespread and was detected in two individuals from the Cape region and in one sample from Natal and Kenya. HJ is in this regarded the hypothesized ancestral haplotype.

University of Cape Town



University of Cape Town

Figure 10. TCS haplotype network for *L smithii* based on *trn*L-F sequence data. Open circles indicate an intermediate haplotype state that was not found in the sample. Each line between the haplotypes represents a mutational step. The frequency of each haplotype is in parentheses.

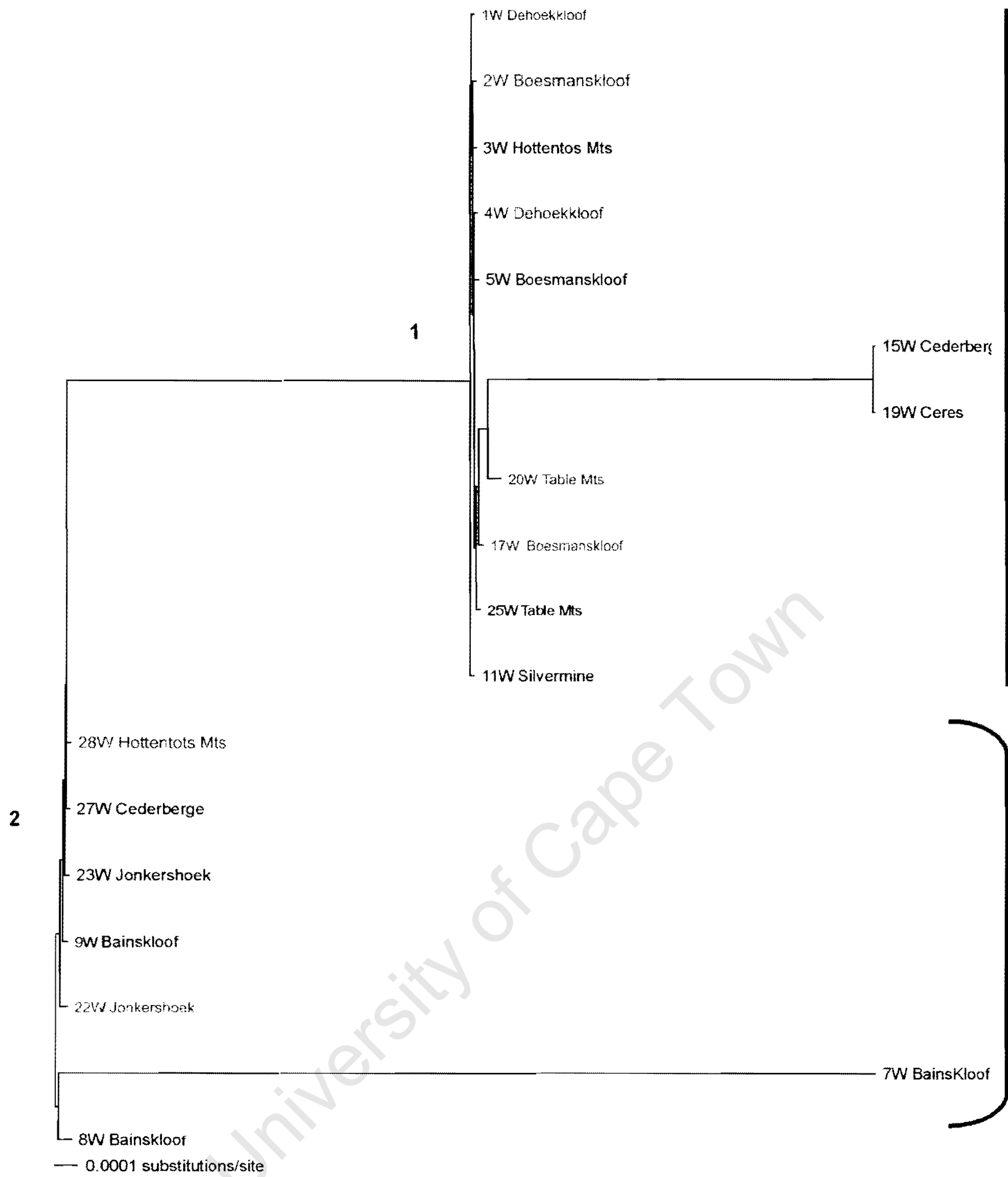
3.2 *Wardia hygrometrica*

3.2.1 Neighbour-joining analyses

Separate analysis of ITS data

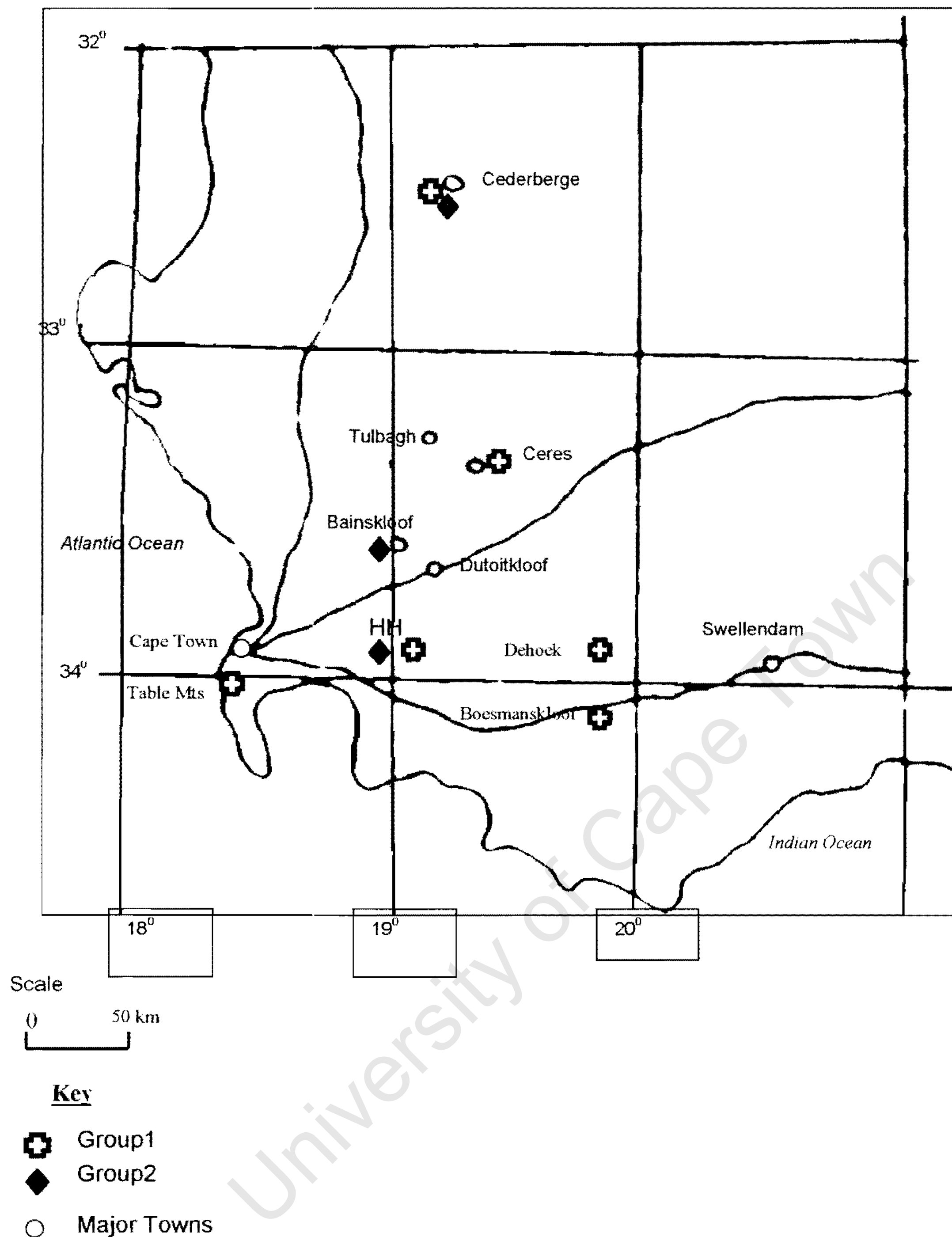
The neighbour joining tree recovered from the ITS data is depicted in Figure 11. The major feature that emerges from this tree is that it separates the 18 samples into two lineages. Lineage 1 was the largest consisting of samples from all the sampled areas except Jonkershoek and Bains Kloof. Within group 1, there are two samples from Cederberge and Ceres, which are highly divergent (0.002). The same applies to group two in which there is a sample from Bainskloof which alone is highly divergent. These samples are located in the North and North East of the Western Cape region. The distribution of the most prominent two groups is given in Figure12. Group 1 occurs in a north/south succession and the more eastern group (2) are restricted to single geographical areas. The two groups have high specificity to different geographic regions and thus can be regarded as phylogeographic groups. Of particular interest is the presence of highly divergent lineages in particularly the two from Bainskloof (group 2) with a genetic distance of 0.007. Results from the ITS analysis show that there is significant population structuring in *W. hygrometrica*.

University of Cape Town



University of Cape Town

Figure11. ITS Neighbour-joining tree using maximum likelihood distances of *W. hygrometrica*. Numbers correspond to sample numbers in Table 1.



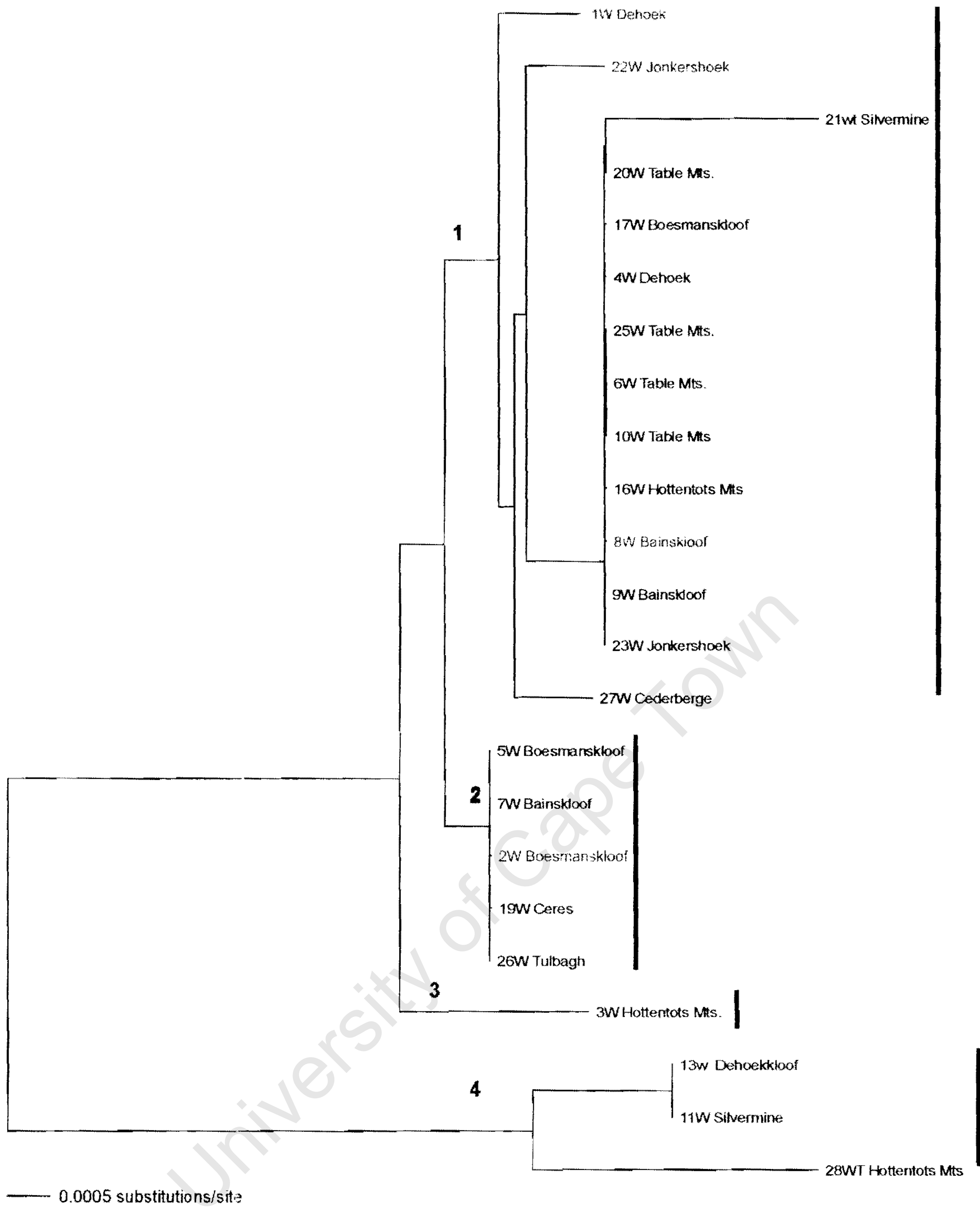
University of Cape Town

Figure12. Geographical distribution of the ITS lineages in *W. hygrometrica* . Lineage 1 was wide spread. HH refers to Hottetonts Holland Mountains. TM refers to Table Mountains.

Separate analysis of trnL-F data

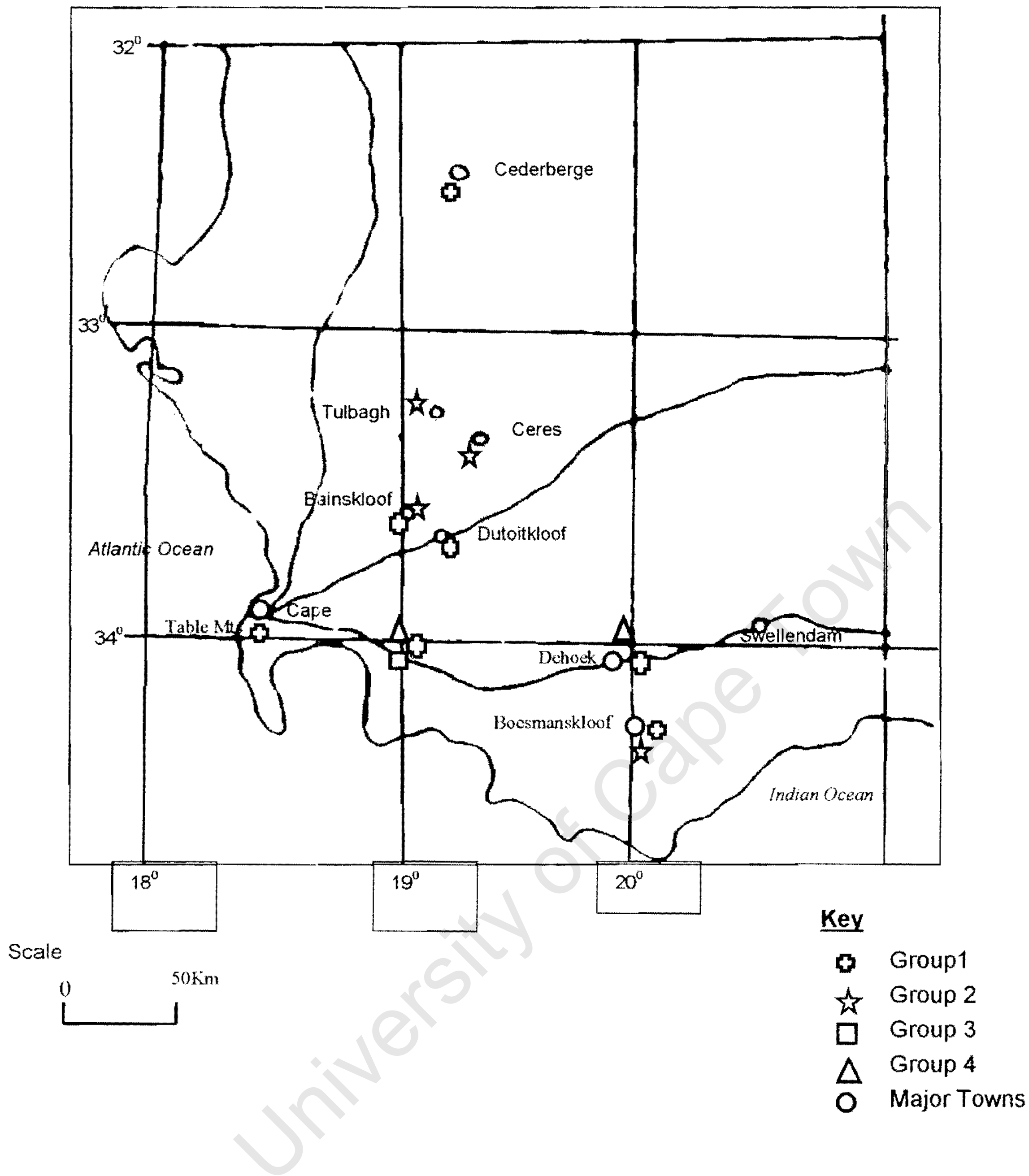
Figure 13 shows a neighbour-joining tree for the trnL-F data set. Twenty-three trnL-F sequences revealed four major lineages (1, 2, 3, and 4). The geographic distribution and location of all the clusters is presented in Figure 14. The most widely distributed group (1) is found in seven geographically distant areas. Group 2 is found in only four of the sampled areas and its distribution overlaps with that of the major group (1). This lineage is located Northern and Eastern areas of the Western cape region. Group 3 consists of only one sample and is unique to the Hottentots Holland Mountains. The specimen was very divergent from the rest (d= 0.02). Group 4 is distributed in the southern part of the western Cape region and its distribution overlaps with that of group 1. In this data set, there is an overlap in the distribution of the lineages compared with the ITS data set. However, the genetic distinctiveness within and among populations is emphasized in this data set as well as in the ITS data.

University of Cape Town



University of Cape Town

Figure13. trnL-F Neighbour-joining tree using maximum likelihood distances of *W. hygrometrica*. Sample numbers correspond to sample numbers in Table 1.



University of Cape Town

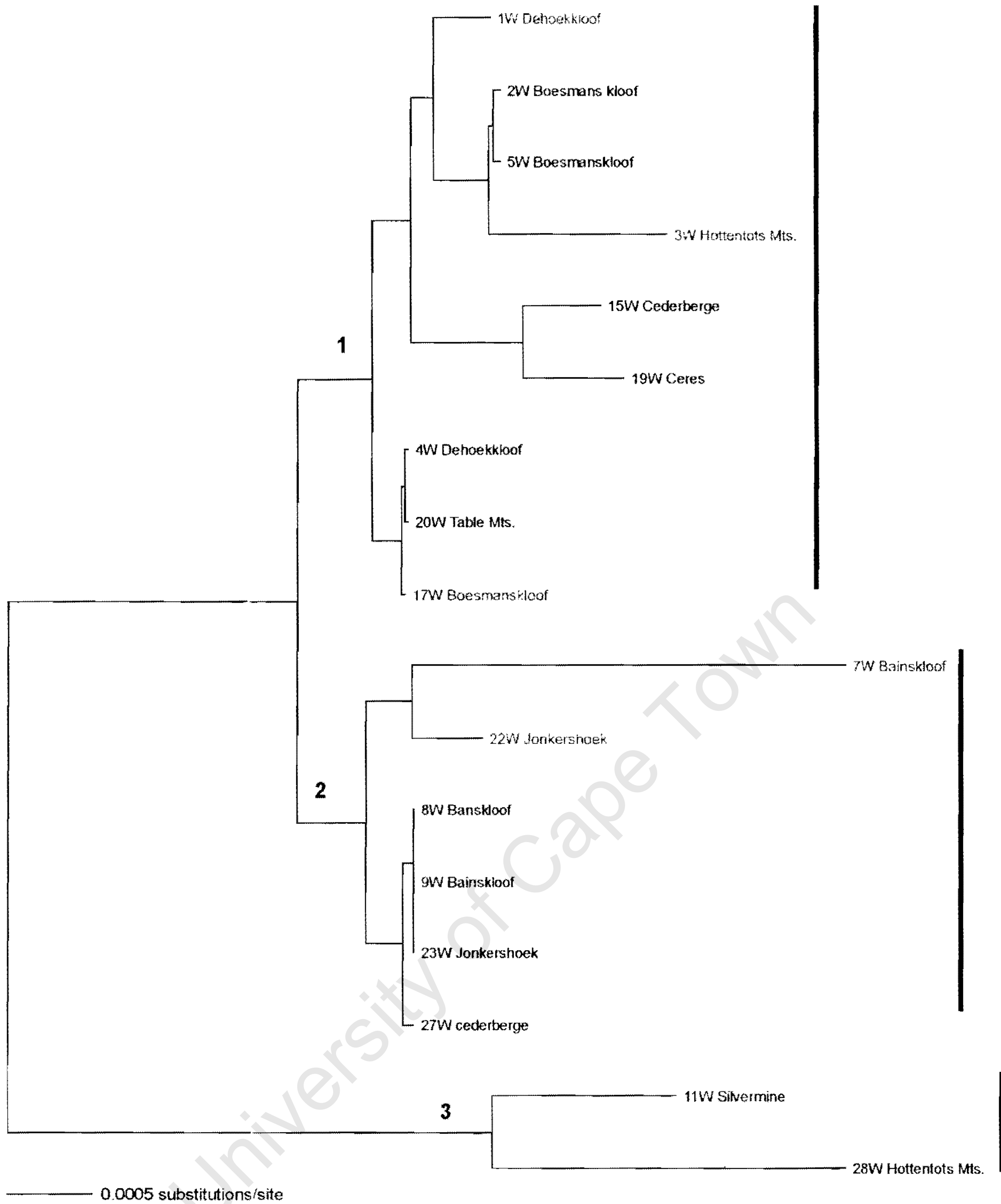
Figure 14. Geographical distribution of the trnL-F lineages in *W. hygrometrica* . Lineage 1 was widespread. HH refers to Hottetonts Holland Mountains. TM stands for Table Mts.

Combined (ITS+trnL-F)

The neighbour-joining tree recovered from this data is depicted in Figure 15. The major feature that emerges from this NJ tree is that it separates the 17 samples into three groups (1, 2, & 3). The geographic distribution of the three major clusters is shown on Figure 16.

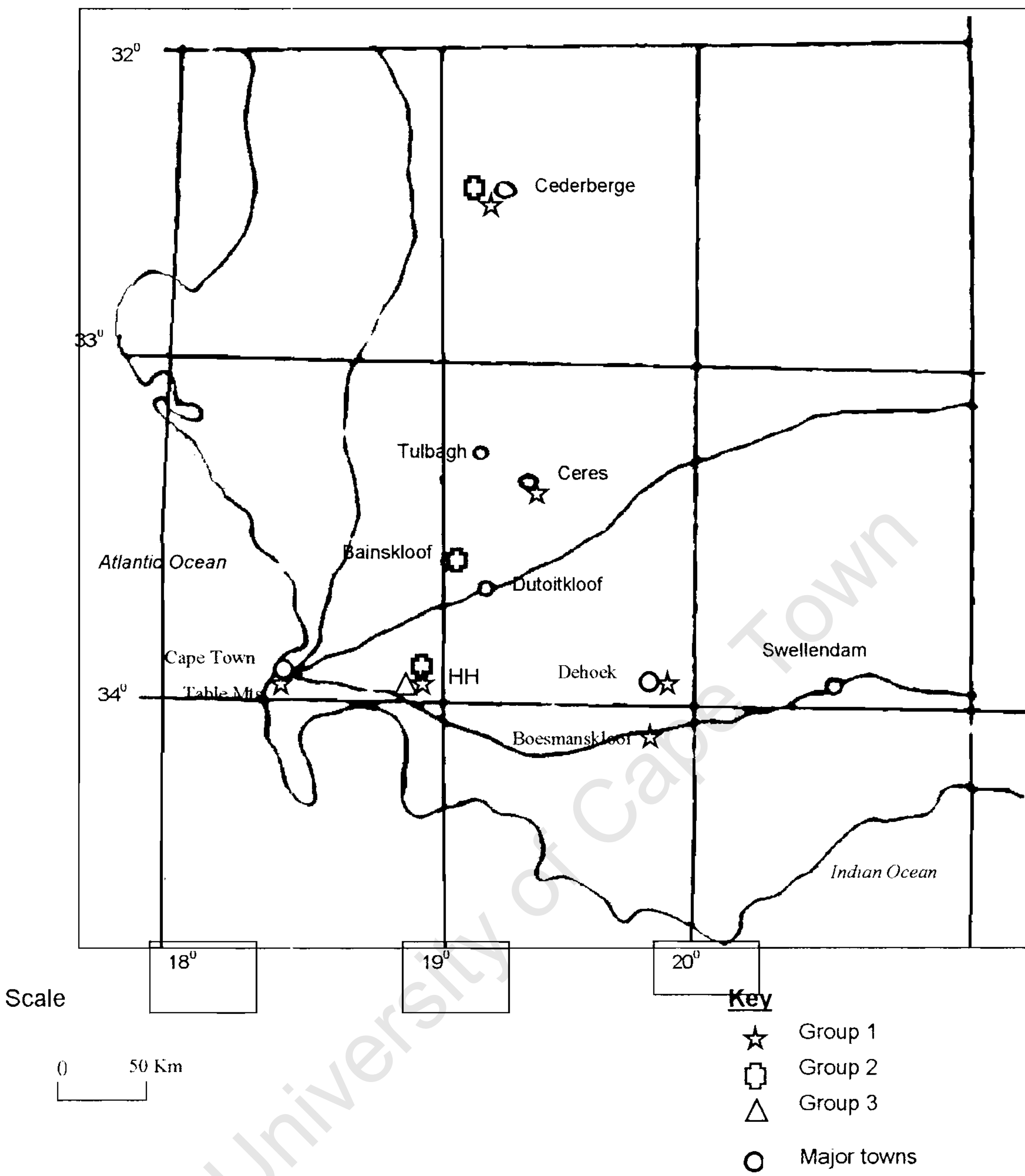
Group 1 was the largest consisting of samples from all the sampled areas except Jonkershoek. Group 2 is a more North-Eastern group and is restricted to this particular geographical area. Group 3 is also unique to the Hottentots Holland and Silvermine area. Although the three lineages overlap, the genetic distinctiveness between and among populations is still evident in this larger data set in that it is able to give some distinction within and among samples.

University of Cape Town



University of Cape Town

Figure 15. Combined ITS+trnL-F Neighbour-joining tree using maximum likelihood distances of *W. hygrometrica*. Sample numbers correspond to sample numbers in Table 1.



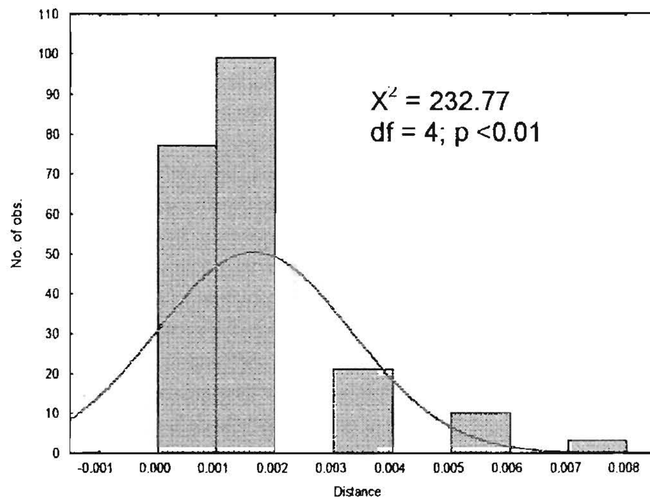
University of Cape Town

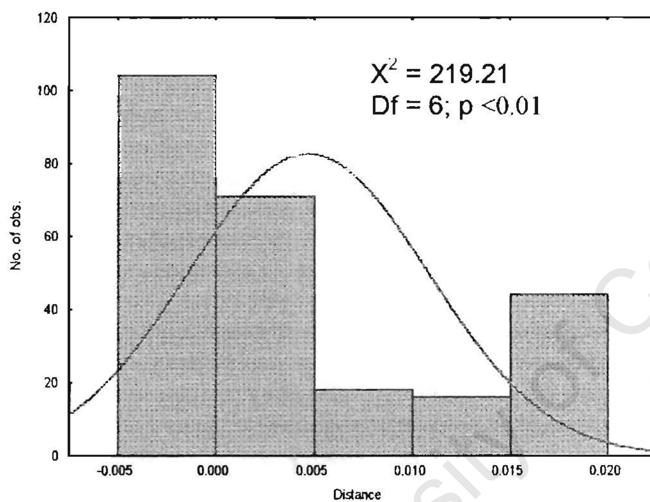
Figure 16. Geographical distribution of the combined ITS+trnL-F lineages in *W. hygrometrica*. Lineage 1 was wide spread. HH refers to Hottetonts Holland Mountains. TM refers to Table Mts

3.2.2 Population demography

Figs. 17a, b, & c show the mismatch distributions of ITS, trnL-F, and the combined data of *Wardia hygrometrica*. All the three graphs are not smooth and unimodal, but rather bimodal hence ruling out the possibility of demographic expansions in the remote past. The p-values, which test the discrepancy of the observed data from respective theoretical distributions, also show significant differences between observed and expected frequencies ($p<0.01$ in data partitions). For *Wardia*, such significant X^2 values are therefore, an indicator of no bottleneck in the remote past.

University of Cape Town





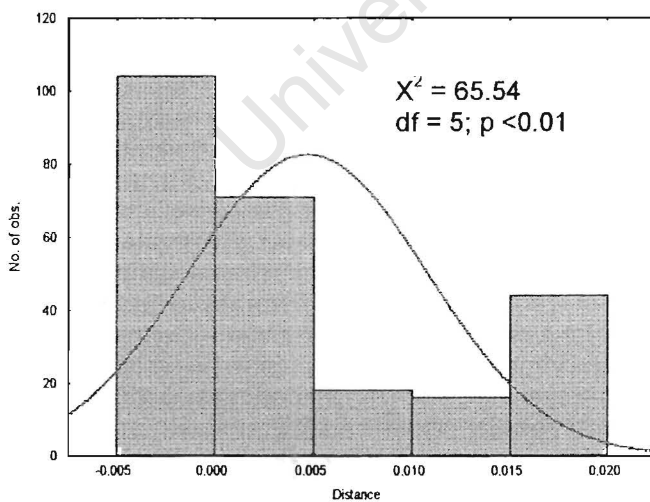


Figure 17. Mismatch distribution per population of *W. hygrometrica*. a" , ITS data, "b", trnL-F "c" , combined ITS+trnL-F. Departures from the expected value were tested with X^2 goodness of fit

University of Cape Town

3.2.3 Haplotype distribution in *W. hygrometrica*

After determining the nucleotide sequences of both the ITS and trnL-F gene regions for 21 and 23 individuals respectively, eight haplotypes were resolved from the ITS data and 13 haplotypes from the trnL-F data. Pair wise comparisons of the ITS and trnL-F sequences using the HKY model of sequence evolution yielded divergences ranging from 0.00-0.0057 among individuals for ITS data and 0.00-0.018 for the trnL-F data.

ITS haplotype distribution

The TCS analysis resolved eight haplotypes for the ITS dataset (Fig. 18). This haplotype network is star-like having the central most common haplotype (HW4) observed in nine individuals. Haplotype HW4 is the commonest haplotype and is separated from HW2 by a single nucleotide substitution. The pattern shown by ITS data involves common lineages that are widespread plus closely related lineages that are private (each confined to one or a few nearby localities). This phylogeographic outcome intimates low or modest contemporary gene flow between populations that are connected tightly in history (Avise 2000).

University of Cape Town

The second most common haplotype (HW2) was observed in six individuals from four different areas in the Cape. This haplotype was observed in samples from the Cederberg to Hottentots Holland Mountains. Five mutations separate haplotype HW2 from HW1 (from Bainskloof), and haplotype HW4 is separated from haplotype HW7 by one mutational step (Fig. 18). Haplotype HW7 (from the

Cederberg) is separated from haplotype HW8 (from Ceres) by five mutational steps. Two of the haplotypes are common (HW2 and HW4) and widespread in most population samples while the remaining six were rare and confined to one samples. Samples predominating in haplotype HW4, were from Boemanskloof (Table 6). Samples from Jonkershoek dominated in haplotype HW2. Silvermine had one unique haplotype (HW5). Other less common haplotypes (HW1, HW3, HW7, & HW8) were scattered in populations throughout the Western Cape region.

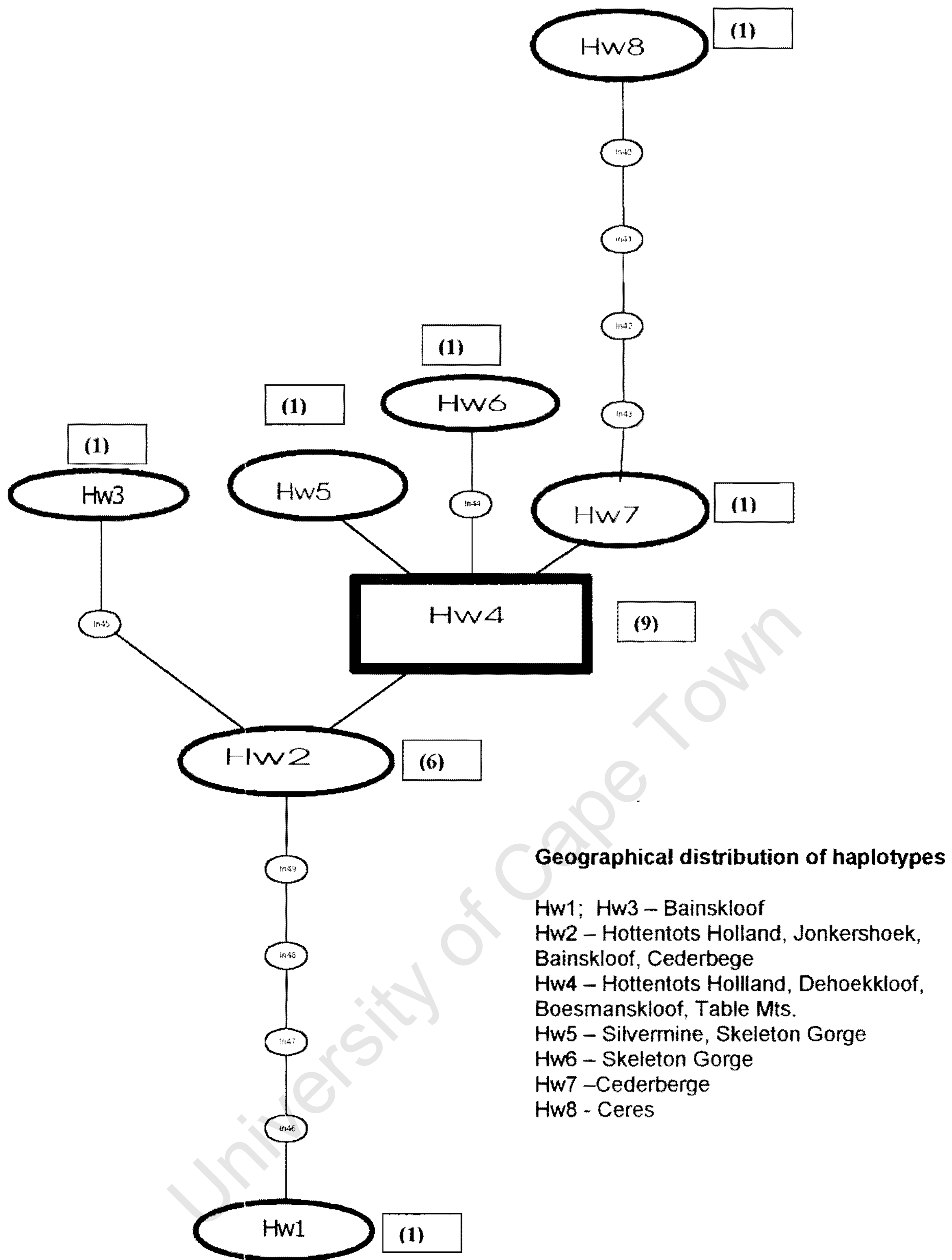
The haplotype distribution in the ITS data shows a pattern of increasing genetic divergence as one moves to the north and east of the region. For example, samples from Ceres and Cederberg were separated from the rest by an average of five mutational steps. Like the pattern in the NJ tree, this network is also characterised by regionally unique halotypes. There is also a high degree of structuring within populations evidenced by the presence of more than one haplotype per area

The two most common haplotypes representing the area of occurrence of the two widespread haplotypes have high specificity to different geographic regions and hence can be regarded as the major phylogeographic groups. Although haplotypes HW2 and HW4 were restricted to different geographic regions, they did overlap in the Hottentots Holland area. This indicates that the Hottentots Holland area is a zone of contact for the two most widely spread haplotypes.

University of Cape Town

The most common haplotype (HW4) in this case is the hypothesized ancestral haplotype (plesiomorphic) based on the following lines of evidence; a) it is by far the most common haplotype, occurring in 9 of the 21 assayed specimens; b) it was geographically widespread, observed in 4 of the 12 surveyed areas; c) it formed a core of a star like phylogeny whose rays (branches) connected separately to seven other nuclear DNA haplotypes.

University of Cape Town



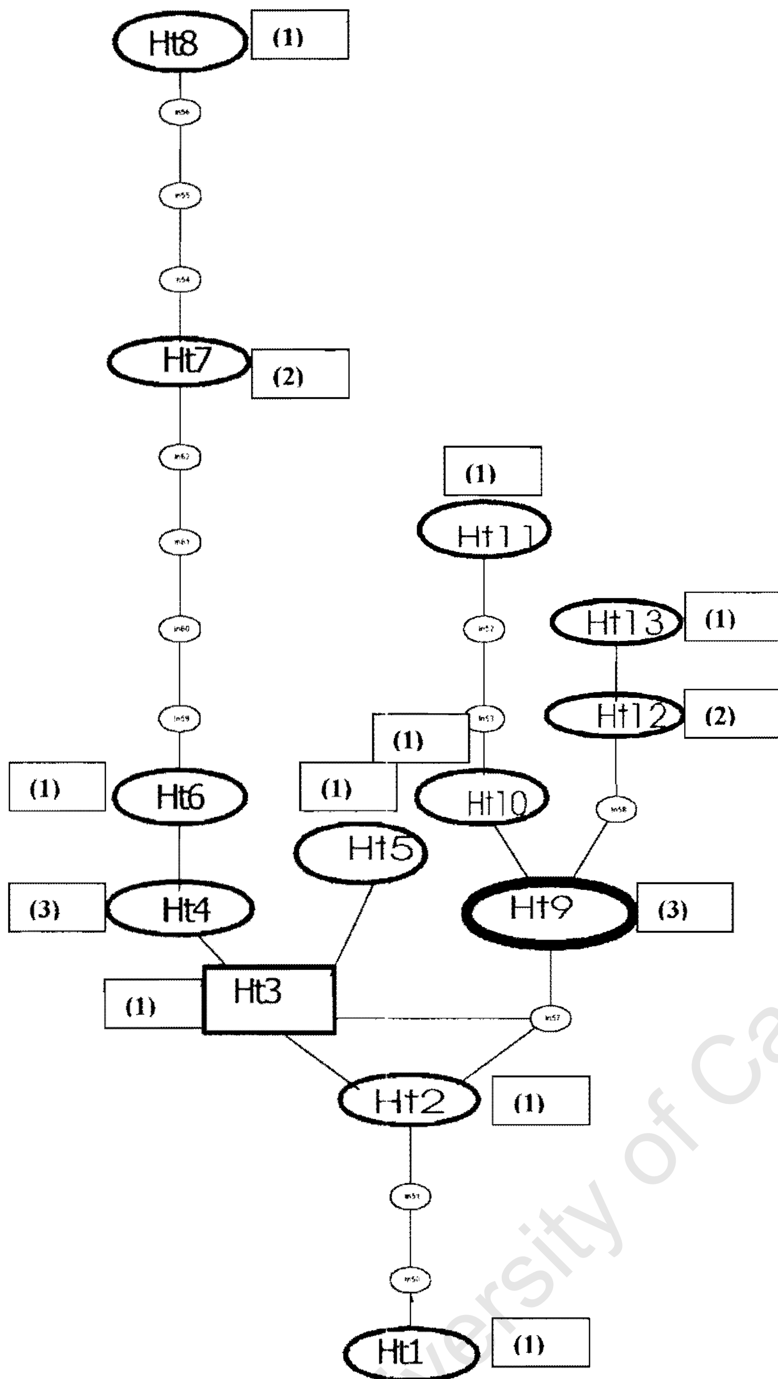
University of Cape Town

Figure 18. TCS haplotype network for *W. hygrometrica* based on ITS sequence data. Open circles indicate an intermediate haplotype state that was not found inthe sample. Each line between the haplotypes represents a mutational step. The frequency of each haplotype is in parentheses.

TrnL-F haplotype distribution

The trnL-F sequence data from 23 individuals was resolved into 13 different haplotypes (Fig. 19). Haplotype Ht4 is the most widespread haplotype occurring in most populations from Tulbagh to Boemanskloof while the remaining twelve were rare and confined to one to three individual samples. Ht7 was observed in single specimens from Table Mountain and Dehoek. Ht7 occurred disjunct□□□ly in populations from Dehoek and Bainskloof. Ht4 is therefore the hypothesized ancestral haplotype in this case. On Table Mountain alone, there were five unique haplotypes (Ht8, Ht10, Ht11, Ht12, Ht13). The remaining haplotypes were mostly confined to one, two or three samples. When considered in a geographical context, the most common haplotype (Ht4), was shared among populations found on the eastern side of the Table Mountain. This result is consistent with results obtained from neighbour joining analyses as well as ITS haplotype network i.e they all show genetic structure within and among populations and as well emphasizes the uniqueness of the populations found in the eastern and northern areas of the Western Cape region.

University of Cape Town



Geographical distribution of haplotypes
Ht1 - Jonkershoek

Ht8, Ht10, Ht11, Ht12, Ht13 – Table Mts.
Ht2 – Houtbais
Ht3 – Ceres
Ht4 – Bainskloof, Boesmanskloof
Ht5 – Hottentots Holland Mts.
Ht6 – Boesmanskloof
Ht7 – Dehoekkloof, Silvermine
Ht9 – Dehoekkloof, Bainskloof

Figure 19. TCS haplotype network for *W. hygrometrica* based on trnL-F sequence data. Open circles indicate an intermediate haplotype state that was not found in the sample. Each line between the haplotypes represents a mutational step. The frequency of each haplotype is in parentheses. The heavy surround line for haplotype Ht9 signifies a haplotype with the highest frequency of samples.

University of Cape Town

CHAPTER 4. GENERAL DISCUSSION AND CONCLUSIONS

University of Cape Town

Assessment of the degree to which codistributed taxa have congruent phylogeographies is critical because, in instances where congruence is detected, shared history has led to similar partitioning of genetic variation in the taxa being compared. This enables the prediction of geographic patterns of genetic variation for other codistributed taxa that have yet to be examined (Sullivan et al. 2000).

In southern Africa, in plant studies, as in studies with other organisms, much research has focused on phylogenetic studies of higher-level taxa. In essence, phylogeographic studies are lacking. The present study is the first phylogeographic study on plants in southern Africa; hence, it will form a benchmark for all of the future studies in this field.

The haplotype information from this study reveals a complex pattern among forest dwelling mosses in Southern Africa and suggests that different evolutionary factors have been important within each of the species examined. *L. smithii* and *W. hygrometrica* showed genetic differentiation in both the ITS and trnL-F regions whereas *P. gracile* showed no variation in either region. This leads to the following question: why do these forest species have different patterns and degrees of nuclear and chloroplast DNA differentiation when they seem to share the same ecological niche, and have closely corresponding distributions in southern Africa?

University of Cape Town

4.1 Major lineages in *L. smithii*

Neighbour-joining trees and haplotype networks were used to infer relationships among populations of *L. smithii* specimens from both Europe and Africa. Neighbour-joining trees inferred from, trnL-F and combined ITS+trnL-F agree with the basic division of *L. smithii* into European and African lineages (Figs. 5 & 6). The genetic distinctiveness of the European samples was evident in all the three data sets. Furthermore, the erratic mismatch distributions (Fig. 8), when statistically tested, indicate stable populations, rather than expanding ones. Thus regardless of the time this species has been resident in the forest patches, enough time has elapsed for its populations to reach equilibrium between genetic drift and migration.

The TCS haplotype networks of both the ITS and trnL-F data recognised 14 and 12 haplotypes respectively (Figs. 9 & 10). The most common haplotypes inferred from the two data sets were very widespread in Africa and were shared between western Cape, Natal, and Kenya. Such a pattern might reflect a recent common history or ongoing gene flow between the studied populations. There was also sharing of trnL-F haplotypes between Madeira and the Portuguese mainland suggesting possible gene flow between the two regions. Furthermore, results from the haplotype analyses emphasize the genetic distinctiveness of the African and European lineages as evidenced by lack of haplotype sharing between the two regions. In the hierarchical analysis of *L. smithii* populations, the greatest part of variation was found among regions. Meanwhile, a marked structure was

University of Cape Town

also depicted among forest patches. On population genetic grounds, these results suggest that genetic drift is effective at all levels and genetic exchange is limited (Arctander et al. 1999).

4.2 Phylogeographic pattern in *L. smithii*

In this study both the ITS and trnL-F data sets exhibit discontinuous genetic divergence patterns with spatial separation with regard to specimens from Europe and Africa. This is a very common phylogeographic pattern in both plants and animals and is usually a result of long-term extrinsic barriers to gene flow but can also be due to extinction of intermediate haplotypes especially in species with wide distributions but limited gene flow (Avise 2000). If we consider the phylogeographic patterns within each continent, we encounter a very different pattern i.e. a continuous genetic divergence pattern with moderate levels of gene flow characterised by sharing of haplotypes and fewer mutations between lineages.

Caution should however be exercised when one is interpreting such results from a dataset with inadequate sampling. In molecular surveys, the power to detect statistically significant lineage structure increases with larger numbers of individuals surveyed, all else being equal. (Avise 2000). Thus, the same species that bear a molecular footprint of moderate to high gene flow might convert to an appearance of low or intermediate gene flow under more stringent sampling conditions.

University of Cape Town

4.3 Major lineages in *Wardia hygrometrica*

The ITS region in *W. hygrometrica* was less variable than trnL-F gene region suggesting a potential for more resolving power in the chloroplast gene. In the NJ analysis, the ITS data divided the 18 *Wardia* samples into two lineages (Fig. 11). The major lineage (1) consisted of samples from most of the sampled ravines and was very widespread. However, lineages occurring on the northern and eastern part of the Western Cape region were more divergent than the rest. This could be a result of low gene flow among forest patches or a completion of lineage sorting.

The high genetic structure among populations and the genetic distinctiveness of samples from the north and east western Cape, remains evident in the trn and the combined data sets (Figs. 13, 14, 15, &16) although the distinction is not clear-cut. The trnL-F recognised four major groups (Fig. 13). The biggest group "1" was very common and widespread. The combined data set recognised three major lineages (Fig. 13) with group 1, which was widely distributed. The combined tree was more resolved than either of separate analyses. The tree also preserves some of the geographic groups depicted in the combined ITS+trnL-F data. Therefore, the joint analysis resulted in greater resolution in all the groups; moreover, the topologies of the combined tree boast a stronger correlation than the separate analyses.

University of Cape Town

The mismatch distribution (Fig. 17) of the three data sets when statistically tested by Chi-squared statistic for goodness of fit did not indicate demographic expansion. The haplotype network also revealed some geographic structuring in the two data sets. In the trnL-F haplotype network (Fig. 19) the Table mountain area did not share any haplotypes with the rest of the areas. This result agrees with the trnL-F NJ tree results, which clustered all the Table mountain specimens together. This shows that there is little or no genetic exchange between the Table Mountains ravines and populations from the rest of the species area. The sharing of alleles was also evident in the NJ and haplotype trees.

In this study, chloroplast and nuclear haplotype networks of *W. hygrometrica* reveal two main points. The first is that the two gene regions in this group are geographically structured despite partial overlap. The second point is that the separation of the haplotypes on a geographical basis in both ITS and trnL-F is still evident although the distinction is not clear-cut. Though only nine different river systems were analysed, these results are likely to be broadly representative of the species situation. In this study trnL-F was more structured than ITS.

This study has identified some aspects of the broad phylogeography of *W. hygrometrica*. One interesting finding concerns the overlap between common haplotypes suggesting shared ancestral polymorphisms that predated the divergence of populations.

University of Cape Town

4.4 Lack of variation in *P. gracile*

The lack of variation in *P. gracile* suggests that the species might have arrived very recently so it has not undergone any differentiation. Although isolated populations would be expected to differentiate via genetic drift, the lack of differentiation can also be explained in terms of possible ongoing gene flow between populations. The species may not be differeciating if predominantly asexual which is the case with most mosses

4.5 Phylogeographic hypotheses

Owing to the relevance of comparative phylogeography to ecological and historical biogeography, numerous studies have taken a lead from Avise in attempting to assess phylogeography on a regional scale (reviewed in Sullivan et al. 2000). Thus, with different packages for analysis of population level processes, phylogeographers now have the ability to study the history of populations.

In the current study, phylogeographic structuring in *L. smithii* and *W. hygrometrica* cannot be matched with that of *P. gracile* which showed no variation in both the ITS and trnL-F gene regions. Thus, contrary to the original prediction, it seems these similarly distributed moss species have not been affected in the same way by historical biogeographic features. Zink (1996) stated that such patterns usually result from species' differences in response to barriers or selective gradients, levels of gene flow, rates of molecular evolution, effective

University of Cape Town

population size, or generation time. That is, species might have been historically codistributed but variable responses to historical events produced conflicting phylogeographic patterns.

4.6 History and age of the forests in southern Africa

The lack of phylogeographic concordance in forest dwelling mosses could mean that the species arrived at different times at their current distribution. In the absence of fossil records, it is difficult to predict their actual time of arrival into their current habitat and consequently the age of the forests themselves.

The sharing of ITS and trnL-F haplotypes in *L. smithii* among widely distributed populations is consistent with the hypothesis of the existence of a continuous forest which existed in east Africa all the way down to Natal up to the Western Cape region. Fragmentation of this forest during the Pleistocene climatic changes would account for the shrinking of forests. Each forest patch in this regard has kept their ancestral haplotype, which was widely distributed before the Pleistocene events. These observations are in agreement with the link shown by higher plants such as *Protea* (Rourke 1980) and *Disa* (Linder 1983).

University of Cape Town

The ITS and trnL-F haplotype networks for *W. hygrometrica* reveal that all the haplotypes within western Cape are closely related, mostly differing from each other by a single or two mutations. The wide spread occurrence of HW4 and HW2 (ITS) and Ht4 (trnL-F) suggests that these haplotypes are the most

probable ancestral lineages within this species. In addition, the phylogeographic pattern shown by *W. hygrometrica* is very different from that of the two forest mosses (*L. smithii* & *P. gracile*) which is expected due to the differing in the ecologies of the three species. *W. hygrometrica* is restricted to streams within the kloofs that the forests occupy whereas the remaining two species are restricted to forests. Their unique ecologies might have led to their different responses to climatic changes.

4.7 Sequence variation in the ITS and trnL-F regions and their suitability for phylogeographic studies

In the present study, two molecular markers were used to examine the geographical relationships of three South African forest mosses. ITS and trnL-F sequences have been the most prevalent sources of sequence data in plants at lower taxonomic levels and have yielded good resolution in phylogenetic studies.

In the current study, the trnL-F marker was more divergent in *W. hygrometrica* and *L. smithii* compared with the ITS gene. This might be attributable to "greater discrimatory power inherent in assays of cytoplasmic inherited genomes (due to smaller effective population sizes for uniparentally inherited genes than nuclear genes). Therefore, the trnL-F region evolves faster in the studied mosses than the ITS gene region. However, in *P. gracile*, there was no genetic differentiation in the two gene regions.

University of Cape Town

4.8 Conclusions

The current geographical survey demonstrates the use of ITS and trnL-F genes to examine the phylogeography of plant species, in this case, forest dwelling mosses. The analyses in the present study are, however, subject to restriction of small sample sizes. Despite this, they represent the first comparative ITS and trnL-F based investigations in forest mosses. Essentially, the study used two forest mosses (*L. smithii* & *P. gracile*). *L. smithii* was variable in both regions while *P. gracile* was not. In order to determine if there is a general pattern in these forest species , there is need for more species to be assayed.

The current research has also shown a genetic discontinuity between African and European populations of *L. smithii*. However, in Africa, there is sharing of haplotypes between populations from western Cape, Natal and Kenya. This suggests possible gene flow between the southern African populations and a probable continuous existence of a continuous forest in the past. Furthermore, the study has shown genetic structure in *W. hygrometrica* a moss species restricted to the western Cape region.

University of Cape Town

4.8 Limitations of the study and future prospects

Sampling considerations

The successful application of phylogeographic estimation requires an appropriate sampling regime for the species being studied. This entails choosing sample locations across the entire region of interest at a density that reflects the natural

density of the subject species. The analyses in the present study are subject to the restriction of small sample sizes. This problem was inevitable because of the reliance on herbarium material which were too old hence could not amplify in most cases. The best option would have been to collect fresh samples but this activity was limited by the short period in which the study had to be conducted (six months).

Due to the absence of a good fossil record for mosses, there is little data for calibration of molecular clocks. Therefore, neither the actual rate of nuclear nor chloroplast DNA evolution nor the period of divergence could be determined for any of the species.

University of Cape Town

The future of phylogeography in southern Africa

Results from this study represent the first phylogeographic studies of plants in southern Africa. The current data indicate that investigating mosses would be very rewarding in terms of insights into the general biogeographic patterns of southern African plants, which are little understood at present. While there were some signs of geographic structuring in two of the studied mosses, the level of common history in forest mosses is not evident. What still remains in the species studied here is to widen the sampling, especially in *L. smithii*.

Further insight into the historical biogeography of southern African flora will require more extensive studies in both higher and lower plant taxa. The result will

be a mutual enrichment of southern African biogeographic history, knowledge of which is lacking at present.

University of Cape Town

References

ACOCKS, J.P.H. 1953. Veld types of South Africa, *Memoirs of the botanical survey of South Africa,*. 28: 1-192.

AKIYAMA, H. 1999. Genetic variation of the asexually reproducing moss, *Takakia lepidoziodides. J.Bryol*, 21:177-182.

AMOS, W. & J. HARWOOD. 1998. Factors affecting levels of genetic diversity in natural populations. *Philos. Trans. R. Soc. Lond. B* . 353: 177-186.

ARBOGAST, B.S. & J. KENAGY, 2001. Comparative phylogeography as an integrative approach in biogeography. Guest Editorial, *J.Biog.* , 28(7).

ARCTANDER, P., C. JOHANSEN, & M. COUTELLEC-VRETO. 1999. Phylogeography of three closely related African Bovids (Tribe Alcelaphini). *Molecular Biology & Evolution* 16(12):1724-1739.

ARNOLD, M.L. & S.K. EMMS. 1998. Molecular markers, gene flow, and natural selection. In SOLTIS, P.S., D.E. SOLTIS, & J.J. DOYLE (Eds). Molecular systematics of plants, pp. 117-150. Chapman & Hall. New York.

AVISE, J.C. 1998. Conservation genetics in the marine realm. *J. of Heredity*, 89:377-382.

AVISE, J.C. 2000. Phylogeography. The history of species. Havard University Press. Cambridge, Massachusetts.

BAKKER, F.T.& J.L. OLSEN, W.T. STAM. 1995b. Global phylogeography in the cosmopolitan species *Cladophora vagabunda* (Chlorophyta) based on nuclear rRNA ITS sequences. *Eur. J. Phycol,* 30: 197- 209.

University of Cape Town

BALDWIN, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molecular Phylog and Evol.,* 1: 3-16

BALDWIN, B.G. 1993. Molecular phylogenetics of Calycadenia (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evidence re-examined. *Am, J. of Bot.,* 80: 222-238.

BALDWIN, B.G., M.J. SANDERSON, J.M. PORTER, M.F. WOJCIECHOWSSKI, C.S. CAMPBEL, & M.J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA; a valuable source of evidence on angiosperm phylogeny. *Ann.Missouri Bot. Gard,* 82: 247-277.

BALL, I.R. 1976. Nature and formulation of biogeographical hypotheses. *Syst. Zool.,* 24: 407-430.

BRITTEN, R.J., & D.E. KOEHN. 1968. Repeated sequences in DNA. *Science,* 161: 529-540.

BROWER, A.V.Z., R. DEDALLE, & A. VOGLER. 1996. Gene trees, species trees, and systematics: a cladistic perspective. *Annu. Rev. Ecol. Syst.*, 27: 423-450.

CLEMENT, M. D. POSADA & K.A. CRANDALL. 2000. TCS: A computer programme to estimate gene genealogies. *Mol.r Ecol.*, 9(10): 1657-1660.

COVE, D. 2000. Molecular genetic studies of moss species. In SHAW, A.J. & B. GOFFINET. Bryophyte biology. University Press, Cambridge.

CRANDAL, K.A., & A.R.TEMPLETON. 1996. Application of intraspecific phylogenetics. Pp 81-99. In HARVEY P.H., A.J. LEIGH BROWN, J.

University of Cape Town

MAYNARD SMITH, & S. NEE, (eds). New uses for new phylogenies. Oxford University Press, New York.

CRANDAL, K.A.1994. Intraspecific cladogram estimation: accuracy at higher level of divergence. *Syst. Biol,*. 43:222-235.

CRUZAN, M.B. 1998. Genetic markers in plant evolutionary ecology. *Ecology*,79:400-412.

DE QUEIROZ, A. 1993. For consensus (sometimes). *Syst. Biol,* 42:368-372.

DONOGHUE , M.J. & M.J. SANDERSON. 1992. The suitability of molecular and morphological evidence in reconstructing plant phylogeny. Pp. 340-368. In SOLTIS, P.S., SOLTIS D.E. & DOYLE, J.J. (eds).Molecular Systematics of plants, Chapman & Hall.

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

FUJI, N., K. UEDA, Y. WATANO, & T. SHIMUZI. 1995. Intraspecific sequence variation in chloroplast DNA of *Primula cuneifolia* Ledeb. *J. Phytogeogr. Taxon.*, 43: 15-24.

GAWEL, N.J. & R.L JARRET. 1991. A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea. Plant Mol. Biol. Reports,* 9: 262-266.

GOLDEN, J.L. & J.F. BAIN. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera (Asteraceae)* species in southern Alberta. *Evol.,* 54(5): 1566-1579.

University of Cape Town

HAMBY, R.K., L. SIMS, L. ISSEL & E. ZIMMER. 1988. Direct ribosomal RNA sequencing: optimisation of extraction and sequencing methods for work with higher plants. *Pl. Mol. Biol. Rep., 6*: 175-192.

HARDING, R.M. 1996. New phylogenies: Pp. 15-22. an introductory look at coalescent. In New uses for new phylogenies, (ed. BROWN, E.J., P.H. HARVEY, J.M. SMITH, & S. NEE). Oxford University Press, New York.

HARE, M.P. 2001. Prospects for nuclear gene phylogeography. *Trends in Ecol. & Evol.,* 16 (12): 700-706.

HARNER, J.L. 1991. The meaning of rarity. Cited in LONGTON, R.E. & T.A. HEDDERSON. 2000. What are rare species and why conserve them? *Lindbergia,S* 25(2-3): 53-61.

HASEGAWA, M., H. KISHINO, & T. YANO. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. of Mol.Evol.,,* 22: 160-174.

HAUK, W.D., C.R. PARKS, & M.W. CHASE. 1996. A comparison between trnL-F intergenic spacer and rbcL DNA sequence data: an example form Ophioglosssaceae. *American Journal of Botany* (suppl.). 83:126.

HEDDERSON, T.A., C.J. COX, & J.G. GIBBINGS. 1999. Phylogenetic relationships of Wardiaceae (Musci): Evidence from 18s rRNA and rps4 gene sequences. *The Bryologist,* 102(1): 26-31.

HEWITT, G.M. 1999. Post-glacial re-colonization of Eouropean biota. *Biological J. of the Linn. Soc.,* 68: 87-112.

University of Cape Town

KELCHNER, S.A. 2000. The evolution of Non-coding chloroplast DNA and its application in plant systematics. *Ann. Missouri. Bot. Gard.*, 87: 482-498.

KIMURA, M. & T. OHTA. 1972. On the stochastic model for estimation of mutational distance between homologous proteins. *J. of Mol.Evol.*,, 2:87-90.

KIMURA, M. 1968) Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genetic Research*, 11:247-269.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. of Mol. Evol.,* 16: 111-120.

KIMURA, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

KVIST, L. 2000. Phylogeny and Phylogeography of European Parids. Acta Univ. Oulu. http://herkules.oulu.fi/isbn9514255364/

LEWONTIN, R.C., 1991. Twenty-five years ago in genetics electrophoresis in the development of evolutionary genetics: milestone or millstone? *Genetics,* 128: 657-662.

LINDER, H.P. 1983. The historical phytogeography of Disnae (Orchidaceae). *Bothalia* 14: 565-570.

LINDER, H.P. M.E. MEADOWS & R.M. COWLING. 1992. History of the Cape flora. In COWLING, R.M. ED. The ecology of fynbos: Nutrients, fire and diversity. Oxford University Press, Oxford.

University of Cape Town

LISTON, A. & J.W. KADEREIT. Chloroplast DNA evidence for introgression and long distance dispersal in the desert annual *Senecio flavus* (Asteraceae). *Pl. Syst. and Evol.*. In press

LISTON, A. L.H. RIESEBERG, & M.A. HANSON. 1992. Geographic partitioning of chloroplast DNA variation in the genus *Datisca (Datiscaceae)*. *Pl. Syst. Evol.*, 181:121-132.

LONGTON, R.E. & T.A. HEDDERSON. 2000. What are rare species and why conserve them? *Lindbergia,* 25(2-3): 53-61.

MAGILL, R.E. & J. van Rooy (1998). Bryophyta. Part 1 Musci. Fascicle 3: Erpodiaceae – Hookeriaceae. National Botanical Institute, South Africa

MAGILL, R.E. 1981. Flora of southern Africa. Bryophyta, Part 1. Mosses. Botanical Research Institute, Department of Agriculture and fisheries, Republic of South Africa.

MANOS, P.S. 1997. Systematics of *Nothofagus* (Northofagaceae) based on rDNA spacer sequencies (ITS): taxonomic congruence with morphology and plastid sequences. *Amer. J. Bot.,* 84: 1137-1155

MEADOWS, M.E. & H.P. LINDER. 1993. A palaoecological perspective on the origin of Afromontane grasslands. *Journal of Biogeography.* 20:345-355.

MIDGLEY, J.J., R.M. COWLING, A.H.W. SEYDACK, & G.F. van WYK. 1997. pp. 278-299 In COWLING, R.M., D.M. RICHARDSON, & S.M. PIERCE (Eds). Vegetation of Southern Africa,. Cambridge University Press, Cambridge.

University of Cape Town

MISHLER, B.D. 1997. Monographic research in the moss family Calymperaceae. University Herbarium, University of California Berkeley. http://ucjeps.berkeley.edu/bryolab/PEET/index.html.

MITTON, J.B. 1994. Molecular approaches to population biology. *Annual Rev. Ecol. Syst,*25: 45-69

MIYAMOTO, M.M. 1985. Consesus cladogram and general classification. *Syst. Biol.,* 44: 64-76

MMATTAJ, J.W., D. TOLLERVEY, & B. SERAPHIN. 1993. Small nuclear RNAs in messenger RNA and ribosomal RNA processing. *FASEB Journal,* 7:47-53.

MYERS, A.A. & P.S. GILLER. 1988. Analytical biogeography. An intergarated approach to the study of animal and plant distributions. Chapman & Hall, London.

NEWTON, A.E., C.J. COX, J.G. DUCKETT, J.A. WHEELER, B. GOFFINET, T.A. HEDDERSON, & B.D. MISHLER. 2000. Evolution of the major moss lineages: Phylogenetic analyses based on multiple gene sequences and morphology. *The Bryologists,*103(2): 187-121.

OLSEN, K.M. & B.A. SCHAAL .. 1999. Evidence of the origin of cassava: Phylogeography of *Manihot esculenta. Proc. Natl. Acad.Sci. USA.*, 96: 55 86-5591.

PALMER, J.D. 1987. Chloroplast DNA evolution and biosystematics uses of chloroplast DNA variation. *Amer. Nat.,* 130 (Suppl.): S6-S29.

University of Cape Town

PALMER, J.D. 1991. Plastid chromosomes: structure and evolution. Pp 5-53 In BOGORAD, L. & I.K. VASIL (eds.), Cell Culture and Somatic Cell Genetics of plants, Vol. 7A: The Molecular Biology of plastids. Academic Press. Orlando.POSADA, D. & K.A. CRANDAL. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecol. & Evol.*, 16:37-45

RIESEBERGE, L.H., & S.J. BRUNSFELD. 1992. Molecular evidence and plant introgression. Pp.151-176. In Molecular Plant Systematics, Soltis, P.S. D.E. Soltis, and J.D. Doyle. (eds.). Chapman and Hall, New York.

ROGERS, A.R. & H. HARPENDING. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. and Evol.,* 9: 552-569

ROSEN, D.E. 1975. A vicariance of Caribbean biogeography. *Syst. Zool.,* 24: 431-464.

ROURKE, J.P. 1980. The proteas of southern Africa. Purnell, Cape Town.

SAITOU, N. & M. NEI. 1987. Neighbour-joining method. *Mol. Biol. & Evol.,* 4: 406-425.

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

SCHAAL, B.A. & K.M. Olsen. 2000. Gene genealogies and population variation in plants. *Proc. Natl. Acad. Sci.* USA, 97:13: 7024-7029.

University of Cape Town

SCHL□TTERE, C., M.T. HAUSER, A. VON HAESELER, & D. TAUZ. 1994. Comparative evolutionary analysis of rDNA ITS regions in the Drosophilla. *Mol. Biol. & Evol.*, 11:513-522.

SCOTT, H.M. H.M. ANDERSON & J.M. ANDERSON. 1997. Vegetation history. Pp. 278-299. In COWLING, R.M., D.M. RICHARDSON & S.M. PIERCE (Eds). Vegetation of Southern Africa,. Cambridge University Press, UK.

SHACKLETON, N.J. & J.P. KENNET. 1995. Palaeotemperature history of the Cenozoic and the initiation of the Antactic glaciation: oxygen and carbon isotope analyses in DSDP sites 277, 279 and 281. *Initttial reports of the DSDP,* 29:743-755.

SHAW, A.J. 1995. Genetic biogeography of the rare "copper moss, *Scopelophila cataractae* (Pottiaceae). *Pl. Syst. Evol.*, 197: 43-58.

SIM, T.R. 1918. Geographical distribution of the South African Bryophyta. South African Journal of Science.

SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution,* 47: 264-279.

SOLTIS, P.S. & S.J. NOVAK, 1997. Intraspecific cpDNA variation in four widespread species of *Lomatium* (Umbelliferae). *Syst. Bot* 22(1): 99-112

SOLTIS, D. E. & P. S. SOLTIS. 2000. Contributions of molecular systematics to studies of molecular evolution. *Pl. Mol.r Biol.,* 42: 45-75.

SOLTIS, D.E., & R.K. KUZOFF. 1995. Discordance between nuclear and chloroplast phylogenies in the Heuchera group (Saxifragaceae). *Evolution,* 49: 727-742.

University of Cape Town

SOLTIS, D.E., P.S. Soltis & B.G. Millgan. 1992. Intraspecific chloroplast variation: Systematic and phylogenetic implications, Pp. 117-150. In P.S. Soltis, D.E. Soltis & J.J. Doyle (eds.). Plant Molecular Systematics. Chapman & Hall, New York.

SOLTIS, D.E., P.S. SOLTIS, & B.D. NESS. 1989a. Chloroplast DNA variation and multiple origins of autoploidy in *Heuchera micrantha* (Saxifragaceae). *Evolution,* 43 : 650-656.

SOLTIS, D.E., P.S. SOLTIS, 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. Pp. 1-42. In SOLTIS, P.S. D.E. SOLTIS & J.J. DOYLE (Eds). Moleculer Systematics of plants II, . Chapman & Hall. New York.

SOLTIS, D.E., P.S. SOLTIS, T.G. COLLIER & M.L. EDGERTON. 1991. Chloroplast variation within and among genera of the Heuchera group (Saxifragaceace): Evidence for chloroplast transfer and paraphyly. *Am. J. Bot.,* 79: 1150-1161.

STATISTICA 5.5. 1999-2000. Statsoft Inc.

SULLIVAN, J., E. ARELLANO, & D.S. ROGERS. 2000. Comparative phylogeography of Mesoamerican Highland Rodents: Concerted versus independent response to climatic fluctuations. *American Naturalist,* 155(6): 755-768.

SWOFFORD D. L.. 1991. Phylogenetic Analysis Using Parsimony (PAUP), version 3.0s. Illinois Natural History Survey, Champaign, Illinois, USA.

University of Cape Town

SWOFFORD, D.L. & G.J.OLSEN. 1991. Phylogeny reconstruction. Pp. 411-501. In HILLIS, D.M. & C. MORITZ (EDS). Molecular Systematics. Massachucetts, Sunderland.

SWOFFORD, D.L. 2000. PAUP (Phylogenetic Analysis Using Parsimony).Version 4. Edn. Massachusetts Sunderland.

SWOFFORD, D.L., G.J. OLSEN, P.G. WADDELL, & D.M. HILLIS. 1996. Phylogenetic inference. In HILLIS, D.M., C. MORITZ, & B.K. MABLE (eds). Sinauer Associates, Sunderland, U.S.A.

SYTSMA, K.J. & B.A. SCHAAL. 1990. Ribosomal DNA variation within and among individuals of *Lisianthius* (Gentiaceae) populations. *Pl. Syst. and Evol., 170*: 97-106

TABERLET, P., L. GIELLY, G. PAUTTOU, & J. BOUVET. 1991. Universal primers for the amplification of three non-coding regions of chloroplast DNA. *Pl. Mol.Biol.,* 17: 1105-1109.

TREWICK, S. & G. WALLIS. 1999. Invertebrate molecular phylogeography. http:www.otoago.ac.za.

VAN NUES, R.W., J.M.J. RIENTJES, C.A.F.M. VAN DER SANDE, F.F. ZERP, C. SLUITER, J. VENEMA, R.J. PLANTA, & H.A. RAUE. 1994. Separate structural elements within internal transcribed spacers 1 of *Saccharomyces cerevisiae* precursor ribosomal RNA direct the formation of 17S and 26S rRNA. *Nucleic Acid Research*, 22:912-919

WALKER, P.J. & C.A. PARIS. 1997. Biogeography, speciation, and concerted evolution in American beachgrass (*Ammophilla* Host): Infraspecific

University of Cape Town

sequence variation in the internal transcribed spacer of nuclear ribosomal DNA. *Am. J. Bot.,* 84(6, supplement): 241.

WATSON, E.V. 1971. The structure and life of Bryophytes. Hutchinson & Co. London..

WHITE, T.J., T. BRUNS, S. LEE, & J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In INNIS, M.A., D.H. GELF , J.J. SNINSKY & T.J. WHITE, (eds). Academic Press, San Diego. WOLFE, K.H., W.H. LI, & P.M. SHARP. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci.*, 84:9054-9058.

WYATT, R. & I. J. ODRZYKOSKI. 1998. On the origins of the allopolyploid moss *Plagiomnium cuspidatum. Bryologist,*101: 263-271.

WYATT, R., A. STONEBURNER & I.J. ODRZYKOSKI. 1989. Bryophyte isozymes: systematics and evolutionary implications. Pp 221-240. In SOLTIS, D.E. & P.S. SOLTIS, (eds). Isozymes in plant biology. Portland: *Dioscorides Press*,.

ZANTEN, B.O. & T. van PÓCS. 1981. Distribution and dispersal of bryophytes. *Adv. Bryol.* 1: 479-562.

ZINK, R.M. 1996. Comparative phylogeography in North American birds. *Evolution,* 50: 308-317.

University of Cape Town

Appendices

University of Cape Town

Appendix 1: DNA sequence data (available on request)

1. *Wardia hygrometrica*
    - ITS data
    - trnL-F
    - Combined ITS+trnL-F
2. *L. smithii*
    - ITS data
    - trnL-F
    - Combined ITS+trnL-F

3. *P. gracile*
    - ITS data
    - trnL-F

University of Cape Town