Phylogeography and Population Genetics of two forest endemic mosses in the Cape Floristic Region

Kate Mc Grath
MCGKAT003

Systematics Honours Project
November 2002
Supervisor: Terry Hedderson

Abstract

Different histories for forests in Southern Africa have been hypothesized from vegetation biogeography and pollen analysis. However the history of forests is still controversial. Phylogeography uses gene genealogies to infer history of distributions. Two forest endemic moss species were sampled: *Leptodon smithii,* and *Neckera valentiniana*. Two gene regions were used, *trn*LF (chloroplast genome) and ITS1 (nuclear genome). *Neckera valentiniana* showed no variation from the populations sampled. Results from *Leptodon smithii* based on the *trn* and ITS region suggest that forests once were widespread, but then became fragmented. Dispersal corridors still exist between Southern Cape populations and Western Cape populations.

Introduction

History of forests in Southern Africa

Forests occur in patches throughout Africa, covering less than 5000km^2 in southern Africa (Midgley *et al*). In the Cape Floristic Region (CFR) afromontane forests cover 4.5 % of the area, while thickets cover 2.4 % (Cowling and Holmes 1992). There are three reasons given as to why forests are fragmented in the landscape today. 1) Forests may represent a climax community and existing forests patches relictual (Acocks 1953). 2) Fragmented forests are a result of a recent invasion (Meadows and Linder 1989, 1993); or 3) they have always been patchy (Midgley *et al* 1997).

The following is a summary of what could have happened to forests. The hypotheses about forests come from fossil pollen analysis and biogeographical studies of vegetation.

During the Late Cretaceous-Paleocene widespread temperate forest covered South Africa. With the northward shift of the continent, these were replaced by tropical forests. Conditions at this time favoured continuous distribution of forests (Scott 1997), while later sea level change and climatic changes reduced them to refugial patches (Linder, Meadows and Cowling 1992).

During the Miocene, forest occurred even in places like the karoo (Scott 1997). The increased aridity of the Pliocene could have caused fragmentation of forests; fire has also been linked to fragmentation during this time (Geldenhuys 1989). Forests became the dominant vegetation type in the Cape Floristic Region towards the end of the Tertiary. During this time the climate was warm and temperate. It then deteriorated rapidly to a dry Meditterranean climate that did not favour trees. It is

thought that the habitats of forests in refugia reflect the climate of the Tertiary before it deteriorated (Linder, Meadows and Cowling 1992).

During the Pleistocene, glacial periods became more prevalent and forests were more widely distributed. Not much is known about the last interglacial, but it is thought to have been similar to the Holocene (Scott 1997). Forests were more widespread before the last Glacial Maximum, probably due to moister weather conditions (Scott 1997). During the last Glacial Maximum (18000 yr BP) climates became cold and dry (Geldenhuys 1989; Meadows and Linder 1989, 1993; Scott 1997). These conditions favour fire, subsequently leading to reduction of forests. Forests would have been confined to refugia during this time.

Pollen analysis of afromontane plateaus in Malawi, Zimbabwe and South Africa suggest that forests only appear after the Holocene. Pollen cores dating back to 12000BP, and clay deposits, indicate that forests covered a small area even then, and grasslands were much more dominant. Therefore the grasslands are not relictual. Forest pollen appears only after the onset of the Holocene (Meadows and Linder 1989, 1993).

Endemism is low in forests. There are two alternative explanations for this: that forests were once widespread and interconnected, or that they are a recent Holocene expansion (Meadows and Linder 1989, 1993). It has also been hypothesized that fynbos islands that occur in a sea of forest in the Knysna area are caused by post Pleistocene expansion of forests.

Phylogeography

Most of these hypotheses involve ice ages, which have an approximate 100-kyr cycles, with relatively short, warm, interglacials (which we are in now) (Ferris 1999). Ice-cores in Greenland provide data on the temperature changes that have occurred. Temperature could have changed by as much as 10-12 % over 5 -10 yrs. These temperature changes would modify species distribution. Effects of the Pleistocene ice ages can be addressed in two ways, 1) the fossil record and paleoenvironmental science can be used to gain information regarding climatic change. The climatic change can be correlated with the distribution of the populations, the extent of range changes and the rate of migrations; or 2) or the present geographic genetic structure can be used to determine rates of migration, rates of hybridization, introgression and divergence (Hewitt 1999; Ferris 1999).

Extensive studies have been done on European biota, to gain knowledge on glacial refugia. The studies have used DNA sequence information to elucidate genetic variation and subdivision within species across Europe (Hewitt 1999). The studies superimposed gene genealogies of single or closely related species onto maps of Europe. In this way divergence could be inferred. The data confirmed that southern areas were refugia during the ice ages; they also revealed more refugia than had previously been postulated.

Combining intraspecific DNA/protein variation and geographical distribution is known as phylogeography. The method was pioneered by Avise (1987). It can be defined as 'a field of study concerned with the principles and processes governing the geographic distribution of genealogical lineages, especially those within and among closely related species' (Avise 2000). Phylogeography applies systematic theory to

population genetic processes and is a link between macro and microevolution. The field makes use of coalescent theory in its conceptual framework. Coalescent theory is the name applied to the formal mathematical and statistical treatment of gene genealogies within and among related species. Looking back in time, haplotypes eventually coalesce to a common ancestor (Avise 2000), and the coalescent is a model of genetic drift run backward in time. Coalescent theory is used to model the ancestry of a random sample of genes (Posoda & Crandall 2001). The coalescent is more powerful than diffusion theory (which is based on effective population size) for modeling demographic history (Harding 1996).

Interspecific trees show hierarchical relationships between species. This is because species are a product of reproductive isolation. Populations diverge and mutate, which creates hierarchical structures in their genetic variation. Intraspecific trees cannot be hierarchical, as they are the result of reproduction, and possibly recombination. Haplotype trees are different from phylogenies because haplotype ancestors are usually present (Crandall and Templeton 1996). The older the haplotype the more descendants that are associated with it. Loops in a network might indicate recombination, or homoplasy (reverse parallel mutations). From coalescent theory, a direct relationship is expected between age of haplotypes and frequency; older alleles should be more frequent. They are also usually more widespread (Posada and Crandall 2001). The phylogeny of alleles within species can reveal previously inaccessible details of historical biogeography and population processes (Moritz 1996)

Phylogeographic work was first done using allozyme data. However, these have the following problems; their output is unordered, qualitative, multistate traits. Allozymes are also encoded by nuclear genes, which undergo recombination. Allozyme data was largely replaced by mitochondrial DNA (mtDNA), at least in animals. mtDNA can be seen to be the perfect marker, as it does not recombine, it mutates at a rapid rate, it is maternally inherited (therefore has a smaller effective population size), it is relatively easy to isolate homologous DNA sequences and it has a simple genetic structure (Avise *et al* 1987; Harrison, 1989). Unfortunately the desirable characters of mtDNA are only present in the animal genome. In plants mtDNA mutates very slowly. It therefore does not contain enough variation for phylogeographic studies. mtDNA in plants also undergoes many structural alterations, thus making homology assessment very difficult (Parker *et al* 1998). MtDNA in plants is also hard to purify (Palmer 1992).

An alternative marker to use in plant studies is from the chloroplast genome (cpDNA). It is also maternally inherited (Cruzan 1998). However it mutates at a slower rate (Parker *et al* 1998). The *trn*L-F chloroplast gene has been used to examine hypotheses of long distance dispersal or slow rates of morphological evolution in the trans-Antarctic moss *Pyrrhobryum mnioides*. Chloroplast spacer genes are known to be variable at the species level in several mosses (McDaniel & Shaw 2002).

Nuclear genomes are now also being used in phylogeographic studies, in both plants and animals. A study using the G3pdh nuclear gene found that the nuclear genome can provide multiple, unlinked allele genealogies at the intraspecific level (Olsen & Schaal 1999). Their usefulness is limited by fact that the dominant life history stage

in most plants is diploid. There is therefore the problem of having two different copies of a haplotype within one individual. Diploid nuclear DNA because of its multiple heterozygous sites can have ambiguous linkage sites (Harding 1996; Parker *et al* 1998). Bryophytes do not suffer this problem, as their dominant life history phase is haploid. They are thus ideal organism for studying phylogeographic patterns. The nuclear ribosomal region ITS1 (internal transcribed spacer) has been found to have sufficient variation for phylogeographic resolution in the moss *Hylocomium splendens* (Chiang and Schaal 1999).

Objectives of the study

Phylogeographic techniques have been used to study gene flow and mutation rates in mosses (Chiang and Schaal 1999) and to test hypotheses of dispersal and evolution (McDaniel & Shaw 2002). In this study I will be looking at pattern of genes in two forest endemic mosses, *Neckera valentiniana,* and *Leptodon smithii.* In a previous study (Mwafongo 2002) it was found that *Leptodon smithii* does show variation at this level. The genetic variation in *Neckera valentiniana* is unknown. Looking at phylogeographic patterns of multiple co-distributed groups can contribute to understanding of relationship between earth history and biotic diversification. If multiple groups share common patterns it could be because they have shared biogeographic history (Arbogast and Kenagy 2001). Organisms may not share the same phylogeographic patterns because of differences in life history patterns. Studying gene genealogies of organisms will definitely contribute to our knowledge of the biology of a system, and will hopefully contribute to our knowledge of biogeography. Congruent patterns of phylogeography in vertebrates in Australian rainforests have confirmed previous theory about historical barriers predicted by climatic modelling (Moritz 1996). My objectives in this study were to 1) discover

phylogeographic patterns of forest endemic mosses 2) use the phylogeographic pattern to infer the history of the forest patches and 3) use genetic parameters to determine genetic diversity, and genetic variation among and within populations.

Materials and Methods

The study species

Leptodon smithii (Hedw.) Web & Mohr.

L.smithii is found in forests and woodlands, on bark and on rocks. It is found throughout Southern Africa, from the northern Transvaal, along the coast and in the Northern Cape. It is also found in Europe, southwestern Asia, northern and eastern Africa, Australasia, southern South America and the Juan Fernandez islands. *Leptodon smithii* is dioicous-with archegonia and antheridia on separate plants. The moss forms tufts, with a creeping primary stem. The secondary stems are upright, and display regular branching. Both the stems and the branches curl over when dry, which helps in identification in the field. The distribution of *Leptodon smithii* in Southern Africa is shown in Figure 1.

Neckera valentiniana

N. valentiniana is a mat forming moss which is found in wet forests throughout southern Africa. It forms mats on rocks and trees. It is also found on east African islands. The plats have a flat appearance, and undulate leaves. *Neckera* is autoicous (archegonia and antheridia on separate inflorescences on the same plant) (Magill and van Rooy 1998). The distribution of *Neckera valentiniana* in Southern Africa is shown in Figure 2.

Sampling

This study is restricted to the Cape Floristic Region. This was because of time pressure and the unavailability of fresh material from further afield. Herbarium material could not be used, as it did not amplify. Thirty-one samples from four locations were studied for *Leptodon smithii*. Seven samples from three locations were sampled for *Neckera valentiniana*.

DNA extraction, amplification and sequencing

DNA was extracted using a CTAB method modified from Gawel and Jarrett (1991). About 20 mg of fresh or dried material was placed in 1.5ml microcentrifuge tube. The plant tissue was flash frozen using liquid nitrogen. Plant material was finely ground with a plastic pestle, in the microcentrifuge tube. 700 µl of pre-heated 2XCTAB buffer and 1µl of ß-mercaptoethanol were used for each sample. The ground tissue and buffer were vortexed and the tubes were placed in a 65°C water bath for at least one hour. After incubation 600µl of chloroform:isoamyl alcohol (24:1 v/v) was added to each sample. The tubes were centrifuged for five minutes at 12 000rpm. The supernatant was transferred to a clean 1µl microcentrifuge tube, and an equal volume of ice-cold isopropanol was added. The samples were left overnight in a -20°C freezer, for the DNA to precipitate. When the samples were removed from the freezer they were centrifuged for 5 minutes at 12000 rpm. The pellet containing the DNA was rinsed three times using 75% ethanol. It was then dried for an hour in a desiccator. The DNA pellet was resuspended in 50µl of elution buffer. Extracted DNA was visualized on a 1% agarose gel. Depending on the strength of the DNA extraction, the stock was diluted either 10^{-1} or 10^{-2}.

Initial amplification of the DNA was done in 30µl reactions. Each tube contained 17.65µl distilled H_2O, 3µl 10X NH_4 buffer, 3µl 50mM $MgCl_2$ buffer, 1.2 µl dNTPs (2nm

each nucleotide, Promega), 1il each of the forward and reverse primer (10mM) and 0.15 *Taq* polymerase (Bioline, Biotaq). 3 µl of diluted DNA template were used. The thermal cycle used for initial amplification was: 94°C for 2minutes, 30 cycles of 94°C for1 minute, 52°C for 1 minute, and 72°C for 2 minutes, then 72 °C for 7minutes. PCR products were checked on a 1% gel, and visualised using ethidium bromide. PCR products were cleaned using Qiaquick spin columns (QIAGEN). The DNA was eluted for at least 10minutes, using 30 µl of elution buffer. Concentration of PCR products was estimated by running a 1% agarose gel with a 20ng, 40ng and 60ng PGEM standard.

Cycle sequencing was performed using 10µl reaction volumes. Between 1 and 3 µl of cleaned product was used for each reaction, with the balance been made up with distilled purified water. The cycle sequencing reaction mixture contained1.84 µl of H_20, 2 µl Terminator Ready Reaction mix, 0.16 µl 10mM primer and 2,5 X buffer. The Terminator Ready Reaction and sequencing buffer were from the ABI prism Big Dye kit (Applied Biosystems, Inc. Palo Alto, CA). The cycle sequencing thermal profile used was 25 cycles of: 30 seconds at 96°C, 15 seconds at 50°C and 4 minutes at 60°C. Initial amplification and cycle sequencing were done on either a GeneAmp PCR system (Applied Biosystems, Inc. Palo Alto, CA) or a Hybaid PCR Sprint Temperature Cycling System (Ashford, Middlesex TW15 1XB, UK). Final products from the cycle sequencing reaction were cleaned in spin columns and resolved on an ABI 3900 autosequencer.

The following primers were used for initial amplification and cycle sequencing: ITS (nuclear) reverse ITS2 GCT GCG TTC TTC ATC GAT GC (White *et al* 1990); forward 18KRC GCA CGC GCG CTA CAC TGA (Hamby *et al* 1991). *trn*L-F (chloroplast)

forward *trnC* CGA AAT CGG TAG ACG CTA CG; reverse *trnF* ATT TGA ACT GGT GAC ACG AG (Taberlet *et al* 1991). Two DNA regions were used in this study as gene genealogies give the genealogy of the gene used, and not necessarily the organismal genealogy. As different genomes undergo different history, a study is more robust if it includes genes from different genomes (Parker *et al* 1998).

Forward and reverse sequences were assembled using Seqman (LaserGene System Software, DNAstar Inc.). The consensus sequence was saved as a single file, and assembled sequences were aligned manually in MegAlign (LaserGene System Software, DNAstar Inc). All sites that showed variation were checked in Seqman to make sure that variation was unambiguous. For analyses gaps were treated as a 5th character state. Aligned sequences were trimmed in MacClade in order that they could be used in TCS. Phylogeographic analysis was done using TCS 1.13 (Clement *et al* 2000). TCS estimates gene genealogies from sequence data, estimating a set of probable relationships between haplotypes whose cumulative probability is 0.95.

Population parameters were estimated using the program Arlequin 2.0 (Schneider, Roessli and Excoffier 2000). Populations were defined as: Jonkershoek, Nursery Ravine, Fernkloof and Sedgefield. Groups for AMOVA were designated Southwestern Cape: Jonkershoek, Nursery Ravine and Fernkloof, and Southern Cape: Sedgefield. The following population analyses were performed on the sequence data:

1) AMOVA (Analysis of Molecular Variance) – this uses the allelic content of haplotypes, as well as frequencies of the haplotypes.
2) Tajima selective neutrality test – compares two estimates of population parameter theta, one based on the number of segregating sites in a sample, and the other based on the mean number of pairwise differences between

the haplotypes. Under the infinite site model, both estimators should estimate same quantity; differences can arise under selection, population non-stationary, and heterogeneity of mutation rates among sites. theta = $2M\mu$, where M=N, the population size, in a haploid population and μ represents the overall mutation rate at haplotypic level.

3) Overall pairwise nuclear diversity was calculated to determine nucleotide diversity within populations.
4) Corrected Tamura Nei distances were calculated to determine diversity among populations.
5) To test the differences among populations based on haplotype frequencies an Exact Test of sample Differentiation based on Haplotype Frequencies was performed (Raymond, M. & Rousset, F. 1995). This tests the random distribution of individuals between pairs of populations.

Results

Phylogeographic analysis

Neither of the DNA regions showed variation among the seven samples of *Neckera valentiniana* analysed. Therefore no further phylogeographic or population analyses were attempted.

For *L. smithii* the amplified ITS region was approximately 600 base pairs long. Because some samples were missing sequences at the end of the region, they had to be trimmed to 533 base pairs. The *trn*L-F region was approximately 400 base pairs long. It was trimmed to 381 base pairs.

Seven haplotypes were identified for the ITS1 region (Figure 3, Table 1). The A haplotype is the most frequent (12 out of 31samples) and the most widespread. The A haplotype has a gap at position 615 in relation to the other samples. The second most frequent haplotype is B (A instead of G at position 546). It is only found in the southwestern cape populations. The samples from Fernkloof comprise one haplotype, which was not found in any of the other populations. The Sedgefield population contains two haplotypes that are not shared by the remaining populations. The Swartboskloof population also contains a unique haplotype. The ITS tree contains one loop, indicating that haplotype F is two mutational steps from both B and C and therefore it is not possible to discern which gave rise to it.

The *trn*-LF sequences only revealed four haplotypes (Figure 4, Table 1). Haplotype 1 is found in all the populations. Nursery Ravine and Swartboskloof share Haplotype 2. Both Sedgefield and Swartboskloof each have a unique haplotype. The Fernkloof population contains only the most common haplotype. The individual moss sampled from Montagues' pass has the most frequent haplotype for both the trnLF and ITS region.

The distribution of the haplotypes, and their relative frequencies' are shown in Figure 5 and Figure 6. Figure 5, haplotype frequencies for ITS1 region, shows clearly that the Fernkloof haplotype is restricted to Fernkloof. It also shows the haplotype A (Black) is found throughout the region sampled.

Figure 6 shows the haplotype frequencies for the *trn*LF region. Haplotype 1 (black) is found in all the populations, but its frequency varies throughout the populations.

Population genetics

The Tajima selective neutrality probabilities are all approximately 1. This indicates a stable population, equal mutation rates within a population and no selection.

As with the phylogeographic analysis there is a difference between the two gene regions. The AMOVA for ITS sequences (Table 2A) shows that most of the variation occurred within populations (86%). Among group variation, and variation among populations, within groups, accounts for approximately seven percent each. The Fst value was low, close to zero, indicating that there was a lot of similarity between the sequences. Another measure of difference was given by nucleotide diversity. For both the ITS and *trn* region the nucleotide diversity was low (Table 2B and 3B). The maximum Tamura and Nei distance was found between Fernkloof and Jonkershoek (with the Sedgefield population not being much different). Tamura and Nei values that are negative are: Jonkershoek and Sedgefield, and Jonkershoek and Nursery Ravine (Table 3C).

For the trnLF region the variation among populations and within population values component are about equal. No extra variation is accounted for by the among region component. The maximum Tamura and Nei distance was found between Nursery Ravine and Sedgefield. There are no negative values between populations based on the *trn* region. However the Tamura and Nei value between Sedgefield and Fernkloof is zero. Neither of the Fst values for the AMOVA is significant.

Exact tests for sample Differentiation show that ITS (Table 2C) has different frequencies. Only the Jonkershoek and Nursery Ravine do not have a significant differentiation between their haplotype frequencies. The *trn* region has very

different results (Table 3C). Nursery Ravine and Sedgefield, and Nursery Ravine and Fernkloof are the only populations with significant differences between their frequencies.

Discussion

The current study reveals a considerable degree of haplotype variation within South African populations of *Leptodon smithii*.

It is unlikely that the F haplotype is derived from C, as C is the unique haplotype found at Fernkloof. It is much more likely to have arisen from B, which is found in the same population. Therefore the loop could have arisen because of homoplasy. Further sampling could help clarify this relationship. That D and G arise from B is unusual, as B is restricted to the South Western Cape and D and G are found exclusively at Sedgefield. More sampling might show that haplotype B is more widespread than is suggested now.

There is definite gene flow between Nursery ravine and Swartboskloof. According to coalescent theory, the haplotype that is the most frequent and widespread is the oldest. Therefore haplotype 1 and haplotype A are the most likely the ancestral haplotypes.

The Peninsula forests are very isolated from the main Western Cape mountain ranges, share more species with forests along the coast. This would suggest that a coastal distribution corridor is very likely (Geldenhuys 1989). These results do not support this theory. Cape peninsula forests show affinity with forests in the Western Cape mountain range (Jonkershoek). Also there is sharing of haplotypes between

Sedgefield and the Cape Peninsula. Furthermore Fernkloof, which would most likely form part of coastal corridor, is completely separate with respect to the ITS haplotype.

Based on the ITS region Jonkershoek and Nursery Ravine have more shared haplotypes. For the *trn* region Nursery Ravine and Sedgefield are the most divergent. However, the Jonkershoek and Fernkloof population is very closely related. The ITS region and the *trn* region show very different patterns. The population analyses support the phylogeographic analyses in that they show that chloroplast genome undergoes a different history to the nuclear genome. If the chloroplast genome does indeed mutate slower, then there is evidence for widespread forests throughout the Cape Floristic Region, which then fragmented. The timing of the fragmentation could not be inferred from this study.
The variation is low among populations. Comparison of Tamura and Nei distance (dA) values in this study with those of McDaniel and Shaw, 2002; dA ~ 0 in this study, dA ~ 37.09 for Mc Daniel and Shaw. The distance between populations in the McDaniel and Shaw is much larger than the distance between the populations in this study, and genetic differentiation usually does increase with geographic variation (Slatkin 1993). The Tamura and Nei values could also have been low because most of the variation was explained within populations, not among.

Although the genetic distances among populations are low, there are significant differences in haplotype frequencies. The *trn* data reveal that gene flow was occurring between most of the populations. The ITS data show significant differences in haplotype frequency between all populations, except between Jonkershoek and Nursery Ravine. The haplotype frequency further suggest that the

pattern revealed in *trn* occurred when forests were widespread, and then the ITS pattern occurred were forests were fragmented.

Chloroplast genes mutate slower than nuclear genes. They therefore can show genetic patterns that have more ancient origins (Hewitt 1999). Therefore the pattern revealed by the Fernkloof population is that there used to be more widespread forests, which allowed for gene flow (as can been seen in the *trn* region). Then the forests fragmented. This led to some forest being restricted, and then diverging from the other forests (e.g. Fernkloof ITS region).

The lack of variation within *Neckera valentiniana* requires some explanation. The lack of variation could be because 1) too little sampling was done; therefore rare haplotypes were never found. 2) There could be long distance dispersal between the populations that would cancel any changes that occur in the population due to natural drift. 3) There is little variation as the forest expansion occurred recently, and there has not been time for divergence. We cannot assume that mutation rates are the same across mosses. Therefore we cannot exclude the possibility that the mosses could have the same history, but are showing different phylogeographic patterns because they have different mutation rates. Life history could determine if an organism will have much or little gene variation. It is unlikely that organisms that reproduce vegetatively will have much genetic variability, as their genes do not recombine. On the other hand, changes that do occur are more likely to persist. Organisms that undergo sexual reproduction are more likely to have more variation, as their genes recombine. *Pterogonium gracile* is a forest endemic moss that seems to depend totally or heavily on asexual reproduction (Mwafongo 2002). Phylogeographic studies of *Pterogonium gracile* with samples from Portugal and

Spain through to samples from Table Mountain showed no variation. This would agree with the hypothesis that asexually reproducing organisms show little variation. *Neckera valentiniana,* however, is a sexually reproducing moss. It does not show variation. Not all the mosses need to have the same history. Perhaps *Pterogonium* and *Neckera* are mosses that arrived recently in the forests.

As the two species studies did not show the same phylogeographic pattern, a comparative phylogeographic analysis could not be done. Therefore strong biogeographical inferences could not be made from the data. The results from *Leptodon smithii* do show intraspecific variation can be found within forest patches of South Africa. Therefore this species is useful for phylogeographic studies.

Future research

Within population sampling from a broad geographic range needs to be conducted: this would allow the use of methods such as nested clade analysis for testing hypotheses. Nested clade analysis allows one to partition population historical factors from population structure (Crandall and Templeton 1996). In addition more mosses need to be identified that show variation, so comparative phylogeographic analyses can be done. It would be useful to discover more on life history. This would help to elucidate whether lack of variation is caused by long distance dispersal, or by recent expansion. Dispersal can negate any changes that happen in a population because of drift (Cruzan 1998). It would also be useful to sample from Afromontane, and Indian Ocean (Tongaland/Pondoland) forests (Midgley *et al* 1997), in order to determine the origin of these forest types.

Acknowledgments

Thank you to Terry Hedderson, my supervisor, for giving me this project and for all ideas and encouragement. Thank you also to Tracey Nowell for all your help in the lab

References

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

Arbogast, B.S. & Kenagy, J. 2001. Comparative phylogeography as an integrative approach in biogeography. Guest editorial, *Journal of Biogeography.* 28 (7).

Avise, J.C. 2000. Phylogegraphy: the history and formation of species. Harvard University Press, Cambridge.

Avise,J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. and Sauders, N.C. 1987. Intraspecific phylogeny: The mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics.* 18: 489 – 522.

Chiang, T.Y. & Schaal, B.A.. 1999. Phylogeography of North American populations of the moss species *Hylocomium splendens* based on the nucleotide sequence of internal transcribed spacer 2 of nuclear ribosomal DNA.. *Molecular Ecology.* 8: 1037 – 1042.

Clement, M.D., Posada, D. & Crandall, K.A. 2000. TCS: A computer programme to estimate gene genealogies. *Molecular Ecology.* 9(10): 1657 – 1660.

Cowling, R.M. & Holmes, P. 1992. Flora and vegetation. In *The ecology of fynbos: nutrients, fire and diversity,* edited by Cowling, R.M. Oxford University Press. Oxford, U.K.

Crandall, K.A, & Templeton, A..R. 1996. Applications of intraspecific phylogenies. In *New Uses for New Phylogenies,* edited by Harvey, P.H., Leigh Brown, A.J., Maynard Smith, J. and Nee, S. Oxford University Press, Oxford. Pp 81 – 99.

Cruzan, M.B. 1998. Genetic markers in plant evolutionary ecology. *Ecology.* 79(2): 400 – 412.

Ferris, C., King, R.A. & Hewitt, G.M. 1999. Isolation within species and the history of glacial refugia. In *Molecular Systematics and Plant Evolution,* edited by Hollingsworth, R.M., Bateman , R.M. & Gornall, R.J. Taylor and Francis. 20 – 34.

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

Geldenhuys, C.J. 1989. Richness, composition and Relationships of floras of selected forests of Southern Africa. In *Biogeography of mixed evergreen forests of South Africa,* edited by C.J. Geldenhuys. FRD, Occasional report 45. Pretoria.

Geldenhuys, C.J. 1989. Phytogeography of Southern Cape Forests and Flora. In *Biogeography of mixed evergreen forests of South Africa,* edited by C.J. Geldenhuys. FRD, Occasional report 45. Pretoria.

Harding, R.M., 1996. New phylogenies: an introductory look at the coalescent. In *New Uses for New Phylogenies,* edited by Harvey, P.H., Leigh Brown, A.J., Maynard Smith, J. and Nee, S. Oxford University Press, Oxford. 18 – 22.

Harrison, R.G. 1989. Animal mitochonadrial DNA as a Genetic Marker in Population and Evolutionary Biology. *Trends in Ecology and Evolution.* 4(1): 6 – 11.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Biologica Journal of the Linnean Soceity.* 68: 87 – 112.

Linder, H.P., Meadows, M.E. & Cowling, R.M. 1992. History of the Cape Flora. In *The ecology of fynbos: nutrients, fire and diversity,* edited by Cowling, R.M. Oxford University Press. Oxford, U.K.

Magill, R.E. & van Rooy, J. 1998. Flora of Southern Africa, Bryophyta Part 1 Fascicle 3 Erpodiaceae – Hookeriaceae edited O.A.. Leistner. National Botanical Institute. Pretoria. 588 – 591

McDaniel, S.F. & Shaw,J. Phylogeographic structure and cryptic speciation in the trans-antartic moss *Pyrrhobryum mniodes.*

Meadows, M.E. & Linder, H.P. 1993. A paleoecological perspective on the origin of Afromontane grasslands. *Journal of Biogeography*. 20: 345 – 355.

Meadows, M.E.& Linder,H.P. 1989. A reassessment of the biogeography and vegetation history of southern afromontane region. In *Biogeography of mixed evergreen forests of South Africa*, edited by C.J. Geldenhuys. FRD, Occasional report 45. Pretoria. 15 – 29

Midgley, J.J., Cowling, R.M., Seydack, A.H.W. & van Wyk, G.F. 1997. Forest. In *Vegetation of Southern Africa* edited by Cowling,R.M., Richardson, D.M. & Pierce, S.M. Cambridge University Press, UK. Pp 278 – 299.

Moritz, C. 1996. Uses of molecular phylogenies for conservation. In *New Uses for New Phylogenies,* edited by Harvey, P.H., Leigh Brown, A.J., Maynard Smith, J. and Nee, S. Oxford University Press, Oxford. 203 – 213.

Mwafongo, E. 2002. Phylogeographic patterns in three South African forest mosses. Unpubl. MSc thesis. University of Cape Town

Olsen, K.M. & Schaal, B.A. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences*. 96: 5586 – 5591.

Palmer, J. 1992. mtDNA in Plant Systematics: Applications and Limitations. In Molecular Systematics of Plants, edited by Soltis, P.S., Soltis, D. and Doyle, J. Chapman and Hall, New York. 36 – 48.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C. & Fuerst, P.A.. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology*. 72(9): 361 – 382.

Posada, D. & Crandall, K.A.. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*. 16 (1): 37 – 45.

Schneider,S., Roessli, D. & Excoffier, L. 2000. Arlequin ver 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Scott, H.M., Anderson, H.M. & Anderson, J.M. 1997. Vegetation history. In *Vegetation of Southern Africa,* edited by Cowling, R.M., Richardson, D.M. & Pierce, S.M. Cambridge University Press, UK.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*. 47(1): 264 – 279.

Table 1: Sample number, locality and haplotype for *Leptodon smithii* samples

SAMPLE NUMBER	LOCALITY	HAPLOTYPE *trn*L-F	HAPLOTYPE ITS1
N1	Nursery Ravine, Cape Peninsula	2	A
N2	Nursery Ravine, Cape Peninsula	1	B
N3	Nursery Ravine, Cape Peninsula	2	B
N4	Nursery Ravine, Cape Peninsula	1	B
N5	Nursery Ravine, Cape Peninsula	2	A
N6	Nursery Ravine, Cape Peninsula	2	A
N7	Nursery Ravine, Cape Peninsula	2	E
N8	Nursery Ravine, Cape Peninsula	no sample	B
N9	Nursery Ravine, Cape Peninsula	2	A
N10	Nursery Ravine, Cape Peninsula	no sample	A
J1	Swartboskloof, Jonkershoek	1	no sample
J1b	Swartboskloof, Jonkershoek	2	B
J2	Swartboskloof, Jonkershoek	1	A
J3	Swartboskloof, Jonkershoek	1	A
J4	Swartboskloof, Jonkershoek	1	F
J5	Swartboskloof, Jonkershoek	1	A
J6	Swartboskloof, Jonkershoek	1	A
J7	Swartboskloof, Jonkershoek	2	B
J8	Swartboskloof, Jonkershoek	3	B
J9	Swartboskloof, Jonkershoek	2	B
SF1	Sedgefield	4	A
SF2	Sedgefield	1	A
SF3	Sedgefield	1	A
SF4	Sedgefield	1	D
SF5	Sedgefield	1	G
SF7	Sedgefield	1	A
SF8	Sedgefield	1	D
FK3	Fernkloof, Hermanus	1	C
FK4	Fernkloof, Hermanus	1	C
FK5	Fernkloof, Hermanus	1	C
FK7	Fernkloof, Hermanus	1	C
FK8	Fernkloof, Hermanus	no sample	C
WF	Montague's Pass, George	1	A

Table 2A: AMOVA based on ITS1 region.

Source of variation	d.f.	Sum of Squares	Variance components	Percentage of variation
Among groups	1	0.75	0.028	7.56
Among populations, within groups	2	1.01	0.024	6.44
Within populations	26	8.33	0.320	86.00
	Fst = 0.14			

Table 2B: Within population parameters based on the ITS1 region

	nucleotide diversity	Tajima (P)
Jonkershoek	0.00105	0.973
Sedgefield	0.002766	0.712
Nursery Ravine	0.001005	0.940
Fernkloof	0.00000	1.000

Table 2C: Among population genetic parameters based on ITS1. Above the diagonal are Tamura and Nei corrected pairwise differences, and below the diagonal are the significance results from the exact test for sample differentiation. Significant results are given as a '+'.

	Jonkershoek	Sedgefield	Nursery Ravine	Fernkloof
Jonkershoek	~	-0011	-0.03341	0.2800
Sedgefield	+	~	0.00008	0.2700
Nursery Ravine	-	+	~	0.1340
Fernkloof	+	+	+	~

Table 3A: AMOVA based on *trn*LF

Source of variation	d.f.	Sum of Squares	Variance components	Percentage of variation
Among groups	1	0.90	-0.001	-0,37
Among populations, within groups	2	1.74	0.105	41.97
Within populations	25	3.64	0.145	58.40
	Fst=0.41			

Table 3B: Within population parameters for the *trn*LF region

	Nucleotide diversity	Tajima (P)
Jonkershoek	0.0017	0.864
Sedgefield	0.0008	1.00
Nursery Ravine	0.0011	0.809
Fernkloof	0.0000	1.00

Table 3C: *trn* LF Population genetic, above the diagonal is the corrected pairwise differences (Tamura and Nei), below the diagonal are results from the exact test for population differentiation.

	Jonkershoek	Sedgefield	Nursery Ravine	Fernkloof
Jonkershoek	~	0.067	0.153	0.067
Sedgefield	-	~	0.540	0.000
Nursery Ravine	-	+	~	0.054
Fernkloof	-	-	+	~

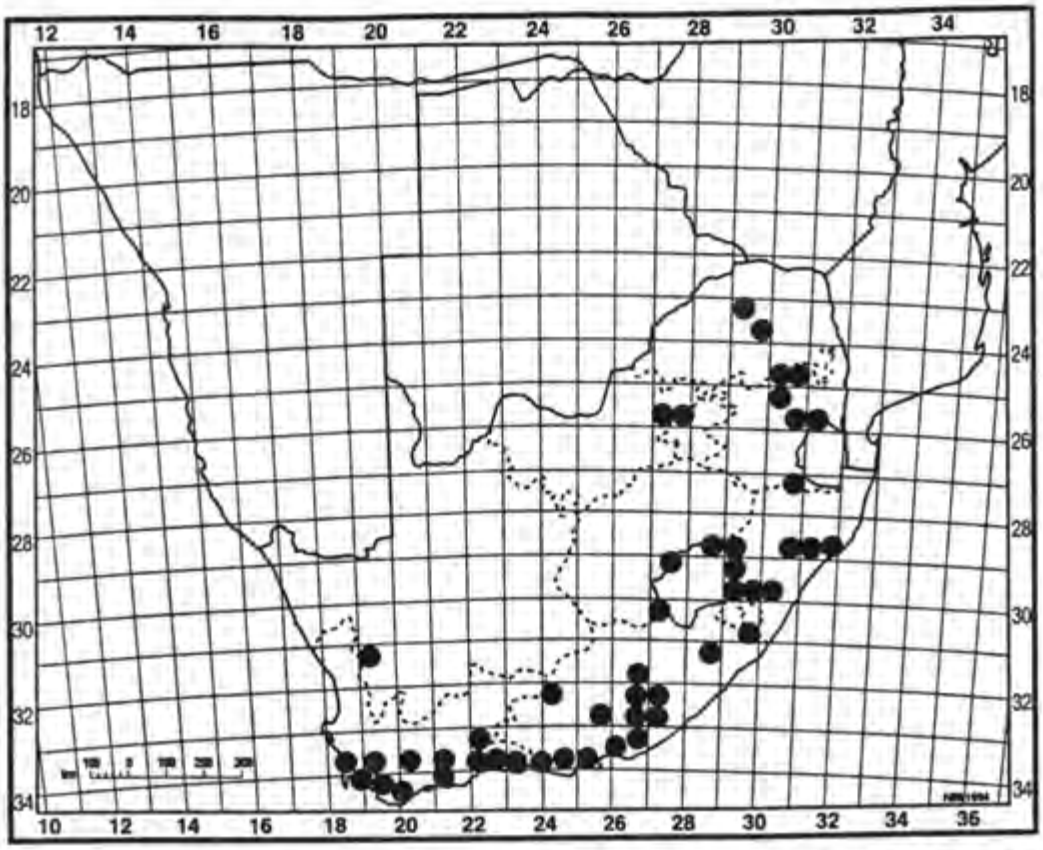

Figure 1: Map showing distribution of *Leptodon smithii* in South Africa (after Magill and Rooy 1998)

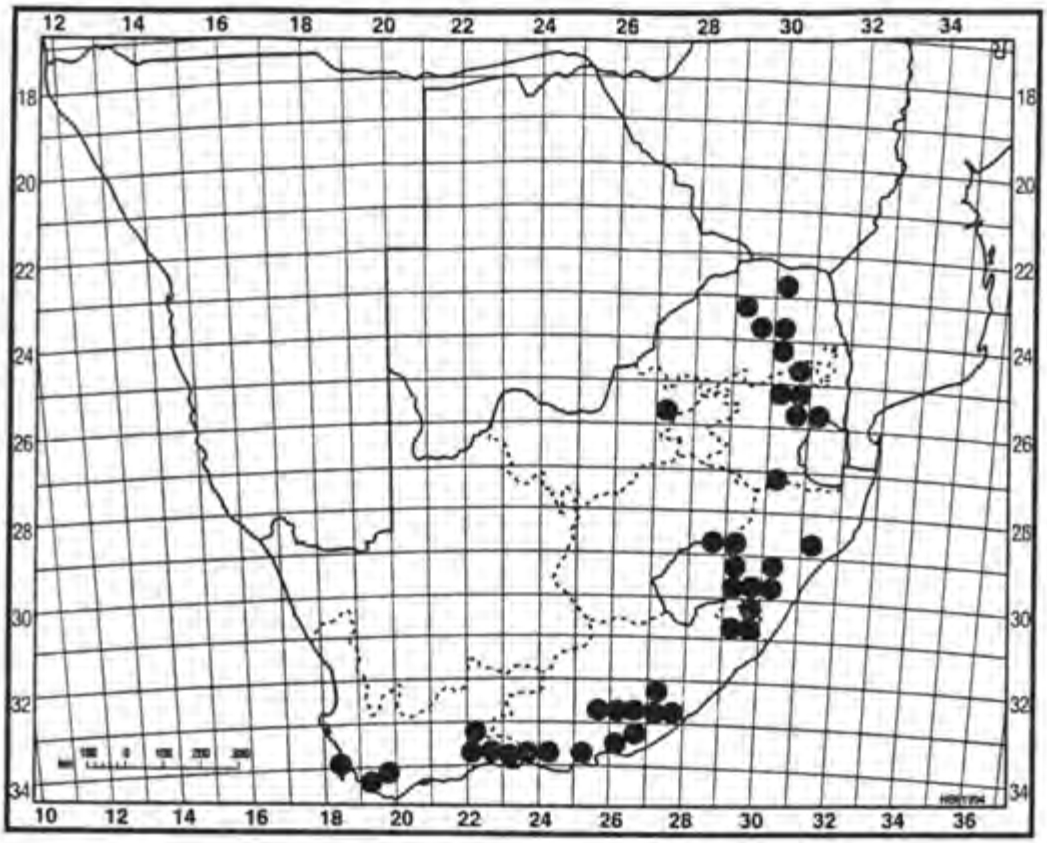

Figure 2: Map showing distribution of *Neckera valentiniana* in South Africa (after Magill and Rooy 1998)

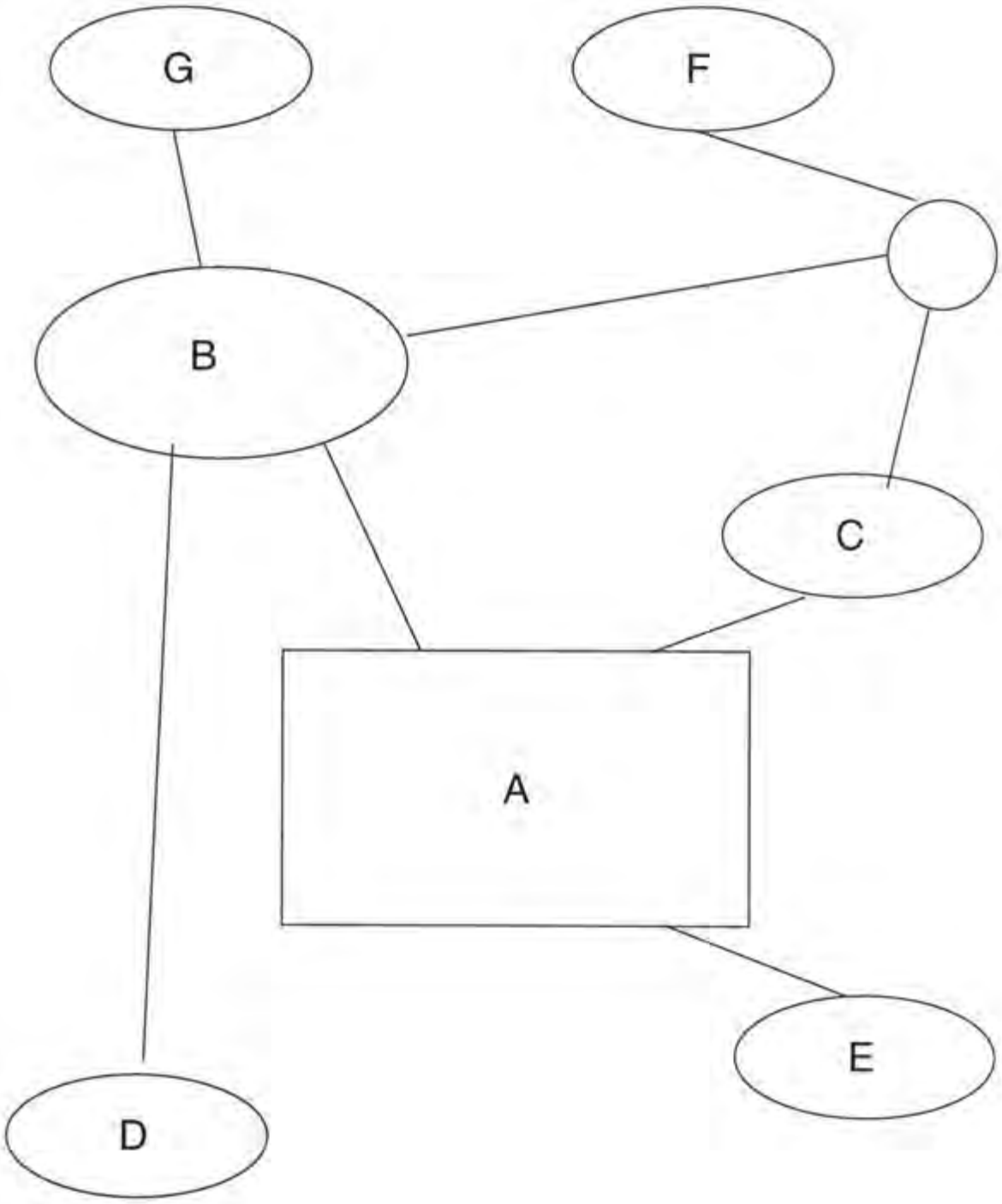

Figure 3: Statistical parsimony network for ITS2.
Size of nodes indicates the frequency of the haplotypes.

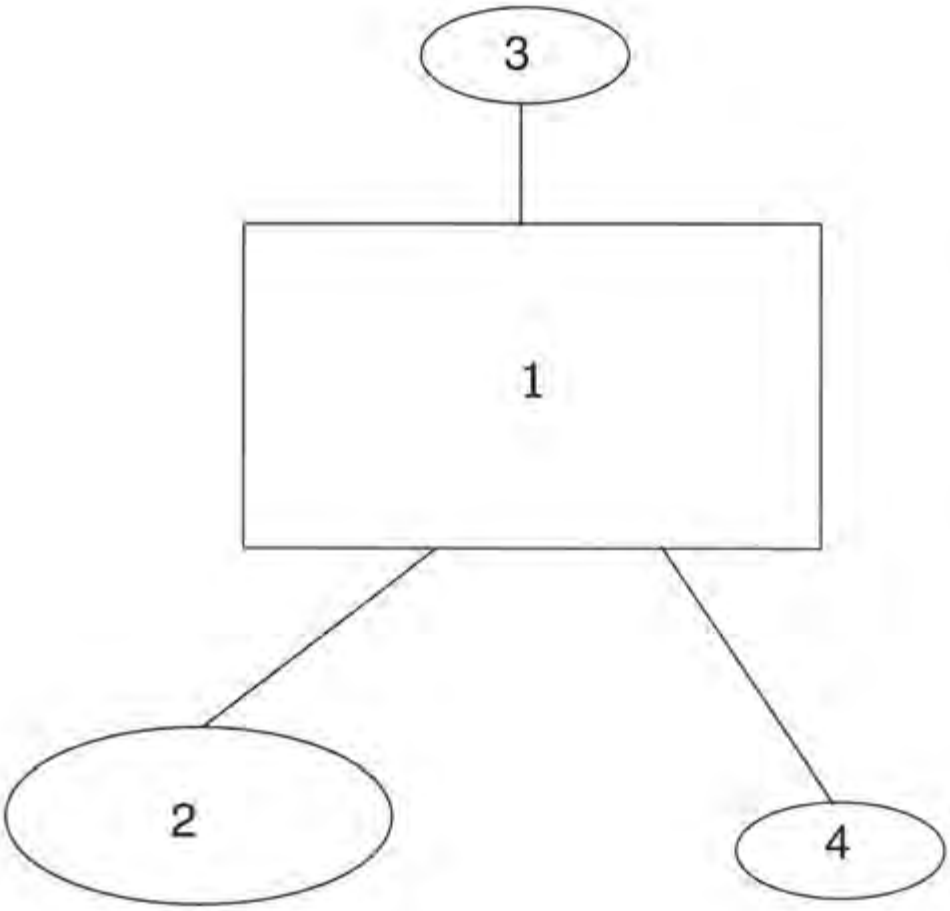

Figure 4: Maximum parsimony network of *trn*L-F haplotypes Size of each haplotype group gives an indication of number of sequences.

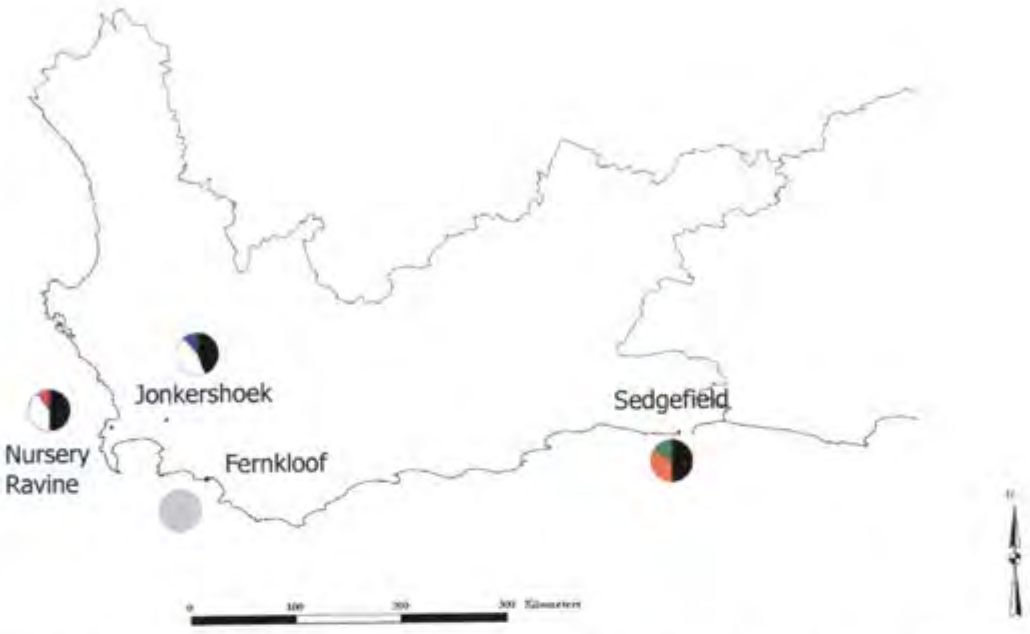

Figure 5: Relative haplotype frequencies for the ITS2 gene. Black represents Haplotype A, white represents haplotype B, grey represents haplotype C, orange represents haplotype D, red represents haplotype E and blue represents haplotype F.

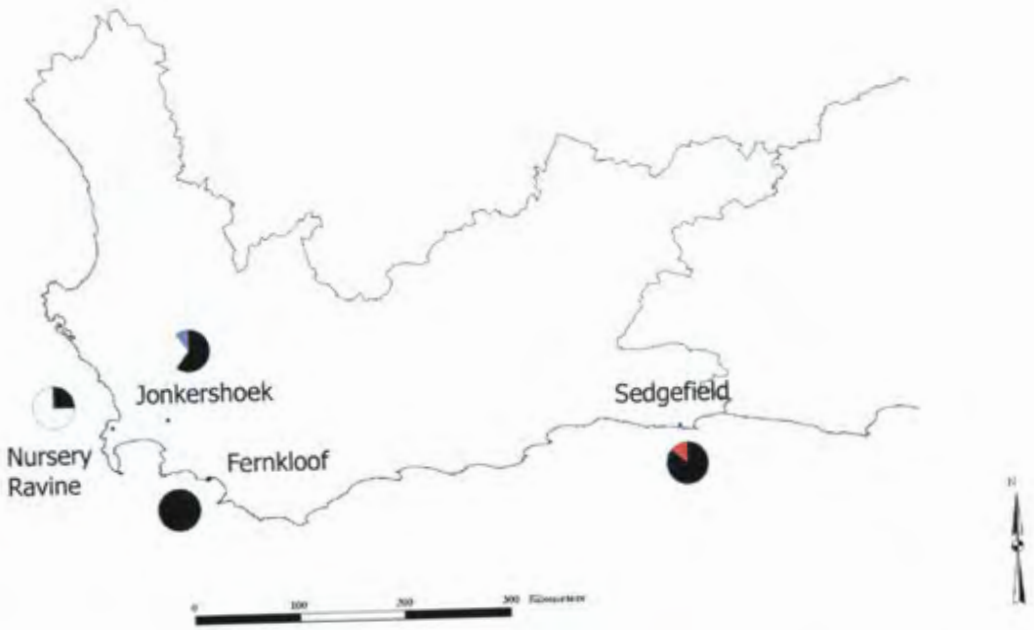

Figure 6: Relative haplotype frequencies for the *trn*L-F spacer. Black represents haplotype 1, white indicates haplotype 2, blue indicates 3 and red represents haplotype 4.