Genetic variation of *Clivia caulescens*

Suzanne Stegmann

Dissertation presented in order to qualify for the degree *Magister Scientiae* Genetics in the Faculty of Natural and Agricultural Sciences, Department of Genetics, at the University of the Free State

June 2011

Supervisor: Prof. J.J Spies

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Most importantly I would like to thank God for making it possible for me to complete this huge task at hand. Without Your blessing it would not have been achievable.

I would like to thank Prof J.J. Spies, my study leader for giving me the opportunity to study under him. Thank you for your guidance and supply of plant material.

To all my colleagues at the laboratory, thank you for all the kind, and encouragement words and help whenever I needed something. I would like to extend a special thank you to Hesmari, who were definitely my guardian angel during this study.

I would like to thank the University of the Free State for the use of their facilities.
I would also like to thank the NRF for their financial support.

Furthermore a big thanks to Attie Le Roux, Johann Schoeman, James and Connie Abel, George Mann, Sean Chubb, the late Bertie Guilluame and Brian Tarr, the suppliers of plant material for this study, without your contributions this study will not have been possible. If I left someone out by mistake, my sincere apology and a very big thank you to you as well.

A very warm thank you goes out to Willie for believing and supporting me in this endeavour. His continuous encouragement meant the world to me. To all my friends for just being there when I needed to blow off steam, thank you.

Lastly I would like to extend a very huge thank you to my parents and family for your love, support and understanding.

ABBREVIATIONS

A	Adenine
AFLP	Amplified Fragment Length Polymorphism
b	extinction coefficient path length
bp	Base pair
BS	Bootstrap
cm	centimeter
C	Cytosine
℃	Degree Celsius
CI	Consistency index
$CaCl_2$	Calcium chloride
cpDNA	Chloroplast DNA
CTAB	Hexadecyltrimethyl Ammonium Bromide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dH_2O	Distilled water
dNTP	Deoxynucleotide Triphosphate
dsDNA	double stranded DNA
EDTA	Ethylene Diamintetra Acetic Acid
e.g.	for example
Ethanol	Ethyl-alcohol
EST	Expressed Sequence Tags
F	Forward primer
Fig.	Figure
G	Guanine
g	gram
g	Gravitational Force
HCL	Hydrochloride acid
INDELS	Insertions/Deletions
IR	Inverted repeats
kb	kilobase
LSC	Large single copy region

M	Molar
matK	maturase
min.	Minute
$MgCl_2$	Magnesium chloride
ml	milliliter
µl	Micro liter
m/v	Mass per Volume
mg/ml	Milligram per Milliliter
mM	Millimolar
mm	millimeter
NaCl	Sodium chloride
ND	Nanodrop
ng	Nanogram
ng/µl	nanogram per microliter
nm	nanometer
ORF	Open reading frame
PCR	Polymerase Chain Reaction
PI	Propidium iodide
pmol	picomol
pmol/µl	Picomol per Microlitre
R	Reverse primer
RAPD	Random Amplified Polymorphic DNA
RI	Retention index
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
rpoB	RNA polymerase beta-subunit-encoding gene
rpoC	RNA polymerase beta-subunit-3' exon
s	Second
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
SSC	Small Single Copy
STR	Short Tandem Repeat
T	Thymine

T_a	Annealing temperature
TAE	Tris; Acetic Acid; EDTA
Taq. Pol.	*Thermus aquaticus* Super Therm DNA Polymerase
TRIS	2-amino-2-(hydroxymethyl)-1, 3-propanediol
trnK	Transfer RNA gene for lysine
U	Units
UV	Ultraviolet
U/µl	Unit per microliter
V	Volts
VNTR	Variable Number Tandem Repeat
v/v	Volume per volume
%	Percentage

APPENDICES

APPENDIX A - Nanodrop readings of the *Clivia caulescens* specimens used in this study.

APPENDIX B - Gel documents of presequence products from all the populations represented in this study of *Clivia caulescens*.

APPENDIX C - Geneious illustrations of the sequences obtained.

APPENDIX D - Cladograms obtained from the other gene regions (*atpH-I*; *matK*; *rpoB* and *rpoC1*) by making use of Maximum Parsimony and Minimum Evolution methods for the populations of *C. caulescens* included in this study.

APPENDIX E - Tables reflecting the mean distance between the gene regions (combined data set; *matK* and *rpoB*).

APPENDIX F - Flow diagrams of the mean distances between populations of *C. caulescens* specimens included in this study.

APPENDIX G - Flow diagrams of mean distances between locations of *C. caulescens* specimens included in this study.

APPENDIX H - Nanodrop readings of the Clivias at the Bearded Man Mountain used in this study.

APPENDIX I - Gel documents of presequence products from the Bearded Man Mountain.

APPENDIX J - Illustration of Geneious output for the different gene regions used in this study.

APPENDIX K - Illustration of SNPs obtained through the different gene regions (*atpH-I*, *matK* and *rpoB*)

Chapter 1:

GENETIC VARIATION IN *CLIVIA CAULESCENS* – Literature review

Suzanne Stegmann

1.1 ABSTRACT

Clivia is a genus of great horticultural importance, as many of its species and cultivars are grown worldwide. There are currently six species, i.e. *Clivia nobilis* Lindl., *C. miniata* (Lindl.) Regel, *C. gardenii* Hook., *C. caulescens* RA Dyer, *C. mirabilis* Rourke and *C. robusta* Murray, Ran, De Lange, Hemmet, Truter & Swanevelder. This study concentrates on the genetic variation in *C. caulescens*.

This chapter contains a literature review on the *Clivia* species in Mpumalanga, the molecular techniques available to determine genetic variation in *C. caulescens*, and an overview of the layout of this dissertation.

1.2 INTRODUCTION

The popularity of ornamental plants fluctuates and this phenomenon is also observed in *Clivia*. The establishment of the *Clivia* Society in 1992 heralded in a new era of interest in this extraordinary genus. Traits of interest for the market include leaf width, leaf variegation, flower form, flower colour and interspecific hybrids (Swanevelder, 2003).

When considering the horticultural industry many of the currently important bulb species e.g. daffodils, tulips etc., have been highly developed by decades of selection and breeding, resulting in large differences from the original wild form or forms which they were derived from. Information concerning the development of modern ornamental cultivars is somewhat uneven in quantity and quality; some genera and species have been studied in detail, others lack comprehensive investigation (Rees, 1992).

The analyses of the genetic diversity and the relationships between and within different populations/species have, therefore, become an important part of breeding programmes. During the seventies and eighties of the previous centuries, classical procedures for evaluating genetic variability became increasingly complemented by molecular techniques (Wiesing & Gardner, 1999). In an attempt to rectify this situation in *Clivia*, this study was conducted.

1.3 GENERAL OVERVIEW

The family, Amaryllidaceae, consists of 61 genera (Meerow *et al.*, 2000), and is concentrated in southern Africa and the Mediterranean (Duncan & Du Plessis, 1989). The genus, *Clivia* Lindl. is a member of the Amaryllidaceae, from the tribe Haemantheae, which includes the baccate-fruited genera *Scadoxus* Raf., *Haemanthus* L., *Clivia, Cryptostephanus* Welw., *Gethyllis* L., *Apodolirion* Baker and *Cyrtanthus* Aiton. (Meerow, 1995; Germishuizen & Meyer, 2000). Although it is not strictly speaking a bulbous plant, it is normally treated as such (Meerow *et al*., 1999). *Clivia* plants are evergreen and have predominantly orange coloured flowers, although yellow, red and pastel coloured flowers are sometimes observed in nature.

At present, the genus *Clivia* consists of six species (Figure 1.1), which include *Clivia nobilis* Lindl., *C. miniata* (Lindl.) Regel, *C. gardenii* Hook., *C. caulescens* RA Dyer, *C. mirabilis* Rourke and *C. robusta* Murray, Ran, De Lange, Hemmet, Truter & Swanevelder. Many of the species and cultivars are extensively grown worldwide, making this group of considerable horticultural importance (Truter *et al*., 2007).

Clivia was originally thought to be a species of *Agapanthus* L'Her., by Dean William Herbert, because of its deep green strappy leaves and the considerable resemblances between the vegetative parts of the two species. Their dried specimens look very similar (Koopowitz, 2002). The difference between these two species were detected when they flower, which is not phenotypically similar.

Clivias are slow growing plants with a reproductive cycle of three to twelve years and are adapted to a fertile humus soil environment.

1.3.1. *Clivia caulescens*

This study focused on *Clivia caulescens*. The natural habitat of *C. caulescens* is on the escarpment from Limpopo to Swaziland through Mpumalanga (Figure 1.2). The plant size of the *C. caulescens*, range from 500mm to 1500mm in height. The flowers of *Clivia caulescens* appear to be similar to those of *C. gardenii*, *C. nobilis*, *C. mirabilis* and *C. robusta*, but the plants differ from those species in several important aspects that justify placing it into its own taxon (Koopowitz, 2002). One of these aspects is the leaves of *C.*

caulescens have acute tips similar to *C. gardenii*, but quite different from the blunted apexes observed on *C. nobilis* leaves, and the leaves have smooth margins. The smooth, soft and pointed leaves are arching, between 35 and 70mm broad and 300 to 600mm long.

Figure 1.1: Photographs of different *Clivia* species. **A.** *C. nobilis*; **B.** *C. miniata*; **C.** *C. gardenii*; **D.** *C. caulescens*; **E.** *C. mirabilis*; **F.** *C. xnimbicola*

Flowers can be produced in immense umbels with over fifty florets. The flowers are pendulous and tubular, coloured orange-red with green tips. In the southern hemisphere,

Clivia caulescens flowers mainly in spring or summer, September to November (Swanevelder & Fisher, 2009).

Figure 1.2: Geographical distribution of *Clivia caulescens* specimens used during this study. **1**. Soutpansberg; **2.** Magoebaskloof; **3-5.** Wonderview, God's Window, Pinnacle; **6.** Mariepskop; **7.** Bearded Man.

A major feature of this species is its distinct thickened stem, about 5cm wide, which often looks like a length of short bamboo. The stem can be over one meter long in some old populations observed at God's Window (Figure 1.3). In mature specimens these stems are capable of being as tall as 3m (Swanevelder & Fisher, 2009). However the *C. caulescens* stem is not unique to this species. Elongated stems occur also in some forms of *C. robusta*, which grows in swampy conditions. This phenomenon was also observed in a few *C.*

gardenii and *C. miniata* specimens. *Clivia caulescens* normally grows on the forest floor, on rocks or on trunks of trees, as observed at Mariepskop (Figure 1.4).

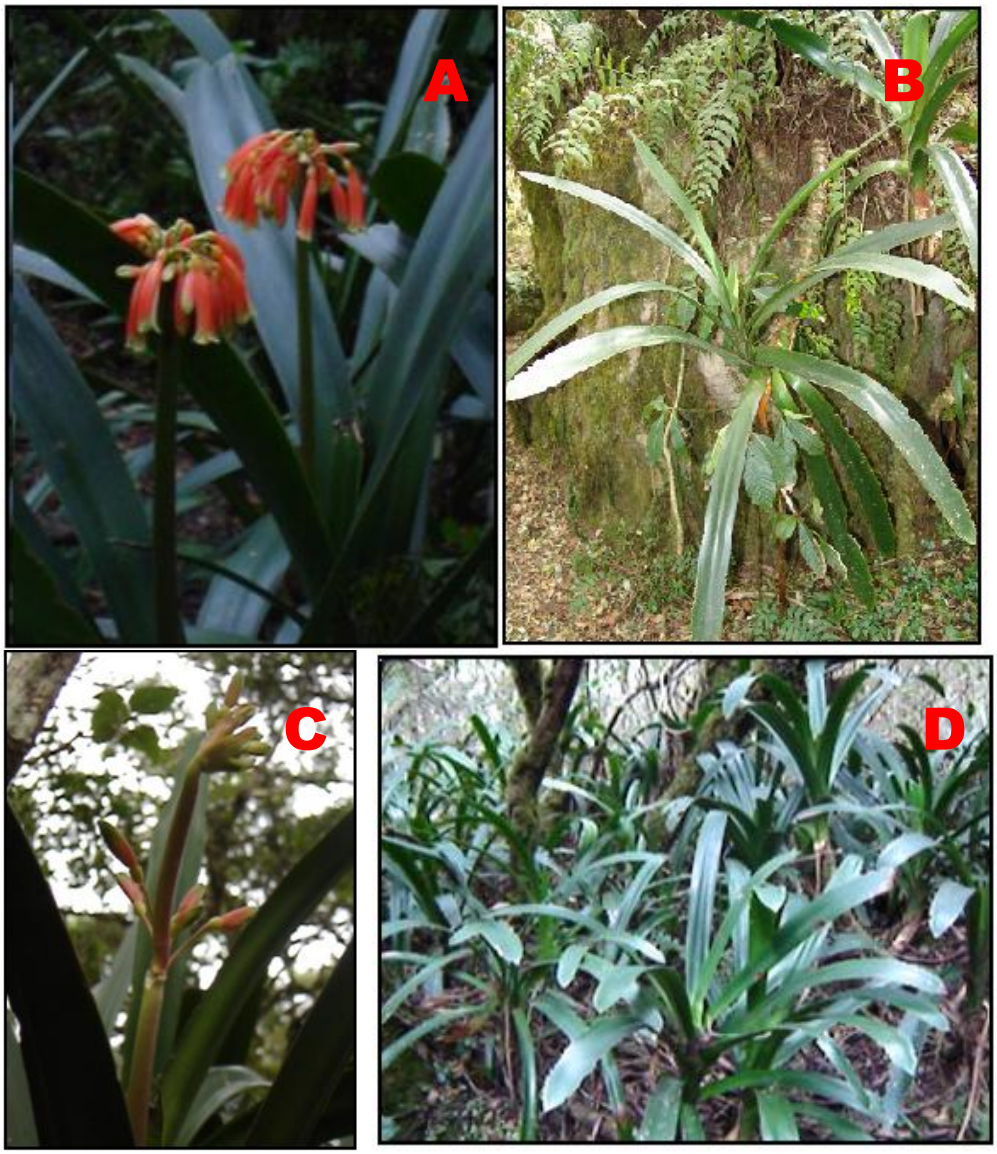


Figure 1.3: Photographs of *Clivia caulescens* in habitat at **God's Window**. **A.** Plants in flower. **B.** Plant with distinct long, thick stem. **C.** Plant with a split flower starting to bloom. **D.** Large number of plants growing together.
These photographs were taken during a visit to this population by the author.

The round to oblong red berries contains 1-4 seeds and is 9-13mm in diameter. Seeds ripen in the winter more or less six months after pollination. Seedlings tend to have erect, pale green leaves. They grow fast, and if fertilized and grown under optimal conditions, the

seedling can often be forced into flower within three or four years. It is a robust species, which establishes itself quickly.

Koopowitz (2002) claimed that, *Clivia caulescens* does not appear to be sought after for medicinal and spiritual purposes, because the populations occur in inaccessible places,

such as vertical cliffs and therefore the species is not regarded as threatened. However, an increasing number of cases are known where *C. caulescens* were confiscated by Nature Conservation from *muthi* healers.

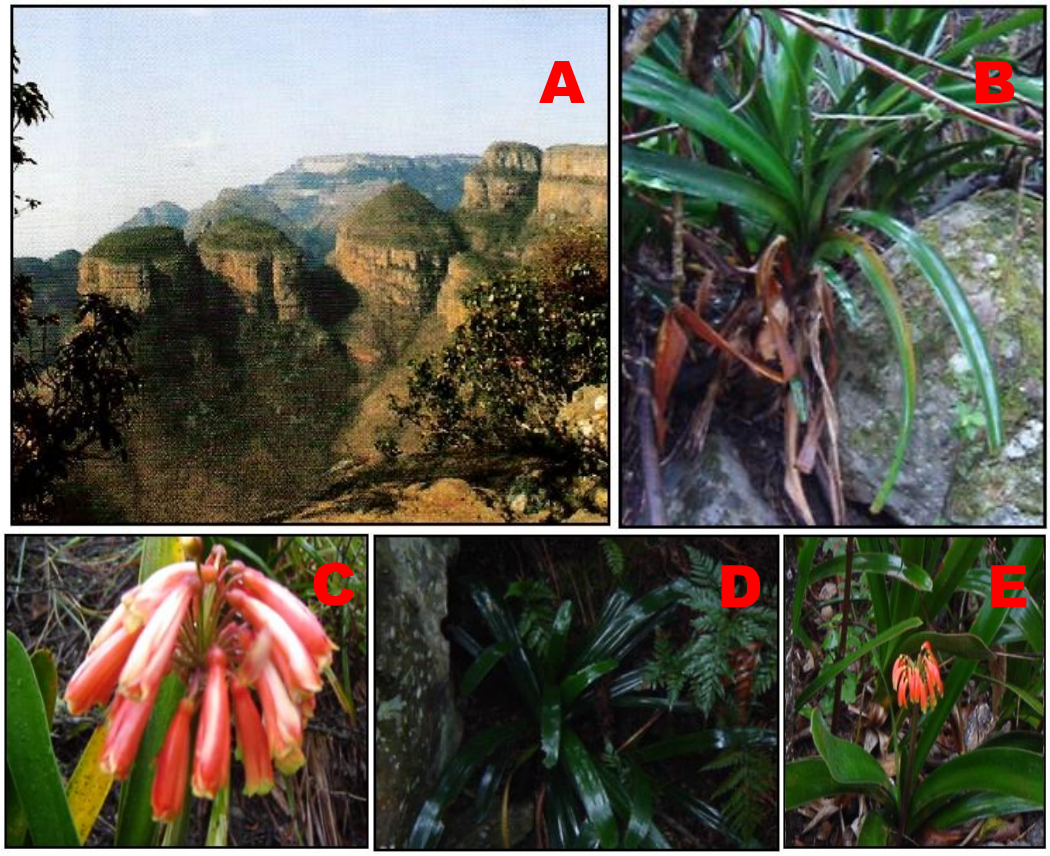


Figure 1.4: Photographs of *Clivia caulescens* in habitat at **Mariepskop**. **A.** General terrain. **B.** Plant in cleft in a rock face. **C.** Flowering umbel. **D.** Plant growing in a tree. **E.** Plant in flower. These photographs were taken during a visit to this population by the author.

1.3.2. *Clivia miniata*

At the southern end of the geographical distribution of *C. caulescens,* it overlaps with *C. miniata*. Consequently it was necessary to include the latter species in the study. *Clivia miniata* has a distribution range from the Kei River in the Eastern Cape Province, through the forested areas of KwaZulu-Natal (Koopowitz, 2002), to the southern part of Mpumalanga. Unlike the other five species, *C. miniata* has flared trumped shaped flowers which form a large umbel. The flower colour can vary from dark orange-red through clear orange to pastels and yellow, although the yellow-centred orange is the most common.

This is the most sought after of all the species, because of it decorative possibilities and thus favoured in cultivation and as a garden plant. *Clivia miniata* is a clump forming with dark green, strap shaped leaves which arise from a fleshy underground stem. The trumpet shaped flowers of brilliant orange (varies from yellow through different shades of orange to red) flowers mainly in spring (August to November) and sporadically at other times of the year. The deep green shiny leaves are a perfect foil for the masses of orange flowers (Koopowitz, 2002).

1.3.3. *Clivia xnimbicola*

The overlapping distribution between *C. miniata* and *C. caulescens* resulted in the formation of a natural hybrid between these species. The occurrences of natural hybrids between the various species are rarely recorded. Man-made hybrids between the different *Clivia* species are currently enjoying great popularity in breeding programs, mainly because of the beautiful progeny they produce – though the first hybrids were made as early as 1856 (*C. nobilis* and *C. miniata*) (Truter *et al.,* 2007). Various references to putative natural hybrids between *C. miniata* and *C. nobilis*; *C. miniata* and *C. gardenii*; and *C. miniata* and *C. caulescens* have been recorded in literature (on the border between Mpumalanga, South Africa and Swaziland) and its subsequent collection and cultivation at Kirstenbosch Botanical Gardens, South Africa (Swanevelder *et al.*, 2006). The Bearded Man Mountain marks the northern limit of *C. miniata* and the southern limit of *C. caulescens,* the only known region in which these two species occur together (Dixon, 2005).

The first and only natural hybrid ever described in the genus *Clivia* occurs in these mountains and is currently known as *Clivia xnimbicola* (Truter *et al.,* 2007). The epithet ‘nimbicola’ means ‘dweller in the mist or cloud’ and refers to the mist belt habitat in which this hybrid

and putative parents are found. The new nothospecies (botanical term to describe a naturally-occurring hybrid) is intended to cover all hybrids between *C. miniata* (including all varieties) and *C. caulescens* (Truter *et al.*, 2007).

The holotype of *C. xnimbicola* was collected on the Bearded Man Mountain, near Barberton, South Africa. In this locality *C. caulescens* grow on steep cliff faces or steep rocky embankments, whereas *C. miniata* generally grow on gentler screen embankments or flatter forest habitats. The hybrid plants are distributed between and amongst both parents, occupying both specific habitats found in the Afromontane Forest (Swanevelder *et al.*, 2006).

Flower colour range from pastel pinks through pastel oranges and deep reds, some specimens with green tepal apices. Flowering is somewhat erratic, from July to December. Some clones even flower twice yearly, the second flush occurring from February to May. In support of the taxon's hybrid origin the extended flowering period of *C. xnimbicola* is regarded as further evidence, keeping in mind that *C. caulescens* flowers October to November in the Bearded Man Mountain. The berries of the hybrid are fertile and produce seedlings that grow close to the parent plants.

1.4 SOME GENETIC STUDIES ON THE GENUS *CLIVIA*

Various genetic techniques have been used in different studies to determine the relationships and variation in the genus, but all these studies were on a small scale. Studies on the karyotypes of the genus (Ran *et al.*, 1999), confirmed the same chromosome number (2n = 2x = 22) and basic chromosome morphology for all species in the genus, *Clivia*. All the named species are cross compatible and produce vigorous, fertile progeny, suggesting a close relationship (Ran *et al.*, 2001b).

Ran *et al*. (2001a, b) used two distinct methods, namely Random Amplified Polymorphic DNA Analysis (RAPD) and DNA sequencing to detect and identify hybrids.

Meerow *et al*. (1999) resolved the cladistic relationships of the family systematic of Amaryllidaceae based on cladistic analysis of chloroplast DNA plastid *rbcL* and *trnL-F* sequence data.

Investigation of genetic relationships among the haplotypes of the different species of *Clivia* was conducted by Conrad & Snijman (2006). Networks were constructed separately for both the individual regions (*trnL-F* and the *rpoB-trnC* region) and combined data matrices were analysed.

1.5 GENETIC VARIATION

Genetic variation refers to the variation in the genetic material of a population, and includes the nuclear, mitochondrial and ribosomal genomes as well as the genomes of other organelles. A study of genetic variation is widely used to examine differences between members of the same species or to differentiate between individuals (for example in forensic analysis) (Dale & Von Schantz, 2003). Alternatively, we can compare the genetic composition of members of different species, even over wide taxonomic ranges, which can throw invaluable light on the process of evolution as well as helping to define the taxonomic relationship between species (Dale & Von Schantz, 2003).

The relative genetic diversity among individuals or populations can be determined using morphological and molecular markers. Morphological characters may be influenced by environmental factors; the developmental stage of the plant and in many plants, particularly at the seedling stage in plants including *Clivia*, morphological variation may not be adequate (Tatineni *et al.*, 1996). In contrast, molecular markers are not directly influenced by environmental effects or epistatic interactions and can provide large numbers of loci. Several methods such as isozyme analysis or restriction fragment length polymorphisms (RFLP's) have been used to investigate genetic relationships between and within different species. Methods that detect variation at the level of the DNA sequence have proved to be extremely effective tool for distinguishing between closely related genotypes (Hartl & Seefelder, 1998) and a variety of these are currently available.

1.6 MOLECULAR TECHNIQUES AVALABLE FOR TESTING GENETIC VARIATION

Since various techniques are available to determine the genetic structure and variation in populations, a short description of some of the techniques follows, but only two of those techniques, namely microsatellites and sequencing have been used during this study and will be discussed in more detail.

Restriction Fragment Length Polymorphism (Botstein *et al.*, 1980) was the principle molecular technique for identifying genetic polymorphisms in the eighties. It has several limitations which include the need for sufficient genomic DNA from each of the large number of samples to do a Southern Blot; need for a probe (short fragment of genomic DNA that has been cloned into a bacterial cell) and the need for radioactive label to achieve the most sensitive detection, although in some cases fluorescent primers can be used.

Random Amplified Polymorphic DNA (Welsch & McClelland, 1990) is a convenient method for identifying genetic polymorphisms, because this particular method does not require probe DNA and no advance information about the genome of the organism is needed. A disadvantage of RAPD's is that the method uses a set of PCR primers of 8 to 10 bases whose sequence is random, therefore resulting in random primers binding on the template. Another disadvantage is that results are not reproducible in other laboratories.

The **Amplified Fragment Length Polymorphism** technique (Vos *et al.*, 1995) is based on the detection of genomic restriction fragments by PCR amplification, and can be used for DNAs of any origin or complexity. Fingerprints are produced without prior sequence knowledge using a limited set of generic primers. The AFLP technique is robust and reliable, because stringent reaction conditions are used for primer annealing. The alleles supporting amplification of AFLP fragments are dominant, which means that a single + allele is sufficient to support amplification, and so homozygous +/+ and heterozygous +/- genotypes cannot be distinguished, which is a disadvantage.

Microsatellites have emerged as one of the most popular choices for these studies in part because they have the resolving power to distinguish relatively high rates of migration from panmixia, have the potential to provide contemporary estimates of migration and can estimate the relatedness of individuals (Selkoe & Toonen, 2006).

Microsatellites are tandem repeats of 1-6 nucleotides found at high frequency in the nuclear genomes of most taxa, and are also known as simple sequence repeats (SSR), variable number tandem repeats (VNTR) and short tandem repeats (STR) (Selkoe & Toonen, 2006).

A microsatellite locus typically varies in length between 5 and 40 repeats, but longer strings of repeats are possible. Dinucleotides, trinucleotide and tetranucleotide repeats are the most common choices for molecular genetic studies. Dinucleotide repeats account for the majority of microsatellites for many species (Li *et al.*, 2002). Trinucleotide and hexanucleotide repeats are the most likely repeat classes to appear in coding regions, because they do not cause a frame shift (Toth *et al.*, 2000). Mononucleotide repeats are less reliable because of problems with amplification; longer repeat types are less common and fewer data exist to examine their evolution (Li *et al.*, 2002).

Primers can be designed to bind to the flanking region and guide the amplification of a microsatellite locus with PCR. A given pair of microsatellite primers rarely works across broad taxonomic groups, and so specific primers are usually developed for each species (Glenn & Schable, 2005). The process of isolating new microsatellite markers has become faster and less expensive, which substantially reduces the failure rate or cost. There are however several disadvantages of using microsatellites and the first being that unclear mutational mechanisms can be complex and the frequency and effects are usually low. Another disadvantage is the occurrence of hidden allelic diversity, but one can make use of a microsatellite screening protocol to overcome this problem. Thirdly there can be problems with amplification, because consistent amplification across all samples can only be assured by trial and error (Selkoe & Toonen, 2006).

The major advantages of microsatellites in this study are that microsatellites allow population genetic parameters to be estimated using alleles at anonymous nuclear loci; the allelic composition of individuals within a population can be assessed using PCR and microsatellites which are species-specific, thus cross contamination by non-target organisms are much less of a problem compared with techniques that employ universal primers, such as AFLPs. Although RAPDs and AFLPs are also multilocus, none of them have the resolution and power of a multilocus microsatellite study (Selkoe & Toonen, 2006).

Using template DNA from populations shown to be genetically different, Swanevelder (2003) developed microsatellites for *Clivia miniata.* Plants used were from Broedershoek farm,

Donkeni, Kentani area, Mzamba River, Oribi Gorge, Port St. Johns and Umtamvuna River in South Africa. Swanevelder (2003) developed 4 primer sets and observed polymorphisms between samples from different localities for primer sets, *CLV*2 and *CLV*4. The other two marker sets, *CLV*1 and *CLV*3, showed no polymorphism between different *C. miniata* localities sampled. Swanevelder (2003) proposed that these might still be useful in studies of other *Clivia* species.

The attention of plant molecular biologists were also attracted by **Single Nucleotide Polymorphisms** (SNPs) during the last few years (Gupta *et al*., 2001). A SNP is the polymorphism occurring between DNA samples with respect to a single base (Jehan & Lankhanpaul, 2006) or otherwise described as variation at a single nucleotide position (Liu & Cordes, 2004; Strachan & Read, 2004). This variation is usually the result of a point mutation (Liu & Cordes, 2004), for example deletion, insertion, or a single nucleotide being substituted by a different nucleotide (Fairbanks & Andersen, 1999). SNPs are less mutable compared to other markers, particularly microsatellites (Jehan & Lankhanpaul, 2006). They may be found both in the non-repetitive coding or regulatory sequences and in the repetitive non-coding sequences (Gupta *et al*., 2001). The international SNP map working group constructed a map of human genome sequence variation containing 1.42 million SNPs i.e. one SNP per 1.9 kb^4 (Gupta *et al*., 2001; Jehan & Lankhanpaul, 2006), to give more insight into genetic variation of humans. In plants, SNPs are found to be present in high density across the genome. In the wheat genome one SNP per 20 bp and in the maize genome, one SNP per 70 bp has been observed in some regions (Jehan & Lankhanpaul, 2006). SNP analysis could be pivotal in the study of genetic variation of *C. caulescens* and sequence analysis is the most direct way of identifying SNPs.

DNA sequencing by DNA polymerase chain reaction termination was introduced by Fred Sanger (Sanger *et al*., 1977) in 1977. In 1993, DNA sequencing studies already accounted for about 50% of all molecular systematic investigations (Sanderson *et al*., 1993). DNA sequencing is considered to be more powerful for evolutionary studies than physiological and morphological data. Firstly, protein and DNA sequences can provide a clearer picture of relationships between organisms independent of physiological and morphological characters. Secondly, statistical and mathematical theories have already been developed for analysing DNA sequence data. Thirdly, molecular data are more abundant. Traditional means of evolutionary enquiry, such as anatomy, morphology, palaeontology and physiology should

not be abandoned all together. Different approaches provide complementary data. Morphological and anatomical data are necessary for constructing a time frame for evolutionary studies (Olmstead & Palmer, 1994; Soltis & Soltis, 1998).

DNA sequencing provides a means for direct comparison (Olmstead & Palmer, 1994). With the advent of PCR technology, DNA sequencing has rapidly become a major source of comparative molecular data. A pragmatic look at DNA sequencing in plant phylogenetic studies have been reported by a number of DNA sequencing studies in plants (Olmstead & Palmer, 1994; Bayer & Starr, 1998; Fennel *et al*., 1998; Meerow *et al*., 1999; Molvray *et al*., 1999; Fay *et al*., 2000; Meerow & Clayton, 2003; Montero-Castro *et al*., 2006). The identification of easily amplifiable and relatively rapid evolving but clearly alignable DNA regions that can provide adequately suitable variation within short sequence segments, is the primary challenge in using nucleotide characters for lower-level phylogenetic studies (Baldwin *et al*., 1995). Some of the criteria and reasons that should be kept in mind for the choice of sequences as the primary data for classification are:

- The sequence should be of adequate length to provide enough phylogenetic informative nucleotide positions. In addition, it is necessary that the rate of sequence divergence be appropriate to the phylogenetic question being addressed. A short sequence with a high substitution rate will not necessarily be comparable to a long sequence with a low substitution rate because the chance of a substitution along a branch of a tree must be relatively low for parsimony to succeed (Olmstead & Palmer, 1994).

- Sequences must be readily aligned. Sequence alignment is essential for correct assessment of character homology.

- Sequences must have evolved orthologous. A severe problem with the phylogenetic analysis of many nuclear genes is distinguishing orthology (genes derived from a speciation event) from paralogy (genes related by gene duplication within a genome). As long as these genes remain within the chloroplast genome, this is not a problem with chloroplast genes, which all evolved as single-copy genes (Olmstead & Palmer, 1994; Soltis & Soltis, 1998).

1.6.1. Chloroplast regions used in this study

The two primary sources of molecular variation tapped for analyses purposes have been the chloroplast genome (cpDNA) and ribosomal DNA repeat regions (Olmstead & Palmer, 1994).

The chloroplast genome in plants and mitochondrial genome in animals are natural counterparts in the phylogenetic study of their respective groups. The chloroplast genome has provided useful intraspecific variation in some, but not all, species (Taberlet *et al*., 1991). In chloroplast genomes, gene orders are highly conserved (Demesure *et al*., 1995; Dumolin-Lapegue *et al*., 1997 and Hamilton, 1999), whereas some spacers show even intra-species variation. Amplified fragments can be analysed by restriction analysis or DNA sequencing.

There are many genes in the chloroplast genome that are widespread and sufficiently large (> 1 kb) to be generally useful in comparative sequencing studies.

These genes are suitable for a wide range of taxonomic levels and encompass a wide range of evolutionary rates (Olmstead & Palmer, 1994). Chloroplast genes are unlikely to be functionally correlated in their evolution, as they code for diverse functions such as photosynthesis, respiration and transcription. The strategy of comparative sequencing will yield two sets of data that are relatively free of functional correlations, but all cpDNA sequences exhibit the characteristic of being inherited as a single linkage group (Olmstead & Palmer, 1994).

The cpDNA from tobacco (*Nicotiana tabacum*) has often served as a reference for plastid genomes (Wakasugi *et al*., 1998). In 1986, the complete nucleotide sequence and gene map was published (Shinozaki *et al*., 1986). Wakasugi *et al.* (1998) constructed the updated gene map which includes 105 different genes. The genome size is 155,939 basepairs (bp). It consists of 86,686 bp of a large single copy region (LSC), 18,571 bp of a small single-copy region (SSC) and two inverted repeats (IR) of 25,341 bp each.

Molecular systematicists have utilized PCR-amplified chloroplast gene sequences for establishing and verifying phylogenies, starting off with the highly conserved *rbcL* gene, and later expanding to e.g. *matK, ndhF, rpH*6 and *atpB* (Heinze, 2007).

Markers of choice must exhibit sufficient variability to link species and groups of species by possessing shared (synapomorhic) substitutions. Unique substitutions (autoapomorphic) are not used in assessing phylogenetic relationships of species and other taxa (but note that they are used in dating of phylogenetic trees, i.e. in molecular clock studies, and establishing overall genetic distances between species) (Chase *et al*., 2005). The DNA barcoding initiative taken by Kew Botanical Gardens in the United Kingdom led to a project where the Consortium for the Barcode of Life (CBOL) aim to give every land plants a barcode (CBOL Plant Working Group, 2009). Three plastid regions were chosen, *matK*, *rpoC1* and *rpoB*, for our sequencing purposes, from their study. These primers can be used to study inter- and probably intraspecific phylogenies of plants because they amplify cpDNA non-coding regions over a wide taxonomic range (Taberlet *et al.*, 1991). Small insertions or deletions (also referred to as INDELS) are relatively frequent, when compared to point mutations that result in restriction site changes. In general the exon sequences are highly conserved, but this depends on the gene in question. Molecular systematicists have utilized PCR-amplified chloroplast gene sequences for establishing and verifying phylogenies (Heinze, 2007).

1.6.1.1 *matK* region

RbcL gene sequences are often employed in plant phylogenetic analysis, but the evolutionary rate of this gene is considered too slow to resolve the lower level phylogeny of angiosperms (Chase *et al*., 1993). Therefore the *matK* gene (Figure 1.5) is used in this study, which is also located in the chloroplast genome and has a faster evolutionary rate than the *rbcL* gene (Olmstead & Palmer, 1994; Johnson & Soltis, 1994; Steele & Vilgalys, 1994; Nakazawu *et al*., 1997). Given *matK*'s adequate rate of variation, easy amplification and alignment, a portion of the plastid *matK* gene has been identified as a universal DNA barcode for flowering plants (Lahaye, 2008). Ito *et al*., (1999) resolved a monophyletic *Haemantheae* by using plastid *mat*K, with a 98% bootstrap support. The *matK* gene is located in the large single-copy region of the chloroplast genome (Soltis & Soltis, 1998). It is a chloroplast intron-specific maturase of higher plants and might have a function in splicing of multiple introns (Vogel *et al.*, 1999).

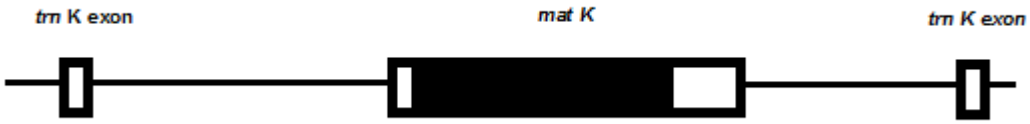


Figure 1.5: Diagram of the *matK* region in the intron region of *trn*K gene (Ito *et al*., 1999). (Not to scale)

1.6.1.2 *atpH-I* region

The ATP synthase complex, occurring in the plastid, consists of nine subunits; six of which are encoded in the plastome. One of the two transcriptional units of the plastid-encoded genes is known as *atpI/H/F/A* (Miyagi *et al.*, 1998). The combination name for the noncoding spacer region is *atpH-I* (Figure 1.6), consisting out of *atpH-P* and *atpI-M*. *atpH* is identified as the F1 sector of membrane-bound ATP synthase, delta subunit (Miyagi *et al.*, 1998), it lies on the inside of the chain and is transcribed clockwise. *atpI* is known as the ATP synthase, membrane-bound accessory subunit IV (Miyagi *et al.*, 1998). *atpI* is situated on the LSC on position 16,000 bp to 15,257 bp of the 86,686 bp region and is transcribed clockwise (Heinze, 2007).

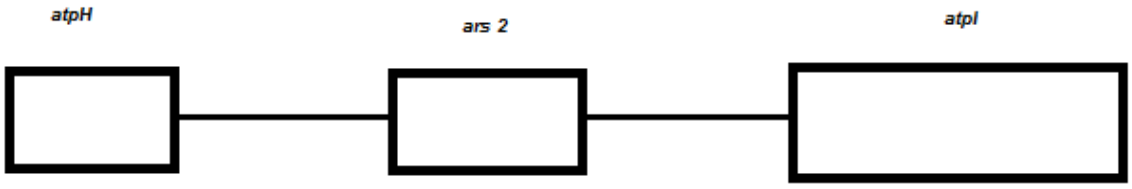


Figure1.6: Diagram of the *atpH-I* region. The region of *ars2* is located between *atpH* and *atpI* (Miyagi *et al.*, 1998). (Not to scale)

1.6.1.3 *rpoB* region

The RNA polymerase beta-subunit-encoding gene, *rpoB* (Figure 1.7) is known as a coding gene (CBOL Plant Working Group, 2009) and lies on the inside of the chain of the cpDNA and is transcribed clockwise. It lies on position 27,511bp to 24,299bp of the LSC (86,686bp), downstream from the split gene *rpoC1* 3' exon intron 738 bp 5' exon (Wakasugi *et al.*, 1998).

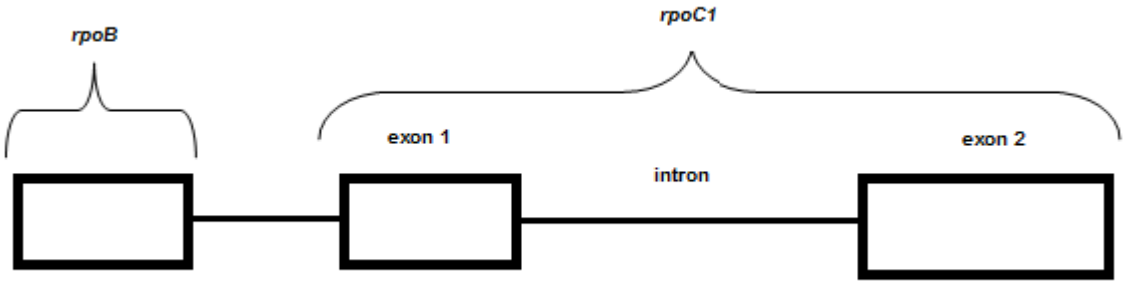


Figure 1.7: Diagram indicating that regions *rpoB* and *rpoC1* are situated adjacent. The region of *rpoC1* comprises of two exons and one intron (Heinze, 2007). (Not to scale)

1.6.1.4 *rpoC1* region

The *rpoC1* (Figure 1.7) gene is known as the RNA polymerase beta' subunit 3' exon, intron 738bp, 5' exon and lies on the inside of the chain cpDNA. The *rpoC1* gene is transcribed clockwise. It is situated on position 23 102bp to 21 486bp (RNA polymerase beta' subunit 3' exon); 23 840bp to 23 103bp (intron 738bp) and from 24 293bp to 23 841bp (5' exon) (Wakasugi *et al.*, 1998).

1.6.1.5 *trn*L-F

The non-coding *trn*L-F (Figure 1.8) region of the chloroplast genome, consist of the *trn*L-(UAA)-intron, and intergenic spacer (IGS) between the *trn*L-(UAA)-3'-intron and *trn*F-(GAA) gene (Pfosser & Speta, 1999). In addition, the *trnT-L* region was also sequenced, an intergenic spacer between *trnT* (UGU) and *trnL* (UAA) 5' exon (Taberlet *et al*., 1991).

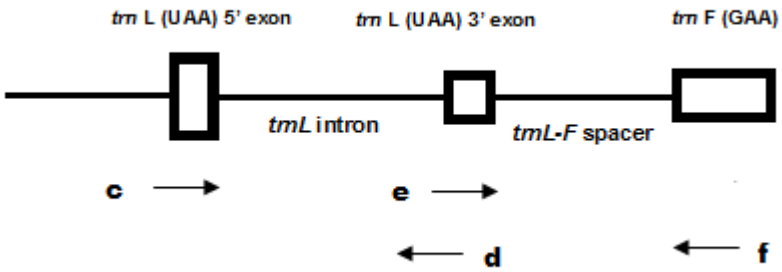


Figure 1.8: Diagram of the *trn*L-F region, which contains an intron and intergenic spacer. The directions of the four universal primers are indicated by the arrows (Taberlet *et al*., 1991). (Not to scale)

1.7 DISSERTATION OUTLINE

This dissertation is presented as a series of individual papers; therefore the references are listed at the end of each chapter. In an attempt to avoid unnecessary duplication, cross referencing between different chapters was occasionally used. The format of the chapters is roughly according to the layout of *Philosophical Transactions in Genetics*.

Genetic variation in the different *Clivia caulescens* populations is the focus of **Chapter 2. Chapter 3** deals with the two species and their natural hybrid growing on the Bearded Man Mountain. The use of cross-species markers in *Clivia caulescens* is discussed in **Chapter 4**. A general discussion and conclusion is presented in **Chapter 5**. To conclude, **Chapter 6** consists of a summary of the whole dissertation.

1.8 AIMS OF THE STUDY

The aims of this study were to determine:

- Techniques:
  - whether SNPs can be used in *Clivia* to determine the degree of genetic variation

- whether microsatellites can be used in *Clivia* to determine the degree of genetic variation

- Practical applications:
  - what genetic variation exists within *C. caulescens* populations
  - what the genetic variation between different *C. caulescens* populations is
  - whether gene flow occurs between the different localities
  - whether molecular markers can be used to identify the geographical origin of a specimen
  - what the genetic variation in the *C. xnimbicola* population is
  - the correlation between the genetic variation in *C. xnimbicola* and the two parental species, *C. caulescens* and *C. miniata,* in the Bearded Man populations
  - whether *C. xnimbicola* is continuously formed by random pollination events.

1.9 REFERENCES

BALDWIN, B. G., SANDERSON, M. J., PORTER, J. M., WOJCIECHOWSKI, M. F., CAMPBELL, C. S. & DONOGHUE, M. J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of Missouri Botanical Garden* **82**: 247-277.

BAYER, R. J. & STARR J. R. 1998. Tribal phylogeny of the Asteraceae based on two non-coding chloroplast sequences the *trn*L and the *trn*L/ *trn*F intergenic spacer. *Annals of the Missouri Botanical Garden* **85**: 242-256.

BOTSTEIN, D., WHITE, R. L., SKOLNICK, M. & DAVIS, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314–331.

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

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

CHASE, M. W., SALAMIN, N., WILKINSON, M., DUNWELL, J. M. KESANAKURTHI, R. P., HAIDAR, N. & SAVOLAINEN, V. 2005. Land plants and DNA barcodes: short term and long-term goals. *Philosophical transactions of the royal society*, published online. Retrieved article in October, 2009.

CONRAD, F. & SNIJMAN, D. 2006. Systematics and phylogeography of *CLIVIA*. South African National Biodiversity Institute, Cape Town.

DALE, J. W. & VON SHANTZ, M. 2002. *From Genes to Genomes: Concepts and Applications of DNA Technology*. John Wiley and Sons Ltd., Baffins Lane, England

DEMESURE, B., SODZI, N. & PETIT, R. J. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**: 129-131.

DIXON, R. 2005. Have genes, will travel on the trail of 'Vico Yellow'. *Clivia Yearbook* **7**: 78-81.

DUMOLIN-LAPENGUE, S., PEMONGE, M. H. & PETIT, R. J. 1997. An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology* **6**: 393-397.

DUNCAN, G. & DU PLESSIS, N. 1989. Bulbous plans of Southern Africa, a guide to their cultivation and propagation. Tafelberg. Cape Town.

DYER, R. A. 1943. *Clivia caulescens*. *Flowering plants of South Africa* **23**: 891.

FAY, M. F., RUDALL, P. J., SULLIVAN, S., STOBART, K. L., DE BRUIJN, A. J., REEVES, G., QUANARUS-ZAMAN, F., HANG, W-P., JOSEPH, J., HAHN, W. J., CONRAN, J. G. & CHASE, M. W. 2000. Phylogenetic studies of Asparagales based on four plastid DNA regions. *In* K.L. Wilson & D.A. Morrison (Eds). *Monocots: Systematics and evolution*, 360-371, CSIRO, Melbourne.

FAIRBANKS & ANDERSON. 1999. Genetics: The Continuity of Life. Brooks/Cole Publishing Company, Pacific Grove, California.

FENNEL, S. R., POWELL, W., WRIGHT, F., RAMSAY, G. & WAUGH, R. 1998. Phylogenetic relationships between *Vicia faba* (Fabaceae) and related species inferred from chloroplast *trn*L sequences. *Plant Systematics and Evolution* **212**: 247-259.

GERMISHUIZEN, G. & MEYER, W. L. 2000. Plants of Southern Africa: an annotated checklist. *Sterlitzia* **14**: 1050-1051.

GLENN, T. C. & SCHABLE, N. A. 2005. Isolating microsatellite DNA loci. *In* Zimmer, E.A. & Roalson, E. *Molecular Evolution: Producing the Biochemical Data, Part B* [eds.], 202-222, Academic Press, San Diego, USA.

GUPTA, K. P., ROY, J. K. & PRASAD, M. 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science* **80**: 4.

HAMILTON, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**: 521-523.

HARTL, L. & SEEFELDER, S. 1998. Diversity of selected hop cultivars detected by fluorescent AFLPs. *Theoretical and Applied Genetics* **96**: 112–116.

HEINZE, B. 2007. A database of PCR primers for the chloroplast genomes of higher plants. *Plant Methods* **3**: 4.

ITO, M., KAWAMOTO, A., YOKO, K., TOMOHISA, Y. & KURITA, S. 1999. Phylogenetic relationships of Amaryllidaceae based on *matK* sequence data. *Journal of Plant Research* **112**: 207-216.

JEHAN, T. & LAKHANPAUL, S. 2006. Single nucleotide polymorphism (SNP) – Methods and applications in plant genetics: A review. *Indian Journal of Botany* **5**: 435-459.

JOHNSON, L. A. & SOLTIS, D. E. 1994. *matK* DNA sequences and phylogenetic reconstructions in *Saxifragaceae*. *Systematics Botany* **19**: 143-156.

KOOPOWITZ, H. 2002. *Clivia.* Timber press, Portland, Oregon, Cambridge.

LAHAYE, R. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of National Academy of Sciences USA* **105**: 2923-2928.

LI, Z. J. & CORDES, J. F. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **238**: 1-37.

LI, Y. C., KOROL, A. B., FAHIMA, T., BEILES, A. & NEVO, E. 2002. Microsatellites: genomic distribution, putative functions, and mutational mechanisms: a review. *Molecular Ecology*., **11**: 2453-2465.

MEEROW, A. W. 1995. Towards the phylogeny of Amaryllidaceae. *In* J. Rudall, P.J. Cribb, D. F. Cutler & C. J. Humphries [eds.], *Monocotyledons: systematic and evolution*, 169-179, Royal Botanic Gardens, Kew.

MEEROW, A. W. & CLAYTON, J. R. 2003**.** Generic relationships among the baccate-fruited Amarillidaceae (tribe Haemanthea) inferred from plastid and nuclear non-coding DNA sequences. *Plant Systematics and Evolution* **244**: 144-155.

MEEROW, A. W., FAY, M. F., GUY, C. L., LI, Q., ZAMAN, F. Q. & CHASE, M. W. 1999. Systematics of Amaryllidaceae based on cladistic analysis of plasmid *rbcL* and *trnL-F* sequence data. *American Journal of Botany* **86**: 1325-1345.

MEEROW, A. W., GUY, C. L., LI, Q. & YONG, S. 2000. Phylogeny of the American Amaryllidaceae based on nrDNA ITS sequences. *Systematic Botany* **25**: 708-726.

MIYAGI, T., KAPOOR, S., SUGITA, M. & SUGIRA, M. 1998. Transcript analysis of the tobacco plastid operon rps2/atpl/H/F/A/ reveals the existence of a non-consensus II (NCII) promoter upstream of the *atpI* coding sequence. *Molecular and General Genetics* **257**: 299-307.

MOLVRAY, M., KORES, P. J. & CHASE, M. W. 1999. Phylogenetic relationships within Korthalsella (Viscaceae) based on nuclear ITS and plastid *trn*L-F sequence data. *American Journal of Botany* **86**: 249-260.

MONTERO-CASTRO, J. C., DELGADO-SALINAS, A., DE LUNA, E. & EGUIARTE, L. E. 2006. Phylogenetic analysis of *Cestrum* section *Habrothamnus* (Solanaceae) based on plastid and nuclear DNA sequences. *Systematic Botany* **31**: 843-850.

MURRAY, B. G., RAN, Y. E., LANGE, P. D., HAMMET, K. R. W., TRUTER, J. T. & SWANEVELDER, Z. H. 2004. A new species of *Clivia* (Amaryllidaceae) endemic to the Pondoland Centre of Endemism, South Africa. *Botanical Journal of the Linnean Society* **146**: 369-374.

NAKAZAWA, M., WAKABAYASHI, M., ONO, M. & MURATA, J. 1997. Molecular phylogenetic analysis of *Chysosplenium* (Saxifragaceae) in Japan. *Journal of Plant Research, Springer Japan* **10**: 265-274.

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

PFOSSER, M. & SPETA, F. 1999. Phylogenetics of Hyacinthaceae based on plastid DNA sequences. *Annals of the Missouri Botanical Garden* **86**: 852-875.

RAN, Y., HAMMETT, K. R. W. & MURRAY, B. G. 1999. Karyotype analysis of the genus *Clivia* by Giemsa and fluorochrome banding and *in situ* hybridization. *Euphytica* **106**: 167-174.

RAN, Y., HAMMETT, K. R. W. & MURRAY, B. G. 2001a. Phylogenetic analysis and karyotype evolution in the genus *Clivia* (Amaryllidaceae). *Annals of Botany* **87**: 823-830.

RAN, Y. MURRAY, B. G. & HAMMETT, K. R. W. 2001b. Evaluating genetic relationships between and within *Clivia* species using RAPDs. *Scientia Horticulturae* **90**: 167-179.

REES, A. R. 1992. Ornamental bulbs, corms and tubers. Wallingford, Oxford.

ROURKE, J. P. 2002. *Clivia mirabilis* (Amaryllidaceae: Haemantheae) a new species from Northern Cape, South Africa. *Bothalia* **32**: 1-7.

SANDERSON, M. J., BALDWIN, B. G., BHARATHAN, G., CAMPBELL, C. S., VON DOHLEN, C. FERGUSON, D., PORTER, J. M. WOJCIECHOWSKI, M. F. & DONOGHUE, M. J. 1993. The growth of phylogenetic information and the need for a phylogenetic data base. *Systematic Biology* **42**: 562-568.

SANGER, F., NICKLEN, S. & COULSAN, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Biotechnology* **24**: 104-108.

SELKOE, K. A. & TOONEN, R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**: 615-629.

SHINOZAKI, K., OHME, M., TANAKA, M., WAKASUGI, T., HAYASHIDA, N., MATSUBAYASHI, T., ZAITA, N., CHUNWONGSE, J., OBOKATA, J., YAMAGUCHI-SHIOZAKI, K., OHTO, C., TORAZAWA, K., MENG, B. Y., SUGITA, M., DENO, H., KAMOGASHIRA, T., YAMADA, K., KUSUDA, J., TAKAIWA, F., KATO, A., TOHDOH, N., SHIMADA, H. & SUGIURA, M. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *Plant Molecular Biology Reporter* **4**: 110-147.

SOLTIS, K. P. & SOLTIS, P. S. 1998. Choosing on approach and on appropriate gene for phylogenetic analysis. *In* D. E. Soltis, P.S. Soltis, & J.J. Doyle [eds.], *Molecular systematic of plants* II, 1-86, Kluwer Academic Publisher, Boston, Dordrecht, London.

STEELE, K. P. & VILGALYS, R. 1994. Phylogenetic analyses of Plemoniaceae using nucleotide sequences of the plastid gene *matK*. *Systemic Botany* **19**: 126-142.

STRACHAN, T. & READ, A. P. 2004. *Human molecular genetics.* (3rd ed.). Garland Science, New York, USA., pp. 635, 639, 641.

SWANEVELDER, Z. H. 2003. Diversity and population structure of *Clivia miniata* Lindl. (Amaryllidaceae): Evidence from molecular genetics and ecology. MSc Dissertation, University of Pretoria.

SWANEVELDER, Z. H., TRUTER, J. T. & VAN WYK, A. E. 2006. A natural hybrid in the genus *Clivia*. *Bothalia* **36**: 77-80.

SWANEVELDER, Z. H. & FISHER, R. 2009. *Clivia* – Nature and Nurture. *Briza Publications*, Pretoria, South Africa.

TABERLET, P., GIELLY, L., PAUTOU, G. & BOUVET, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1109.

TATINENI, V., CANTRELL, R. G. & DAVIS, D. D. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPD. *Crop Sciences* **36**: 186-192.

TOTH, G., GASPARI, Z. & JURKA, J. 2002. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* **10**: 967-981.

TRUTER, J. T., SWANEVELDER, Z. H. & PEARTON, T. N. 2007. *Clivia x nimbicola* – a Stunning Beauty from the Bearded Man. *Clivia Yearbook* **8**: 23-27.

WAKASUGI, T., SUGITA, M., TSUDZUKI, T. & SUGIURA, M. 1998. Updated Gene Map of Tobacco Chloroplast DNA. *Plant Molecular Reporter* **16**: 231-241.

WELSCH, J. & McCLELLAND, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18**: 7213-7218.

WIESING, K. & GARDNER, R. C. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledenous angiosperms. *Genome* **42**: 9-19.

VOGEL, J., BORNER, T. & HESS, W. 1999. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Research* **27**: 3866-3874.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, VAN DER LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. & ZABEAU, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407-4414.

Chapter 2:

GENETIC VARIATION IN *CLIVIA CAULESCENS*

Suzanne Stegmann

2.1 ABSTRACT

Clivia caulescens grow on the edge of the escarpment and rocky outcrops in Mpumalanga. Although populations appear heterogeneous, the extent of gene flow between different populations has not been studied. The number of people arrested for illegal trafficking with this species increase. The aim of this study was consequently to study the genetic variation within populations, as well as the genetic variation between different populations. Thus the extent of gene flow can be determined and possible diagnostic markers linked to specific geographical areas can be identified for forensic purposes.

Five primer sets for the amplification of non-coding regions of chloroplast DNA (*atpH-1*, *trnT-L*, *matK*, *rpoB* and *rpoC1*) were used during this study. The results indicated that some genetic variation occurred in all populations and that the populations could be grouped into three large populations, i.e. the northern (Soutpansberg, Magoebaskloof and Wolkberg), southern (Gods Window, Bearded Man and Swaziland) and Mariepskop populations. Limited gene flow did occur between the different geographical areas. The gene flow from Mariepskop to the escarpment is very restricted.

2.2 INTRODUCTION

The genus *Clivia* currently consists of six species, which include *Clivia nobilis* Lindl., *C. miniata* (Lindl.) Regel, *C. gardenii* Hook, *C. caulescens* RA Dyer, *C. mirabilis* Rourke and *C. robusta* Murray, Ran, De Lange, Hemmet, Truter & Swanevelder. These species are exploited by the *muthi* market and unscrupulous collectors of the genus. In an attempt to get an unique DNA fingerprint for individual localities, this study focused on various *C. caulescens* localities along its natural habitat on the escarpment from Limpopo to Swaziland through Mpumalanga.

In an attempt to identify the different species of *Clivia* and determine if genetic erosion is present in *C. caulescens*, specimens of different populations have been obtained. These specimens were studied to determine the genetic variation between the different populations and within each population. Two primary sources of measuring genetic variation in plants have been the chloroplast genome (Palmer, 1987; Palmer *et al*., 1988; Olmstead *et al*., 1990;

Clegg & Zurawski, 1990) and the nuclear ribosomal DNA region (Knaak *et al*., 1990; Baldwin, 1992; Hamby & Zimmer, 1992.). The mitochondrial genome playes a major role in animal systematics (Moritz *et al*., 1987; Avise, 1991), but has limited use in plants (Palmer, 1992). In chloroplast genomes, gene order is highly conserved (Demesure *et al*., 1995; Dumolin-Lapegue *et al*., 1997; Hamilton, 1999), whereas some spacers even show intra-species variation. An effective marker system should yield the maximum number of polymorphisms for the specific germplasm sampled in terms of fragments amplified per assay, percentage of polymorphic fragments per assay unit and number of unique profiles generated (McGregor *et al*., 2000).

Five chloroplast DNA regions, i.e. *atpH-I*, *matK*, *rpoB*, *rpoC1* and *trnL-F*, were used in an attempt to study the molecular diversity of *C. caulecsens*. This study concentrated on Single Nucleotide Polymorphisms (SNPs) from these regions to study genetic variation. These are polymorphisms based on a single nucleotide difference between different specimens (Rafalski, 2001; Jehan & Lankhanpaul, 2006). Since these areas are all chloroplast regions, no heterozygotes can influence results and each marker represents a haplotype (Niu, 2004).

The aim of this study was to determine the genetic variation between and within the different populations of *C. caulescens*, to determine whether gene flow occur between the different populations and to determine which of the DNA regions included in the study can contribute to the identification of plants from a specific geographical area.

2.3 MATERIALS AND METHODS

2.3.1 Specimens used and DNA extraction

Leaf material was collected from the natural habitat or obtained from several *Clivia* breeders. Material from 20 *C. caulescens* specimens were collected, representing eight localities as well as additional specimens of the other *Clivia* species as outgroups (Table 2.1).

The extractions were based on a method described by Rogstad (1992), with a few modifications. DNA was extracted from leaves that were either fresh, or stored in either CTAB (Hexadecyltrimethyl Ammonium Bromide) or silica gel.

Prior to the extraction, CTAB (3% m/v) and 0.2% (v/v) 2-Beta-mercapto-ethyl-alcohol was added to the extraction buffer (pH 8) (100 mM Tris-HCL; 25 mM EDTA; 1.4 M NaCl) and preheated to 65°C. Approximately 1 g of plant material was cut into a mortar.
Purified sand (0.2 g), as well as 2 ml of the preheated extraction buffer, was added to the leaf material. The mixture was grounded with a pestle and mortar, until it formed a paste. Preheated extraction buffer was added, aiding in the transfer of the paste to a test tube.

Table 2.1: List of the plant specimens included in this study, indicating their geographical origin and voucher numbers. All voucher material is housed in the Geo Potts Herbarium (BLFU).

Species	Locality	Voucher number
C. caulescens	Bearded Man Mountain	*Spies 8565, 8567, 8569, 8571, 8701*
	God's Window	*Spies 8479, 8480, 8481, 8482, 8483*
	Magoebaskloof	*Spies 8893*
	Mariepskop	*Spies 8494, 8495, 8496*
	Soutpansberg	*Spies 8640*
	Swaziland	*Spies 8644, 8757, 8784*
	Wolkberg	*Spies 8487*
	Wonderview	*Spies 8596*
C. gardenii	Greytown	*Spies 8418*
C. miniata	Dwesa	*Spies 8574*
C. mirabilis	Donkerhoek	*Spies 8267*
C. nobilis	Keiskamma	*Spies 8254*
C. robusta	Port Shepstone	*Spies 8440*
C. xnimbicola	Bearded Man Mountain	*Spies 8578*

The leaf material-extraction buffer mixture was incubated in a preheated water bath at 65°C for 30 minutes. Longer times in the water bath could result in the denaturation of the DNA. The test tubes were vortexed briefly every 10 minutes. One volume of chloroform: isoamylalcohol (24:1 v/v) was added after 30 minutes and thoroughly mixed. The test tubes were centrifuged for 10 minutes at 10 000 *g* and the supernatants were transferred to a clean test tube. The DNA was precipitated with cold (-20°C) absolute ethyl-alcohol with 3 M sodium acetate (25:1) for at least 60 minutes at -20°C. The test tubes were centrifuged at 10 000 g for 10 minutes, the supernatant discarded and the pellet washed with 70% (v/v) ethyl-alcohol, containing 10 mM ammonium acetate. The test tubes were centrifuged for 5 minutes at 10 000 g, the supernatant discarded and the pellet were allowed to dry at room temperature or until all alcohol evaporated. The DNA was dissolved (overnight at 4°C) in 50-100 µl sterile water, depending on the size of the pellet.

2.3.2 Sequencing

Five different regions of genes were sequenced in this study to determine whether genetic variation exists between the different populations of *Clivia caulescens*. The primer region of *atpH-I* and *trnL-F* has been optimized in previous publications (Taberlet *et al*., 1991; Pfosser & Speta, 1999) and was used as initial engagement point, but to increase the yield of the amplification product, modifications were made to the concentrations and PCR cycle temperatures. Standard protocols for the amplification of the *matK, rpoB* and *rpoC1* regions were retrieved (Retrieved at http://www.kew.org/barcoding/iupdate.html on February, 2008). The compositions of the different primers are listed in Table 2.2.

The total of 20 µl amplification reaction consisted of: **(a) *atpH-I*** and ***trnL-F* region:** 5x buffer (10 µl mM dNTP, 500 µl 10x buffer, 0.001 g gelatine, 455 µl dH_2O, 5 µl 100x triton), 0.2 µl Super-Therm Taq Polymerase, 1 µl 25 mM $MgCl_2$, 0.2 µl of each primer (50 µM), 11.4 µl dH_2O, 3 µl template DNA (cons. 20 ng/µl). **(b) *matK, rpoB*** and ***rpoC1* regions:** 2 µl 10x buffer, 0.16 µl mM dNTP's, 0.2 µl Super-Therm Taq Polymerase, 1.2 µl 25 mM $MgCl_2$, 4 µl of each primer (10 µM), 0.8 µl DMSO, dH_2O and 3 µl template DNA (cons. 20 ng/µl).

The following two programs for DNA amplification were used: **(a)** for the ***atpH-I*** and ***trnL-F* region:** 4 min at 94ºC; 35 cycles of (1 min at 94ºC, 1 min at 58-50ºC, 2 min at 72ºC); 5 min at 72ºC and stored at 4ºC. **(b) *matK, rpoB*** and ***rpoC1* regions:** 1 min at 94ºC; 30 s at 94ºC, 40 s at 53ºC, 40 s at 72ºC repeated 35 times; 5 min at 72ºC and stored at 4ºC.

Products were cleaned with the BioFlux Biospin Gel Extraction Kit after amplification and visualized by means of gel electrophoresis if any fragments were obtained. The ABI Prism BigDye Terminator v 3.1 Cycle Sequencing Kit was used for sequencing. The total 10 µl sequencing reaction mixture for all regions, consisted of 0.5 µl Premix, 3.2 pmol Primer, 2 µl Buffer, 0.8 µl DMSO, dH_2O and 3 µl Template DNA (diluted). With the PCR, template DNA was initially denatured (3 minutes at 94°C), followed by 25 cycles of annealing (94°C 10s, 50°C 5 s, 60°C 4 minutes) and elongation phase of 5 minutes at 72°C, with the final phase at 4°C for 45 minutes. All 10 µl PCR reactions were performed on either an Applied Biosystems GeneAmp PCR system 9600 or Applied Biosystem 2720 Thermal Cycler. Excess dye terminators were removed by an ethanol based purification reaction as described in the ABI Prism BigDye Terminator version 3.1 Cycle Sequencing Kit protocol.

Table 2.2: List of primer pairs and their composition used during this study.

Region	Primer	Primer sequence 5' 3'	Reference
matK	2.1 f	CCT ATC CAT CTG GAA ATC TTA G	http://www.kew.orgbarcoding/iupdate.html
	5 r	GTT CTA GCA CAA GAA AGT CG	
atpH-I	atpH-P f	CCA GCA GCA ATA ACG GAA GC	Grivet *et al*., 2001
	atpI-M r	ATA GGT GAA TCC ATG GAG GG	
rpoB	2 f	ATG CAA CGT CAA GCA GTT CC	http://www.kew.orgbarcoding/iupdate.html
	4 r	GAT CCC AGC ATC ACA ATT CC	
rpoC1	2 f	GGC AAA GAG GGA AGA TTT CG	http://www.kew.orgbarcoding/iupdate.html
	4 r	CCATAAGCATATCTTGAGTTGG	
trnL-F	c f	CGA AAT CGG TAG ACG CTA CG	Taberlet *et al*., 1991
	d r	GGG GAT AGA GGG ACT TGA AC	
	e f	GGT TCA AGT CCC TCT ATC CC	
	f r	ATT TGA ACT TAA TTG GAT TGA GC	
	PS1 f	CTA CGG AVT GGT GAC ACG AG	Pfosser & Speta, 1999
	PS2 r	GGG GAT AGA GGG ACT TGA AC	
	PS3 f	GGT TCA AGT CCC TCT ATC CC	
	PS4 r	AGG ATT TTC AGT CCT CTG CTC	

2.3.3 Data analyses

For gel electrophoresis the DNA was loaded onto a 1% (m/v) agarose gel {1% (m/v) agarose; 1x TAE buffer [50x TAE (48.44 g Tris; 11.42 ml acetic acid; 2.92 g EDTA)] with pH8; ethidium bromide (10 mg/ml)]}, run at 100 V for 45 minutes and visualized under UV light. GeneRuler™ DNA Ladder Mix, ready-to-use (Fermentas Life Science) was used as reference markers for the genomic DNA as well as the amplification products on the agarose gels (Lonza). The genomic DNA was photographed with the Gel Doc 100 system using the software program Molecular analyst® Software 1.4.1 (Bio Rad Laboratories).

Various computer programmes were used during this study:

- Sequences were aligned with Geneious Pro 4.5.7 (Rozen & Skaletsky, 2000). All the samples' sequenced results were aligned manually with the help of this program.

- Network 4.6 (Anonymous, 2011) enables the reconstruction of phylogenetic trees and networks (Bandelt *et al*., 1995; 1999). Potential types and ancestral types can be inferred and evolutionary branching can be predicted as well.

- Evolutionary relationships were calculated with MEGA 5 (Kumar *et al*., 2011) and this software provided the means to enable statistical analysis.
    - Formulas used:

Mean Diversity within Subpopulations: In a subpopulation, the mean diversity is defined as $\pi_i = \frac{q}{q-1}\sum_{i=1}^{q}\sum_{j=1}^{q} x_i x_j d_{ij}$, where '$x_i$' is the frequency of *i*-th sequence in the sample from subpopulation *i*, and *q* is the number of different sequences in this subpopulation.

Mean Diversity for Entire Population: For the entire population, the mean diversity is defined as $\pi_T = \frac{q}{q-1}\sum_{i=1}^{q}\sum_{j=1}^{q} x_i x_j d_{ij}$, where '$x_i$' is the estimate of average frequency of the *i*-th allele in the entire population, and *q* is the number of different sequences in the entire sample.

Mean Interpopulational Diversity: The estimate of inter-populational diversity is given by $\delta_{ST} = \pi_T - \pi_S$.

Coefficient of Differentiation: The estimate of the proportion of interpopulational diversity is given by $N_{ST} = \delta_{ST} / \pi_T$.

Net average distance: The net average distance between two groups is given by $dA = dXY - ((dX + dY)/2)$. Where, dXY is the average distance between groups X and Y, and dX and dY are the mean within-group distances.

2.4 RESULTS AND DISCUSSION

2.4.1 Specimen variation

The DNA yield from leaves stored in silica was higher than those stored in CTAB or those of fresh leaves (Appendix A). Of the initial five regions that were sequenced, *trnL-F* amplification failed repeatedly, and subsequently was excluded from all analyses. The other four regions amplified (Appendix B) and all showed variation between the different populations of *C. caulescens* (Appendix C).

The *atpH-I* gene region gave a consensus length of 535 basepairs, with a GC content of 33.0%. A total of 46 basepairs differed and 11 of them were informative and two Indels occurred (Table 2.3). The first Indel consisted of a four base pair insertion (ATGT) present in only *C. mirabilis*. The second one was a six base pair (ATATTT) insertion in both *C. gardenii* and *C. robusta*, indicating a close relationship between these species. The *matK* gene region gave a consensus length of 787 basepairs, with a GC content of 30.7%. A total of 23 basepairs differed and nine of them were informative. An eight base pair insertion (TTGTTTTA) was restricted to *C. miniata*. The *rpoB* gene region gave a consensus length of 520 basepairs, with a GC content of 40.4%. A total of 12 basepairs differed and five of them were informative. The region that gave the least information was that of *rpoC1*, which gave a consensus length of 456 basepairs, with a GC content of 42.5%. Only one basepair differed but was uninformative. The combined dataset gave a consensus length of 2298 basepairs, with a GC content of 35.5%. A total of 82 basepairs differed and 25 were informative (Table 2.3).

The aim of this study was to recognize DNA regions that would provide sufficient differences to identify specimens/localities, but would still contain sufficient similarities to indicate that all the specimens belong to the same species. Therefore these results clearly indicated that the diagnostic value of different DNA regions vary between different species. Although all these regions have been used successfully in the past (Ito *et al*., 1999; Lahaye *et al*., 2008; CBOL plant working group, 2009), they provided varied degrees of success during this study. For example *rpoC1* proved to be of no value in this study while it was successfully used in a

number of other studies (CBOL plant working group, 2009). In a similar way *atpH-I* and *matK* proved to provide the most informative characters for a DNA fingerprinting study (Table 2.3).

Three different indels and four SNPs clearly indicate that *C. caulescens* differ from the other species. Sufficient variation within *C. caulescens* exists to attempt the identification of specific specimens/localities. The value of the SNPs for identifying the geographical origin of a specific specimen will be discussed in the next part of this chapter.

Table 2.3: Summary of the DNA sequence data indicating the aligned (consensus) lengths, conserved and variable sites.

	atpH-1	***matK***	***rpoB***	***rpoC1***	Combined
Length	535	787	520	456	2298
Conserved	483	746	508	455	2192
Variable	46	23	12	1	82
Singleton	35	15	7	1	58
Parsimony informative sites	11	9	5	0	25

2.4.2 Geographical variation

Clivia caulescens is distributed over a range of 400 km from Soutpansberg in the north to Swaziland in the south along the escarpment of Mpumalanga. Specimens were collected at Soutpansberg, Magoebaskloof, Wolkberg, Mariepskop, God's Window (including Wonderview and The Pinnacle), Bearded Man and Swaziland. All localities, except Mariepskop, are along the escarpment. Mariepskop is a single isolated locality and represents almost an island effect. The main aim of this study was to determine whether DNA differences are correlated with these geographical areas.

From the Network Tree (Figure 2.1) it is clear that the Mariepskop specimens group together. It appears as if this group developed from the Northern Group (Soutpansberg, Magoebaskloof and Wolkberg). This is quite interesting since it is geographically much closer to God's

Window. Geographical and climatic changes through time should be studied to understand this phenomenon.

The next group (Southern Group) consist of the God's Window, Bearded Man and Swaziland specimens. This group is geologically isolated from the northern range and this isolation is evident in these results. The outgroup (other species) is geographically isolated from the three *C. caulescens* groups, except on the Bearded Man Mountain where it grows sympatrically with *C. miniata* and the natural hybrid between these species, *C. xnimbicola*. This sympatric area will be discussed in the next chapter.

Three specimens did not fit this geographical separation. The DNA profile of *Spies 8757* (SW532 in the diagram) from Swaziland (Malandwene) corresponds with the northern group rather than to the southern group (God's Window, Bearded Man, and Swaziland). The question therefore arises whether this may be a transplant rather than an original plant from the area.

The other deviations were *Spies 8565* (Bearded Man – BM316) and *Spies 8644* (Mbabane - SW391). Both specimens grouped with *C. miniata*. A possible explanation for this phenomenon may be that introgression between *C. miniata* and *C. caulescens* may have occured, rendering species delimitation impossible. If this is true, then *C. miniata* should have been the original "mother plant" and many backcrosses with *C. caulescens* resulted in specimens resembling *C. caulescens* but still containing *C. miniata* chloroplasts (all DNA regions studied are in the chloroplasts).

Another possibility is that the evolutionary separation between these species was relatively recent and both types of chloroplasts are still present in *C. miniata*. However, further studies including more specimens are needed to determine which hypothesis is the most likely. The size of the nodes in the Network tree is according to how many haplotypes corresponds.

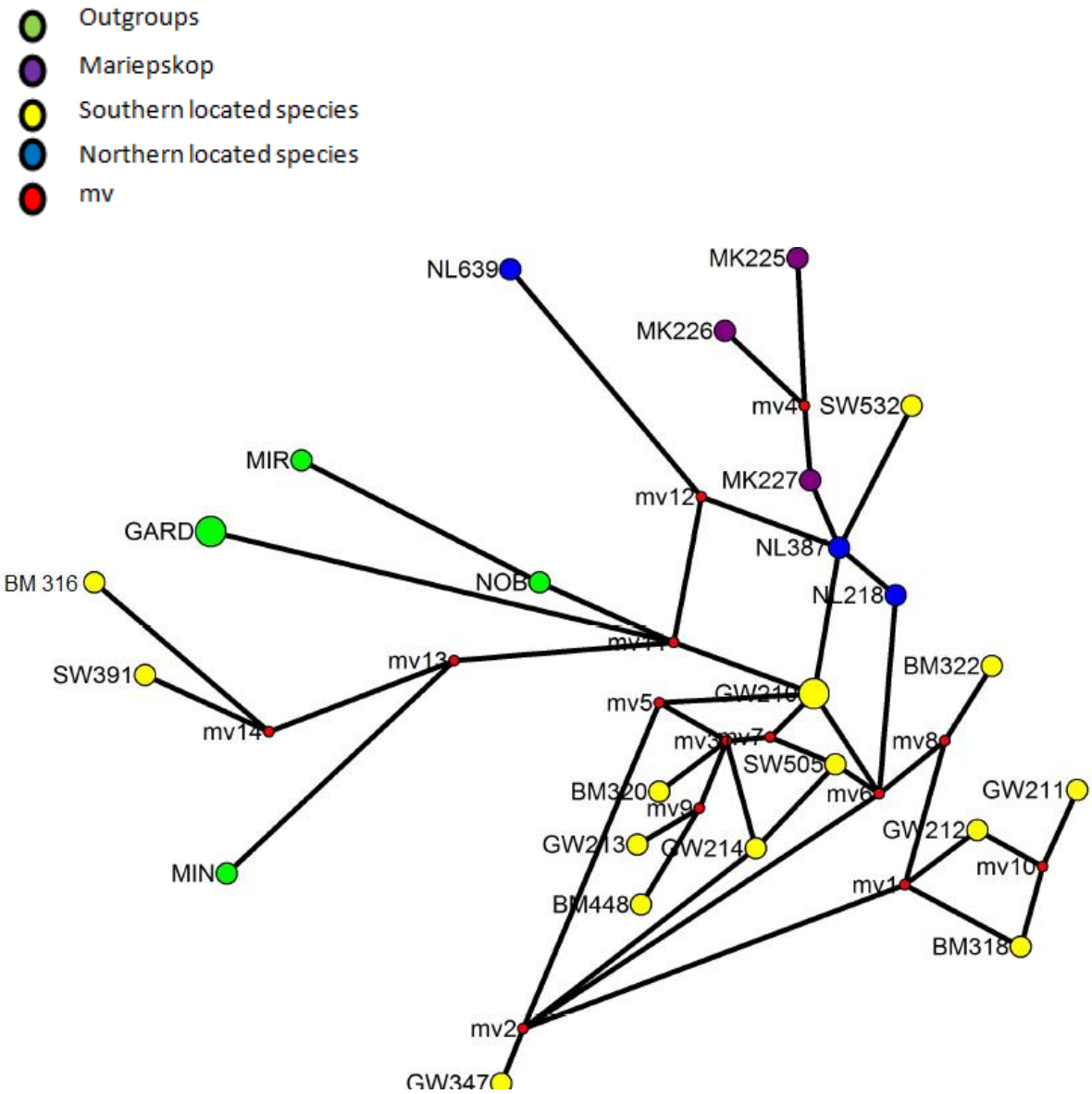


Figure 2.1: Network tree based on SNP's for the different populations of *C. caulescens* included in this study. ('mv' is the abbreviation for 'move')

MK225=Mariepskop, *Spies 8494*; MK226=Mariepskop, *Spies 8495*; MK227=Mariepskop, *Spies 8496*; SW532=Swaziland, *Spies 8784*; SW391=Swaziland, *Spies 8644*; SW505=Swaziland, *Spies 8757*; GW210=God's Window, *Spies 8479*; GW211=God's Window, *Spies 8480*; GW212=God's Window, *Spies 8481*; GW213=God's Window, *Spies 8482*; GW214=God's window, *Spies 8483*; BM316=Bearded Man, *Spies 8565*; BM318=Bearded Man, *Spies 8567*; BM320=Bearded Man, *Spies 8569*; BM322=Bearded Man, *Spies 8571*; BM448=Bearded Man, *Spies 8701*; NL387=Soutpansberg, *Spies 8640*; NL218=Wolkberg, *Spies 8487*; NL639=Magoebaskloof, *Spies 8892*; MIR=*C. mirabilis*, *Spies 8267* ; NOB=*C. nobilis*, *Spies 8254*; GARD=*C. gardenii*, *Spies 8418*; MIN=*C. miniatia*, *Spies 8574*.

Although phylogenetic trees cannot resolve reticulate evolution, two cladograms based on Minimum evolution (Rzhetsky & Nei, 1992) and Maximum parsimony were respectively constructed to determine whether the different geographical areas will predominantly group together.

Bootstrap values (Felsenstein, 1985) from 500 replicates are shown next to the branches (Figure 2.2). As expected the reticulate nature of the species rendered practical results impossible and the geographical distribution is not reflected in the cladogram.

A cladogram was also constructed using the Maximum Parsimony method (Eck & Dayhoff, 1966) (Figure 2.3). Tree #1 out of 236 equally parsimonious trees (length = 62) is shown. The consistency index is (0.67), the retention index is (0.85), and the composite index is 0.75 (0.57) for all sites and parsimony-informative sites (in parentheses). The bootstrap values (Felsenstein, 1985) based on 500 replicates are shown next to the branches.

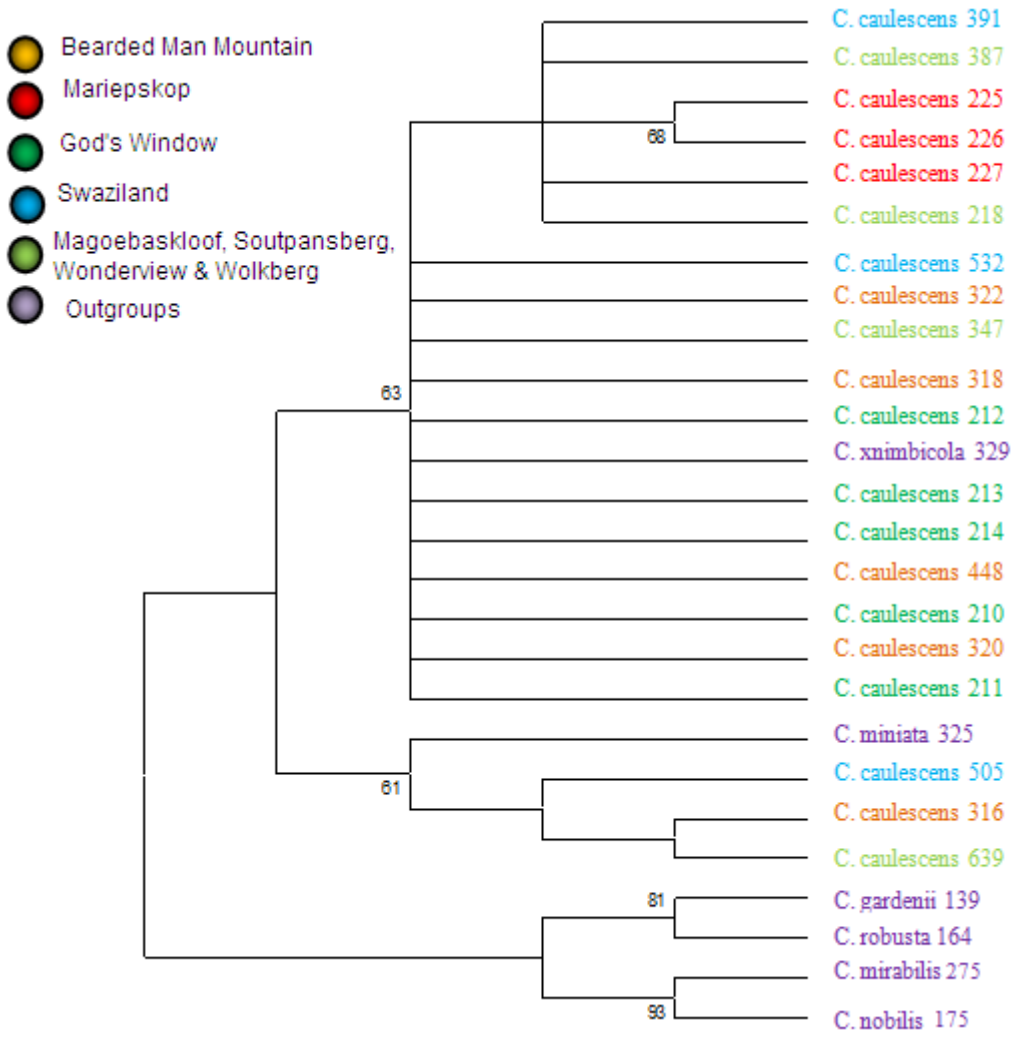


Figure 2.2: Cladogram constructed with the Minimum Evolution method for the combined dataset. The parameters are discussed in the text.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

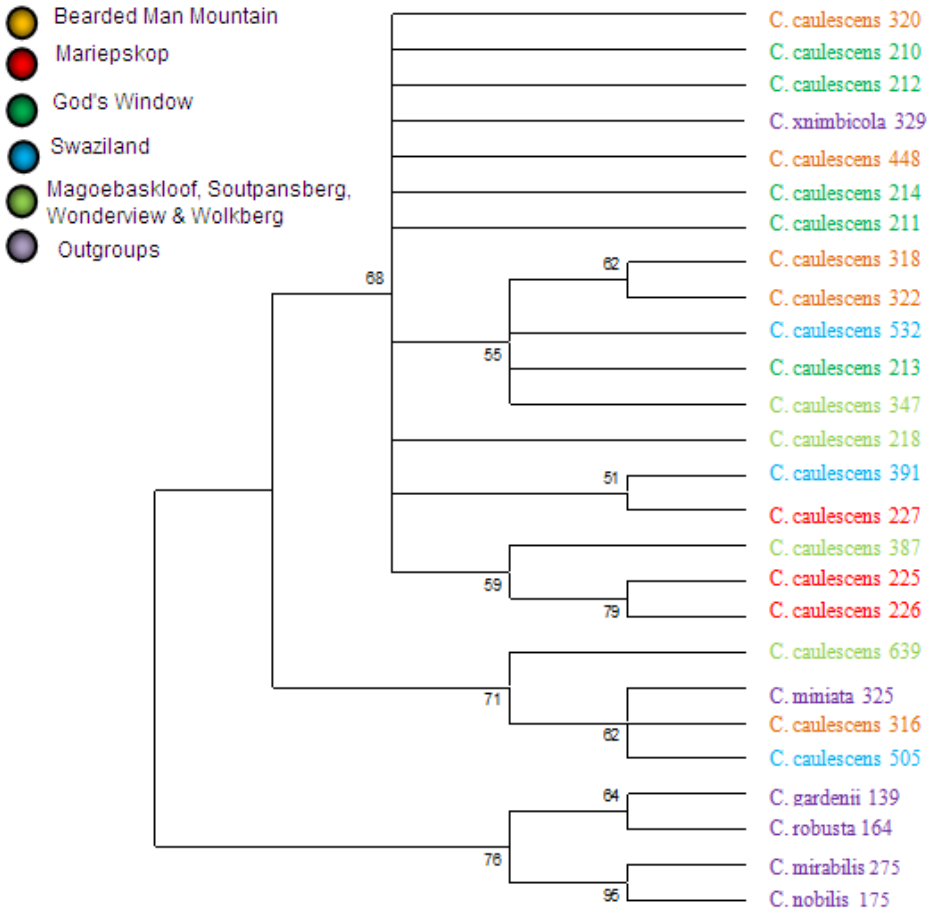


Figure 2.3: Cladogram conducted through Maximum Parsimony method for the combined dataset. The parameters are discussed in the text.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

The MP tree was obtained using the Close-Neighbour-Interchange algorithm (Nei & Kumar, 2000) with search level 3 (Felsenstein, 1985; Nei & Kumar, 2000) in which the initial trees were obtained with the random addition of sequences (10 replicates). All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). In the

final dataset, there were a total of 1642 positions out of which 14 were parsimony informative.

Again the cladogram did not give a true reflection of the geographical distribution of the specimens included in this study. Cladograms obtained for the other gene regions are in Appendix D.

When variation among the different populations are compared for *atpH-I*, Mariepskop differs the most from the other populations (Table 2.4), especially from Wonderview. Magoebaskloof and Wolkberg differs the least from Soutpansberg, thus confirming the Network results. Swaziland and Bearded Man Mountain have a mean distance of 0.007. The largest differences in mean distances are between Wonderview and Mariepskop at 0.015. Magoebaskoof and Wonderview correspond to some extent, but there is still a mean distance of 0.006 between those two populations.

When the populations are grouped together as Northern location (Soutpansberg, Magoebaskloof, Wolkberg); Southern location (God's Window, Wonderview, Bearded Man and Swaziland) and Mariepskop, this diagram below illustrates the mean distances for region *atpH-I* between these populations (Figure 2.4). This region was chosen to be displayed, because it rendered the most differences between the different locations.

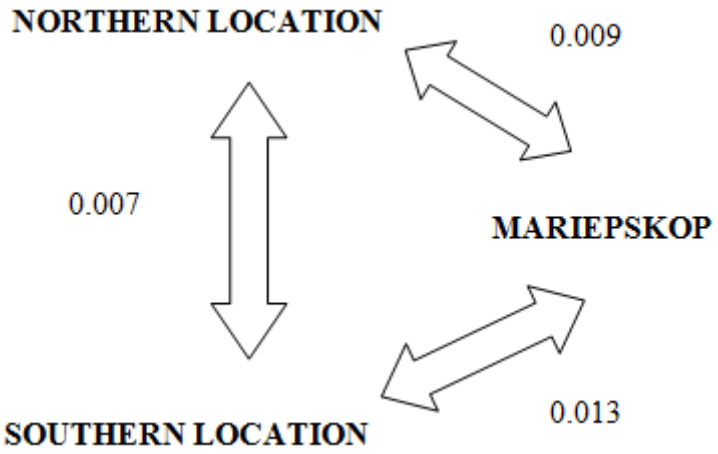


Figure 2.4: Flow diagram indicating gene flow between locations of *C. caulescens* based on region *atpH-I*.

Mariepskop's DNA is the most different from the south located populations of *C. caulescens*, which is interesting because as mentioned before, Mariepskop is closer situated to the south located populations than towards the northern located populations. This evidence support the

hypothesis that Mariepskop seems to have developed from northern located populations as detected from the Network data obtained as well.

The highest mean distances (genetic variation) obtained from the dataset within the different *Clivia caulescens* populations (Table 2.5) is the Mariepskop population with 0.006, 0.015 and 0.008 respectively from the regions used for the combined dataset, *atpH-I* and *rpoB*. The second highest mean distances within a population were obtained for Swaziland by the *atpH-I* and *rpoB* regions respectively 0.007 and 0.010. The God's Window population displayed very little mean distances within the population. Over all *atpH-I* region displayed the highest mean distances within the different populations of *Clivia caulescens*.

The other tables for the rest of the genetic regions (*matK*; *rpoB; rpoC* and combined dataset) are represented in Appendix E and F.

Table 2.4: Mean genetic distances between groups based on *atpH-I* results, indicating DNA differences between the different populations.

(1 – Bearded Man Mountain; 2 – Swaziland; 3 – Soutpansberg; 4 – God's Window; 5 – Mariepskop; 6 – Magoebaskloof; 7 – Wonderview; 8 – Wolkberg)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Bearded Man Mountain								
Swaziland	0.007							
Soutpansberg	0.007	0.005						
God's Window	0.006	0.005	0.004					
Mariepskop	0.014	0.013	0.008	0.012				
Magoebaskloof	0.009	0.007	0.002	0.007	0.010			
Wonderview	0.007	0.006	0.007	0.006	0.015	0.010		
Wolkberg	0.007	0.004	0.002	0.006	0.010	0.005	0.005	
Outgroups	0.004	0.004	0.002	0.002	0.010	0.005	0.005	0.005

Table 2.5: Mean genetic distances within groups of the different populations of *Clivia caulescens* included in this study.

	atpH-I	***matK***	***rpoB***	**Combined dataset**
Bearded Man Mountain	0.009	0.005	0.002	0.004
Swaziland	0.007	0.005	0.010	0.006
God's Window	0.003	0	0	0.002
Mariepskop	0.015	0.001	0.008	0.006
Outgroups	0	0.007	0.003	0.003

The topology of the cladograms based on Minimum Evolution and Maximum Parsimony (MP) corresponds except for the two separate clusters in MP against the one observed for the Minimum Evolution cladogram (Bearded Man Mountain (*Spies 8569*, *Spies 8701*); God's Window (*Spies 8480*, *Spies 8481*, *Spies 8483*); & *C. xnimbicola* cluster 1; Bearded Man Mountain (*Spies 8567*, *Spies 8571*), Swaziland (*Spies 8784*), God's Window (*Spies 8482*) & Wonderview (*Spies 8596*) cluster 2 Maximum Parsimony; and all these populations clustered together Minimum Evolution).

Regarding the statistical analysis, the highest mean distance between the different populations were displayed by the *atpH-I* region and then *matK* region, followed by the combined dataset. There is not much mean distances displayed by *rpoB* region. From all the flow diagrams (Appendix F) can be depicted that there is definitely DNA differences between the different populations. With almost certainty Mariepskop is a population on its own as this population differs the most to all the other populations. Although one would've hypothesized that Bearded Man Mountain and Swaziland are possibly the same population, as they are so closely situated, there is consistently a distinct difference of 0.005 (from the combined dataset, *matK* and *rpoB*) and even 0.007 at *atpH-1* region.

Wonderview and God's Window, which are very closely situated, corresponds at *matK* and *rpoB* regions and mostly with the combined dataset, therefore posing to be more likely to be the same population. The populations situated most northerly e.g. Soutpansberg, Magoebaskloof and Wolkberg, seem to be separate populations because of the relatively high mean distances between them except with the *rpoB* region. All these populations were not well enough represented to say for certain though. Mean distance diagrams indicating DNA differences for the populations divided in the North, South location and Mariepskop are presented in Appendix G.

Regarding the variation within the different populations of *Clivia caulescens* that have been examined, it was detected that Mariepskop population displayed the most variation within the population and God's Window population displayed the least variation within the population. Fortunately the *C. caulescens* population of Mariepskop seems to exhibit genetic variation, whereas the God's Window population data acquired indicate possible genetic erosion in the population. A greater emphasis should be put on the conservation of the God's Window *C. caulescens* population. The gene region of choice would be *atpH-I* to investigate if there is intraspecific variation in *C. caulescens* populations as this region demonstrated the highest mean distances within the different populations. Unfortunately we could not calculate the

mean distances within Magoebaskloof, Soutpansberg, Wonderview and Wolkberg populations, as we only had one specimen of each.

When the results of the phylogenetic and statistical analysis were combined, it was detected that most Bearded Man Mountain specimens and God's Window specimens clustered together in the cladograms and in the mean distances tables. In the flow diagrams it was detected that the mean distance between these two populations was only 0.002 for the combined dataset and *matK*, 0.001 from *rpoB* and 0.006 from *atpH-I*. Mariepskop seemed to be totally separate from the other populations, but when it was closely examined it was detected that Mariepskop clusters with Wolkberg population everytime and that the mean distance between them are 0.004 for combined dataset and *rpoB*; 0.003 for *matK* and 0.010 for *atpH-I*. The deduction can be made again that Mariepskop is more closely related to the northern location of *C. caulescens* although it is more closely situated towards the southern location of *C. caulescens*.

CONCLUSION

One of the key issues in the study of heredity and variation at the molecular level is the detection of associations between DNA sequence variation and the heritable phenotypes (Gupta *et al*., 2001). In this study we made use of SNPs through direct sequencing of the five genetic fragments or regions of choice. Of the five regions, four including the combined data set showed interspecific variation between the different populations of *Clivia caulescens*, as well as intraspecific variation. The *atpH-I* and *matK* regions showed the most variation and indicated that the Mariepskop specimens differ from the other populations. The DNA data indicated the existence of three distinct groups, the northern (Soutpansberg, Magoebaskloof, Wolkberg), Mariepskop and southern (Gods Window, Bearded Man, Swaziland) populations. Evidence that Mariepskop is a single locality was supported by the data acquired. There was also some evidence that seems like Mariepskop developed from the northern locations, although it was more closely situated to the southern locations.

2.7 REFERENCES

ANONYMOUS. 2009. Royal Botanic Gardens: DNA Barcoding. Retrieved November 20, 2009, from http://www.kew.org/barcoding/iupdate.html.

ANONYMOUS. 2011. Network v. 4.5.1.0. Retrieved April 20, 2008, from http://www.fluxus-engineering.com.

AVISE, J. C. 1991. Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual Review of Genetics* **25**: 45-69.

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

BANDELT, H-J., FORSTER, P., SYKES, B. C. & RICHARDS, M. B. 1995. Mitochondrial portraits of human populations. *Genetics* **141**: 743-753.

BANDELT , H-J., FORSTER, P. & RÖHL, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37-48.

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

CLEGG, M. T. & ZURAWSKI, G. 1990. Chloroplast DNA and the study of plant phylogeny: present status and future prospects. *In* P. S. Soltis, D. E. Soltis & J. J. Doyle [eds.], *Molecular systematic of plants*, Chapman & Hall, New York.

DEMESURE, B., SODZI, N. & PETIT, R. J. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**: 129-131.

DUMOLIN-LAPENGUE, S., PEMONGE, M. H. & PETIT, R. J. 1997. An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology* **6**: 393-397.

ECK, R. V. & DAYHOFF, M. O. 1966. *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Silver Springs, Maryland.

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

GRIVET, D., HEINZE, B., VENDRAMIN G. G. & PETIT R. J. 2001. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* **1**: 345-349.

GUPTA, K. P., ROY, J. K. & PRASAD, M. 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science* **80**: 4.

HAMBY, R. K. & ZIMMER, E. A. 1992. Ribosomal RNA as phylogentic tool in plant systematic. *In* P. S. Soltis, D. E. Soltis, and J. J. Doyle [eds.], *Molecular systematic of plants*, 50-91. Chapman and Hall, New York.

HAMILTON, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**: 521-523.

ITO, M., KAWAMOTO, A., YOKO, K., TOMOHISA, Y. & KURITA, S. 1999. Phylogenetic relationships of Amaryllidaceae based on *matK* sequence data. *Journal of Plant Research* **112**: 207-216.

JEHAN, T. & LAKHANPAUL, S. 2006. Single nucleotide polymorphism (SNP) – Methods and applications in plant genetics: A review. *Indian Journal of Botany* **5**: 435-459.

KNAAK, C., HAMBY, R. K., ARNOLD, M. L. LEBLANC, M. D., CHAPMAN, R. L. & ZIMMER, E. A. 1990. Ribosomal DNA variation and its use in plant biosystematics. *In* S. Kawano [ed.], *Biological approaches and evolutionary trends in plants*, 135-158. Academic Press, London.

KUMAR, S., NEI, M., DUDLEY, J. & TAMURA, K. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**: 299-306.

LAHAYE, R., VAN DER BANK, M., BOGARIN, D., WARNER, J., PUPLIN, F., GIGOT, G., MAURIN, O., DUTHOIT, S., BARRACLOUGH, T. G. & SAVOLAINEN, V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Science USA* **105**: 2923-2928.

MCGREGOR, C. E., LAMBERT, C. A., GREYLING, M. M., LOUW, J. H. & WARNICH L. 2000. A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L) germplasm. *Euphytica* **113**: 135-144.

MORITZ, C., DOWLING, T. E. & BROWN, W. M. 1987. Evolution of animal mitochondrial DNA; relevance for population biology and systematic. *Annual Review of Ecology and Systematics* **18**: 269-292.

NEI, M. & KUMAR, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

NIU, T. 2004. Algorithms for inferring haplotypes. *Genetic Epidemiology* **27**: 334-347.

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

PALMER, J. D. 1987. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *American Naturalist* **130**: S6-S29.

PALMER, J. D. 1992. Comparison of chloroplasts and mitochondrial genome evolution in plants. *In* R. G. Hermann [ed.], *Cell organelles*, 99-133. Springer-Verlag, Berlin.

PALMER, J. D., JANSEN, R. K., MICHAELS, H. J., CHASE, M. W. & MANHART, J. R. 1988. Chloroplast DNA variation and plant phylogeny. *Annals of Missouri Botanical Garden* **75**: 1180-1206.

PFOSSER , M. & SPETA, F. 1999. Phylogenetics of Hyacinthaceae based on plastid DNA sequences. *Annals of the Missouri Botanical Garden* **86**: 852-875.

RAFALSKI, J. A. 2001. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Science* **162**: 329-33.

ROGSTAD, S. H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon* **41**: 701-708.

ROZEN, S. & SKALETSKY, H. J. 2000. Primer3 on the www for general users and for biologist programmers. *In* Krawetz, S. and Misener, S. Humana [eds.], *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, 365-386, Humana Press, Totowa, New Jersey.

RZHETSKY, A. & NEI, M. 1992. A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution* **9**: 945-967.

TABERLET, P., GIELLY, L., PAUTOU, G. & BOUVET, J. 1991. Universal primers for the amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1110.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (submitted).

Chapter 3:

POPULATION DYNAMICS IN THE BEARDED MAN *CLIVIA* POPULATION

Suzanne Stegmann

3.1 ABSTRACT

Clivia xnimbicola is a natural hybrid between *C. miniata* and *C. caulescens*, growing along the slopes of Bearded Man Mountain on the border between Swaziland and Mpumalanga. Almost no scientific studies on the origin and directions of this *nothospecies have been done. In an attempt to determine whether the formation of this hybrid species is a repetitive occurrence or a once off event, five different gene regions were sequenced, i.e. the *atpH-I*, *matK*, *rpoB*, *rpoC1* and *trnL-F* regions. The *trnL-F* region did not amplify and *rpoC1* did not show any variation among the three species.

SNPs indicate that *C. caulescens* usually acts as the *pod-parent in crosses but a single exception was observed. The presence of unique nucleotides in *C. xnimbicola* indicates that the original parents of this hybrid species were possibly not included in this study. The insufficient number of specimens due to the lack of obtaining a permit to collect more material inhibited this study and consequently a few hypotheses could be suggested without firm proof to support any of them.

*nothospecies - A hybrid which is formed by direct hybridization of two species, not other hybrids.

*pod-parent - The parent plant that contributes the female reproductive cells. In hybridizing, the parent that forms the pods where the fertilized seeds develop and ripen.

3.2 INTRODUCTION

On the border between Mpumalanga and Swaziland, Bearded Man Mountain reaches an altitude of 1337m. This is the only area where *Clivia caulescens* and *C. miniata* grow sympatrically. Although the presence of a natural occurring interspecific hybrid on Bearded Man was known from as early as 1969 (Rourke, 2004), this hybrid was first described as a nothotaxon (nothotaxa are assemblages of hybrid plants corresponding to a particular hybrid formula), *Clivia xnimbicola*, in 2006 (Swanevelder *et al*., 2006; Truter *et al*., 2007).

Although the three species grow together on the slopes of Bearded Man, *C. miniata* grows at the lowest altitude and *C. caulescens* at the highest (Figure 3.1). The margins of distribution are very close and the two species grow within 10m from one another [locality A – (Le Roux, 2009)]. Among the rocks separating the two species, *C. xnimbicola* grows.

The presence of a natural occurring hybrid raises many questions: how did the two parental species cross (*C. miniata* flowers approximately one month before *C. caulescens*); was the

hybridization process a single event or are new hybrids formed constantly; is the hybridization event occurring only in one way or does reciprocal crosses occur; does

Figure 3.1. Photographs of *Clivias* at the Bearded Man Mountain, Mpumalanga, South Africa. **A**. Suzanne Stegmann with the Bearded Man Mountain in the background. **B**. Flower of *Clivia miniata*. **C**. *Clivia* population at Bearded Man Mountain. **D**. Flower of *Clivia xnimbicola*. **E.** *Clivia miniata* at the end of flowering season. These photographs were taken during a visit to this population by the author.

introgression occur; which pollinators played a role in this phenomenon? Currently these aspects have been touched on but no scientific evidence in support of this exists or has been conducted. The main objective of this study is to use DNA sequence data (sequences of the *atpH-I*, *matK*, *rpoB* and *rpoC1* regions and a combination of theses gene regions) to propose and test certain hypothesis on the questions raised above.

3.3 MATERIALS AND METHODS

3.3.1 Specimen locality and extraction

The sources of specimens used in this analysis are listed in Table 3.1. DNA extractions were carried out as described by Rogstad (1992), with a few modifications.

Table 3.1: List of the plant specimens included in this study, indicating their geographical origin and voucher numbers. All voucher material is housed in the Geo Potts Herbarium (BLFU).

Species	**Locality**	**Voucher number**
C. caulescens	Bearded Man Mountain	*Spies 8565, 8567, 8569, 8571, 8701*
C. miniata		*Spies 8674, 8656, 8558, 8572, 8721*
C. xnimbicola		*Spies 8578, 8641, 8654, 8655, 8648*
C.gardenii	Greytown	*Spies 8418*
C. miniata	Dwesa	*Spies 8574*
C. mirabilis	Donkerhoek	*Spies 8267*
C. nobilis	Keiskamma	*Spies 8254*
C. robusta	Port Shepstone	*Spies 8440*

3.3.2 Sequencing

The same methods, primers (genetic fragments or regions) and data analysis tools as described in Chapter 2**: GENETIC VARIATION IN *CLIVIA CAULESCENS*** were used during the analysis of genetic variation of the Bearded Man Mountain Clivias.

3.4 RESULTS AND DISCUSSION

The natural distribution range of the *C. xnimbicola* is confined to the Baberton Centre of Endemism (Van Wyk & Smith, 2001), the only known region in which the distribution ranges of *C. caulescens* and *C. miniata* overlap (Swanevelder, 2003). The main aim of this study was to determine whether DNA differences are correlated with the hypothesis set in the introduction.

The DNA yields from leaves stored in silica were higher than those stored in CTAB or those of fresh leaves (Appendix H). Of the initial five regions that were sequenced, *trnL-F* amplification failed repeatedly, and these sequences were therefore excluded from all analyses. The other four regions amplified (Appendix I) had some degree of variation between the different specimens of the Bearded Man species (Appendix J).

The *atpH-I* gene region gave a consensus length of 525 basepairs, with a GC content of 33.1%. A total of 44 basepairs differed and 10 of them were informative. The *matK* gene region gave a consensus length of 788 basepairs, with a GC content of 30.9%. A total of 20 basepairs differed and seven of them were informative. The *rpoB* gene region gave a consensus length of 520 basepairs, with a GC content of 40.1%. A total of 19 basepairs differed and seven of them were informative. The region that gave the least information was that of *rpoC1*, which gave a consensus length of 456 basepairs, with a GC content of 42.5%. Only three basepair differed but all were uninformative. The combined dataset gave a consensus length of 2281 basepairs, with a GC content of 35.8%. A total of 84 basepairs differed and 23 were informative (Table 3.2). There were no Indels detected for either of the gene regions used individually or as combined dataset.

The majority of parsimony informative sites are shared between *C. caulescens* and *C. xnimbicola* (Appendix K), suggesting *C. caulescens* to be the pod-parent (chloroplast contributor) in each case. However, not all SNPs corresponded. In some *C. xnimbicola* specimens unique SNPs were observed both in the *matK* and *rpoB* sequences. The aim of this study was to recognize gene regions that would provide sufficient differences to give an indication of how the two putative parental plants, *C. caulescens* and *C. miniata* hybridized to form the natural hybrid, *C. xnimbicola* and if this hybridization process is only occurring in one way or whether reciprocal crosses do occur. These unique SNPs indicate that either the specific parent responsible for the hybridization process was not included in the study, or that new mutations occurred after the speciation event. The variation among *C.*

xnimbicola specimens further suggests that this hybrid species were produced several times and does not represent a single hybridization event. The possibility of several different (genetically different) specimens forming the hybrid species seems consequently the most possible explanation for the differences in DNA sequences from different specimens. This also makes the probability of the "'correct parents' not included in the study hypothesis", more plausible.

Table 3.2: Summary of the DNA sequence data indicating the aligned (consensus) lengths, conserved and variable sites.

	atpH-1	***matK***	***rpoB***	***rpoC1***	Combined
Length	525	788	520	456	2281
Conserved	479	748	501	453	2183
Variable	44	20	19	3	84
Singleton	34	13	17	3	61
Parsimony informative	10	7	7	0	23

From the analysis of the *atpH-I* region's results, it appears as if SNPs are shared by all three the *Clivia* species on the Bearded Man, signifying major hybridization events. Introgression may be present due to the findings of SNPs shared by all three the species at the Bearded Man. The reproduction of natural hybrids in many cases follows a definite pathway known as introgression. Introgressive hybridization, as defined by Anderson (1949), is the repeat backcrossing of a natural hybrid to one or both parental populations. Field observations done by Swanevelder *et al*., (2006), suggested some introgression between *C. xnimbicola* and its putative parents. Where the putative hybrid grew more closely to one of the parents it showed more morphological features with that parent.

Although the analyses of the *rpoB* and *matK* regions indicate that *C. miniata* is mostly a separate entity with its own unique sequences, one *C. caulescens* specimen deviates from this observation. *Spies 8565* is a *C. caulescens* specimen sharing many nucleotides with *C. miniata*. The only explanation for this phenomenon is introgression in the opposite direction

as discussed earlier. The most possible hypothesis seems that it originated as a cross where *C. miniata* acted as the pod-parent and through repeated back-crosses (i.e. introgression) with *C. caulescens* a specimen was produced mimicking the *C. caulescens* phenotype, while still containing the *C. miniata* chloroplasts. This indicates that hybridization may well occur in both directions. Unfortunately all applications to Nature Conservation to obtain permits to collect sufficient material for a proper population study on Bearded Man failed. No application was turned down but no answer, either way, was ever received. The only way to get a more precise answer about hybridization in these species will be to study numerous samples from that area to determine the degree of introgression present.

From the Network tree (Figure 3.2) it is clear that the putative natural hybrid between *C. caulescens* and *C. miniata* clusters with *C. caulescens*. This is further evidence that *C. caulescens* could be the mother plant of *C. xnimbicola*. The *C. miniata* specimens cluster together with a single *C. caulescens* plant (C316 – *Spies 8565*). The outgroup specimens (other species) are geographically isolated from the other groups.

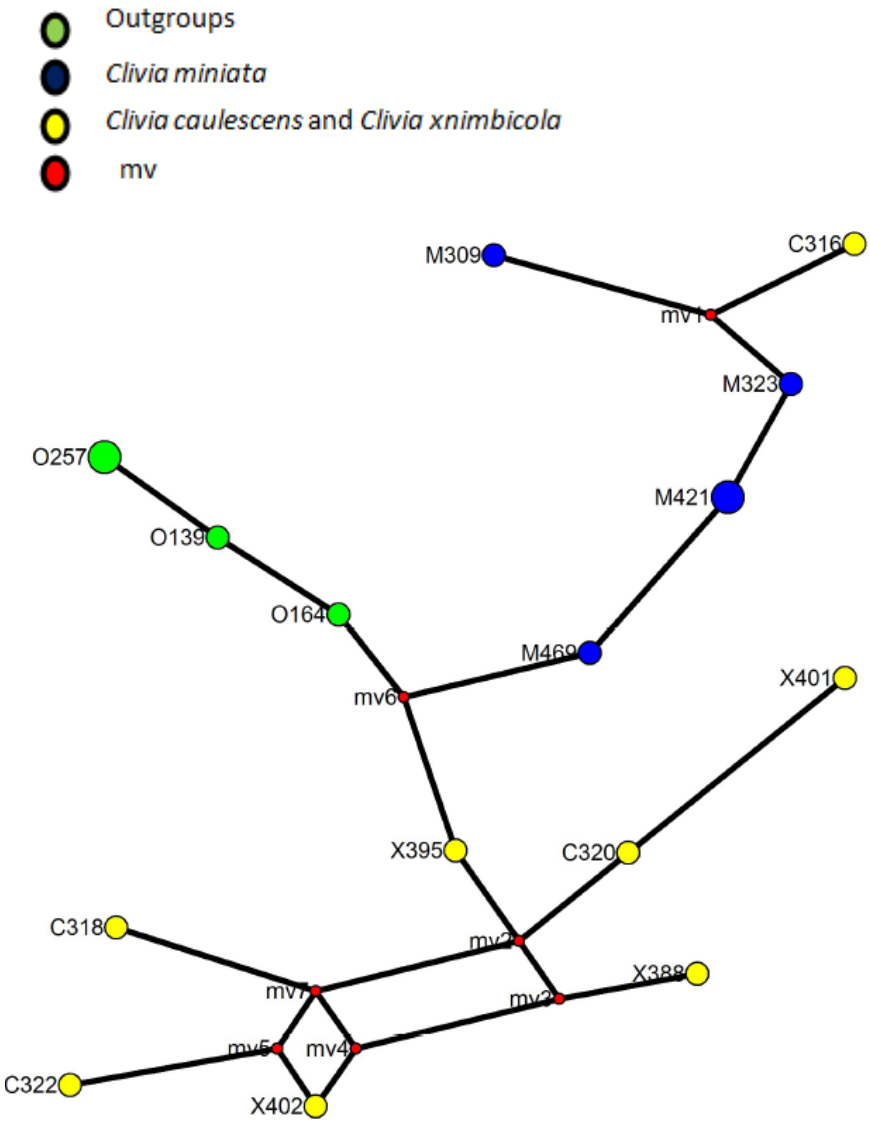


Figure 3.2: Network tree for the different populations of *C. caulescens* included in this study.

C316 =*C. caulescens*, *Spies 8565*; M309=*C. miniata*, *Spies 8558*; M323=*C. miniata*, *Spies 8572*; M421=*C. miniata*, *Spies 8674*; M469=*C. miniata*, *Spies 8721*; X401=*C. xnimbicola*, *Spies 8654*; X395=*C. xnimbicola*, *Spies 8648*; C320=*C. caulescens*, *Spies 8569*; X388=*C. xnimbicola*, *Spies 8641*; X402=*C. xnimbicola*, *Spies 8655*; C318=*C. caulescens*, *Spies 8567*; C322=*C. caulescens*, *Spies 8571*; O257=*C. nobilis*, *Spies 8254*; O139=*C.gardenii*, *Spies 8418*; O164=*C. robusta*, *Spies 8440*.

Two cladograms based on Minimum evolution (Rzhetsky & Nei, 1992) and Maximum parsimony was respectively constructed to determine whether the different species will predominantly group together. The evolutionary history was inferred using the Maximum

Parsimony method (Eck & Dayhoff, 1966). Tree #1 out of 111 most parsimonious trees (length = 65) is shown in Figure 3.3.

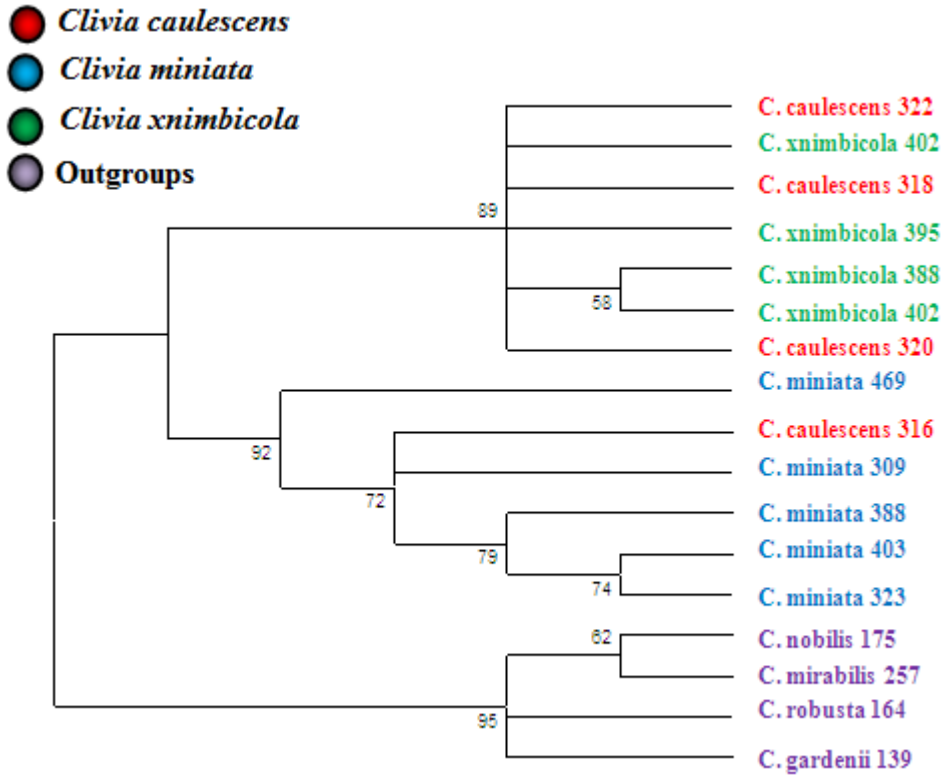


Figure 3.3: Cladogram conducted through Maximum Parsimony method for the combined dataset.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

The consistency index is (0.68), the retention index is (0.85), and the composite index is 0.74 for all sites and parsimony-informative sites (in parentheses). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches (Felsenstein, 1985). The MP tree was obtained using the Close-Neighbour-Interchange algorithm (Nei & Kumar, 2000) with search level 3 (Felsenstein, 1985; Nei & Kumar, 2000) in which the initial trees were obtained with the random addition of sequences (10 replicates). All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There were a total of 1720 positions in the final dataset, of which 17 were parsimony informative. Phylogenetic analyses were conducted in MEGA4 and MEGA5 (Tamura *et al*., 2007; 2011).

The evolutionary history was also inferred using the Minimum Evolution method (Rzhetsky & Nei, 1992). The optimal tree with the sum of branch length = 0.04 is shown in Figure 3.4. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura *et al*., 2004) and are in the units of the number of base substitutions per site. The ME tree was searched using the Close-Neighbour-Interchange (CNI) algorithm (Nei & Kumar, 2000) at a search level of 1. The Neighbour-joining algorithm (Saitou & Nei, 1987) was used to generate the initial tree. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 1720 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 and MEGA5 (Tamura *et al*., 2007; 2011). Cladograms obtained through the use of the other gene regions (*atpH-I*; *matK*; *rpoB* and *rpoC1*) are given in Appendix L.

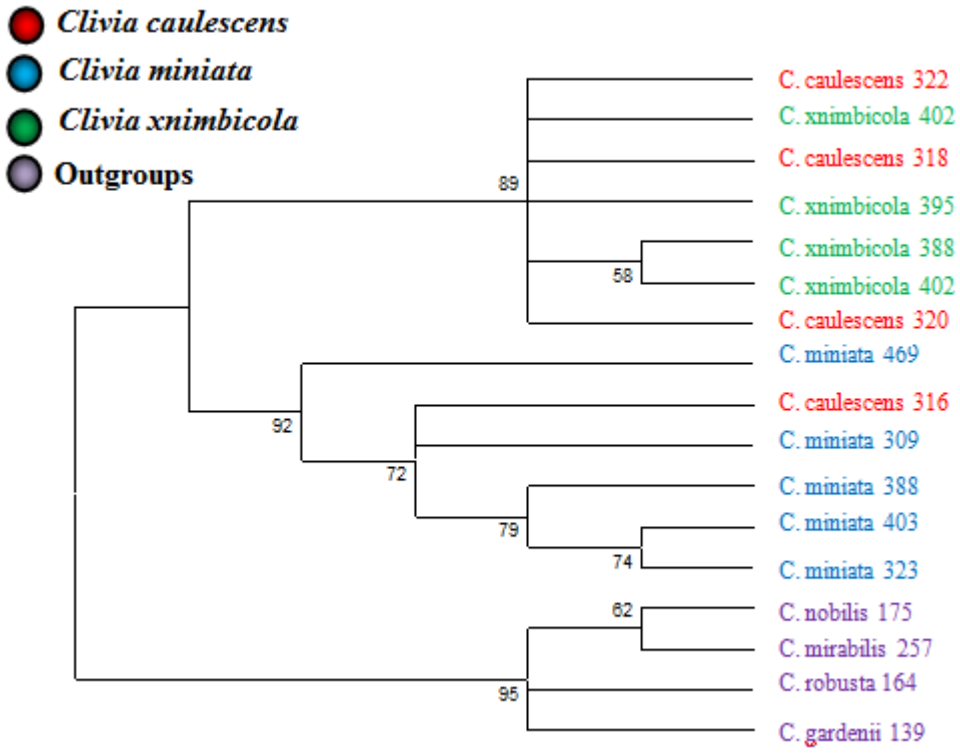


Figure 3.4: Cladogram conducted through Minimum Evolution method for the combined dataset.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

It was apparent that *C. xnimbicola* mostly clustered with *C. caulescens* in both cladograms. This supports the previous conclusion that *C. caulescens* usually acts as the pod-plant of the natural hybrid. The exception, *Spies 8565*, is again clearly separated from the 'norm' in both cladograms.

The results obtained from the *atpH-I* region revealed the largest mean distances between the different populations found at the Bearded Man Mountain. All the other gene regions' results are in Appendix M. Here the most interspecific variation occurs between *C. caulescens* and *C. miniata*, at 0.011 (Table 3.3). The mean distance (0.009) was similar when *C. xnimbicola* were compared to both the putative parents.

Table 3.3: Mean genetic distance between groups for *atpH-I* region.

	1	**2**	**3**
C. caulescens			
C. miniata	0.011		
C. xnimbicola	0.009	0.009	
Outgroups	0.008	0.006	0.005

Consideration was also paid to attempt to unravel if there is any intraspecific variation present for the Clivias situated at the Bearded Man population. The highest mean diversity within a population was detected in the *C. caulescens* population, that revealed 0.012 at *atpH-I* region and 0.007 that was shared by the combined dataset as well as the *matK* and *rpoB* regions. The least intraspecific variation was obtained in *C. miniata*, with a mean diversity of nil at the regions of *rpoB* and *rpoC1*. The natural hybrid, *Clivia xnimbicola* revealed a mean diversity within the population with all the regions except with the *matK* region. All results are based on the pairwise analysis of 17 sequences. Analyses were conducted using the Maximum Composite Likelihood method in MEGA5 (Tamura *et al*., 2004; 2007; 2011). All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion Option). Unfortunately an insufficient number of specimens make these deductions hypothetical.

In order to illustrate possible gene flow between the species at the Bearded Man by using the *atpH-I* region we constructed a flow diagram (Figure 3.5), by using information obtained from MEGA results. The highest mean genetic distance was between *C. caulescens* and *C. miniata*. Both of the putative parents had a mean distance of 0.009 to the natural hybrid. This can be an indication that both parents DNA is equally present in the hybrid.

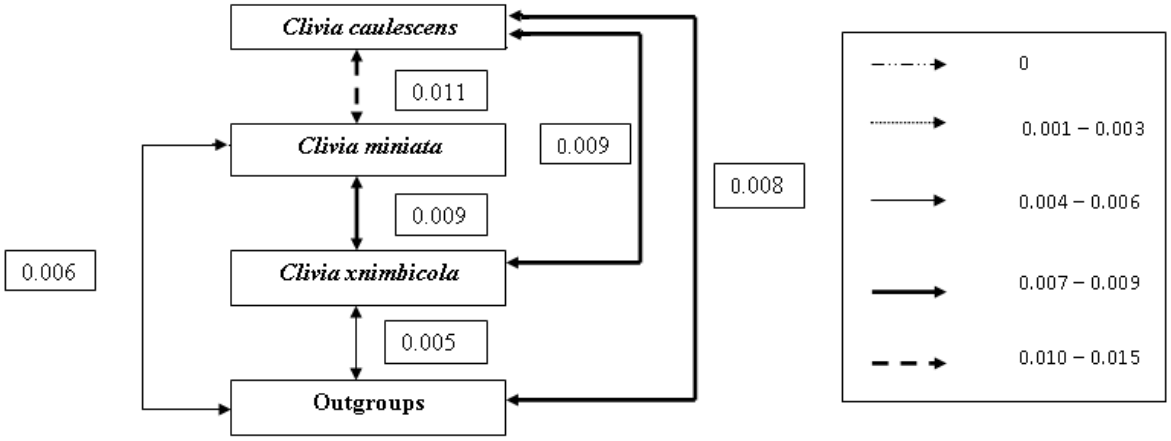


Figure 3.5: Flow diagram of region *atpH-I* indicating gene flow between specimens at the Bearded Man

See Appendix N for diagrams that illustrate the mean distances between the Bearded Man Mountain species of the other gene regions.

The region of *atpH-I* was instrumental in detecting any mean distances (genetic variation) within the species found at the Bearded Man Mountain (Table 3.4), although it is still recommended to rather make use of a combination of gene regions to study the diversity within species to get a true reflection of the genetic diversity present.

Table 3.4: Mean diversity within groups from Bearded Man Mountain Clivias.

	atpH-I	***matK***	***rpoB***	***rpoC1***	**Combined dataset**
C. caulescens	0.012	0.007	0.007	0.000	0.007
C. miniata	0.010	0.003	0.000	0.000	0.004
C. xnimbicola	0.007	0.000	0.012	0.004	0.005
Outgroups	0.002	0.005	0.002	0.002	0.0

The results obtained from the Network tree and the other two cladograms, Minimum Evolution and Maximum Parsimony data corresponded in the sense that the same specimens clustered together, and therefore the data is more reliable as it is supported by different methods of analyses.

3.5 CONCLUSION

Clivia xnimbicola is the natural hybrid between *C. miniata* and *C. caulescens*. All three species grow sympatrically along the slopes of Bearded Man Mountain, on the border between Swaziland and Mpumalanga. Various hypotheses were tested: the hybridization event occurred only once/was repetitive; hybridization occurs in only one direction (introgression). In an attempt to test these hypotheses, five different gene regions were sequenced, i.e. the *atpH-I, matK, rpoB, rpoC1* and *trnL-F* regions. The *trnL-F* region did not amplify and *rpoC1* did not show any variation among the three species.

SNPs indicate that *C. caulescens* usually acts as the pod-parent in crosses but a single exception was observed. The presence of unique nucleotides in *C. xnimbicola* indicates that the original parents of this hybrid species were possibly not included in this study. The insufficient number of specimens due to the lack of obtaining a permit to collect more material inhibited this study and consequently a few hypotheses could be suggested without firm proof to support any of them.

3.6 REFERENCES

ANDERSON, E. 1949. Introgressive Hybridization. John Wiley, New York.

ECK, R. V. & DAYHOFF, M. O. 1966. *Atlas of protein sequence and structure*. National Biomedical Research Foundation, Silver Springs, Maryland.

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

LE ROUX, A. 2009. A Layman's observations of Habitat of Clivia x nimbicola. *Clivia Yearbook* **11**: 50-53. *Clivia* club, Kirstenbosch National Botanical Garden, Cape Town.

NEI, M. & KUMAR, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

ROGSTAD, S. H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon* **41**: 701-708.

ROURKE, J. 2004. Natural interspecific hybrids in Clivia. *Clivia Yearbook 5*, Clivia Society, Kenilworth, South Africa.

RZHETSKY, A. & NEI, M. 1992. A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution* **9**: 945-967.

SAITOU, N. & NEI, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406-425.

SWANEVELDER, Z. H. 2003. Diversity and population structure of *Clivia miniata* Lindl. (Amaryllidaceae): Evidence from molecular genetics and ecology. MSc Dissertation, University of Pretoria.

SWANEVELDER, Z. H., TRUTER, J. T. & VAN WYK, A. E. 2006. A natural hybrid in the genus *Clivia*. *Bothalia* **36**: 77-80.

TAMURA, K., DUDLEY, J., NEI M. & KUMAR S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (submitted).

TRUTER, J. T., SWANEVELDER, Z. H. & PEARTON, T. N. 2007. *Clivia x nimbicola – a Stunning Beauty from the Bearded Man*. In J.T Truter, Z.H Swanevelder & T.N. Pearton, Clivia *Yearbook* **8**: 23-27. *Clivia* club, Kirstenbosch National Botanical Garden, Cape Town.

VAN WYK, A. E. & SMITH, G. F. 2006. *Regions of floristic endemism in southern Africa. A review with emphasis on succulents*. Umdaus Press, Hatfield, Pretoria.

Chapter 4:

THE USE OF CROSS-SPECIES MARKERS IN *CLIVIA CAULESCENS*

Suzanne Stegmann

4.1 ABSTRACT

In an attempt to detect more variation among different *Clivia caulescens* specimens, microsatellite markers were tested on several specimens. Markers developed for *Phaedranassa tunguraguae*, *Hymenocallis coronia* and *Clivia miniata* were used as cross-species markers on *Clivia caulescens*, as they all belong to the family Amaryllidaceae. All attempts during this study to amplify STRs and test allelic diversity in 13 microsatellite loci for 20 specimens failed.

4.2 INTRODUCTION

The markers of choice for plant and animal systems are microsatellites or Simple Sequence Repeats (SSRs), due to their abundance, co-dominance, hypervariablity and transportability between species and multi-allelic characters (Roeder *et al*., 1995; Powell *et al*., 1996; Gupta & Varshney 2000). Development of a microsatellite marker system for a new species requires isolation, cloning, sequencing and characterization of microsatellite loci (Yashodha *et al*., 2005). For the improvement of the efficiency of microsatellite isolation, several procedures are available for the enrichment of microsatellites in a genomic library (Zane *et al*., 2002). In a genome with no or few DNA sequence information, an alternative approach is the sourcing of SSR primers developed for other species (Yashodha *et al*., 2005). Through the screening of primers from different sources, this approach offers a potential for low cost development of SSR markers for species with very little or no genomic information. Rosetto (2001) reviewed the status of transferability of SSR markers to related taxa. Potential transferability of SSR primers across species of the same family was reported in Leguminaceae (Peakall *et al*., 1998), Myrtaceae (Zucchi *et al*., 2002) and Fagaceae (Aldrich *et al*., 2003).

Markers developed for *Phaedranassa tunguraguae* (Oleas *et al*., 2005), *Hymenocallis coronia* (Markwith & Scalon, 2007) and *Clivia miniata* (Swanevelder, 2003) were used as cross-species markers on *Clivia caulescens*, as they all belong to the family Amaryllidaceae. All the species of the genus *Clivia* has a somatic chromosome number of 22 (Gouws, 1949; Ran *et al*., 1999; Murray *et al*., 2004). Although a chromosome number has not yet been determined for *Phaedranassa tunguraguae*, a somatic chromosome number of $2n = 46$ has been determined for both *P. dubia* (Brandham & Durodie, 1981) and *P. viridiflora* (La Cour

& Wells, 1973). *Hymenocallis coronia* has a somatic chromosome number of 44 (Flory, 1976).

The aim of this study was to test if cross-species markers can be used in determining genetic variation in *Clivia caulescens*.

4.3 MATERIALS AND METHODS

4.3.1 Specimens and DNA isolation.

Genomic DNA was extracted from leaves of various *Clivia caulescens* populations and outgroups taxa (Table 4.1). The DNA extractions were based on a method described by Rogstad (1992), with a few modifications. Fresh leaves, leaves stored in CTAB (Hexadecyltrimethyl Ammonium Bromide) or silica gel were used. Genomic DNA was visualized under UV light and were photographed with the Gel Doc 100 system using the software program Molecular analyst® Software 1.4.1 (Bio Rad Laboratories) (Appendix B).

Table 4.1: List of the plant specimens included in this study, indicating their geographical origin and voucher numbers. All voucher material is housed in the Geo Potts Herbarium (BLFU).

Species	**Locality**	**Voucher number**
C. caulescens	Bearded Man Mountain	*Spies 8565, 8567, 8569, 8571, 8701*
	God's Window	*Spies 8479, 8480, 8481, 8482, 8483*
	Magoebaskloof	*Spies 8892*
	Mariepskop	*Spies 8494, 8495, 8496*
	Soutpansberg	*Spies 8640*
	Swaziland	*Spies 8644, 8757, 8784*
	Wolkberg	*Spies 8487*
	Wonderview	*Spies 8596*

C.gardenii	Greytown	*Spies 8418*
C. miniata	Dwesa	*Spies 8574*
C. mirabilis	Donkerhoek	*Spies 8267*
C. nobilis	Keiskamma	*Spies 8254*
C. robusta	Port Shepstone	*Spies 8440*
C. xnimbicola	Bearded Man Mountain	*Spies 8578*

4.3.2 Microsatellite markers

Eight SSR primer sets developed for *Phaedranassa tunguraguae* (Oleas, 2005), four primer sets for *Hymenocallis coronaria* (Markwith & Scalon, 2007) [both members of the family Amaryllidaceae to which *Clivia* also belong] were used in this study, as well as four primer sets developed for *Clivia miniata* (Swanevelder, 2003). Primer names, sequences and corresponding annealing temperatures are listed in Table 4.2.

Table 4.2: Different primers used as cross-species markers in this study.

Locus name	Primer Sequence	T_a	Repeat Motif	Allele size	Fluorescent label
CLV1	F: CAATAATGTGGCTAATGGGTTG	53 ºC	$(T)_4(AT)_6$	± 200 bp	VIC
	R: CTCAAGCTATGCATCCAACG				
CLV2	F: CTTGTTGTAGCTTGTAATAGC	51 ºC	$(GT)_9$	± 225bp	6FAM
	R: CTGAACGGCAGAGGAGTTG				
CLV3	F: ACAACTCCTCTGCCGTTCAG	51 ºC	$(A)_{11}$	± 246 bp	PET
	R: GGGTGCAGTGCACTAGTGC				
CLV4	F: GCATCCCTTGCTCCTCTAC	55 ºC	$(CCT)_2TCT(CCT)_2CGT$	± 210 bp	NED
	R: CTCAAGCTATGCATCCAACG				
[1]Pt4	F: TCCTTGTATCGTATGCTCC	56ºC	$(CT)_{23}$	105-250	NED
	R: CAAACGCTGTATCCCCTTC				
[1]Pt 9	F: TCCTTGTATCGTATGCTCCC	56ºC	$(GA)_{17}$	87-125	VIC
	R: CAAACGCTGTATCCCCTTC				
[1]Pt14	F: GGAGGATGGTAGTACCATGAAC	55ºC	$(GA)_{14}$	153-191	6FAM
	R: TGTATGGTTGGGTATGGGAAC				
[1]Pt36	F: AGAGAATGTGATGGGAGAGAG	52ºC	$(GA)_{22}$	178-199	NED
	R: TCTTCCTTATCCCCTCCACC				
[1]Pt39	F: TCAAAACACTCATACCAACACC	50ºC	$(CA)_{10}$	232-264	VIC
	R: CCTCTCTCTCCAAACTCTCTC				

[2]HcoA10	F: TATGAGTTGAAGTGGAGTTGCA	47ºC	$(TGA)_6$	214-235	PET
	R: ATCCTCCATGATGATGAGACCCAA				
[2]HcoD7	F: AAGCTATGGATCGAGTAGGCCTG	52ºC	$(TGA)_3N_{42}(TGA)_4$	189-195	VIC
	R: CCCTAGAAGGTTATGCTTCCCACA				
[2]HcoD9	F: CCACAGAGAATCCAGGTTCCTA	58ºC	$(CA)_3N_{14}(CA)_4N_{29}(CA)_5$	245-257	6FAM
	R: ACATTCACACACTCACGCCTA				
[2]HcoD12	F: CACTGCAAGTGACACTTACCA	56ºC	$(AGA)_9$ $(TGA)_6$	202-238	6FAM
	R: AATCGCAAAGACCTGCCAA				

[1]Primers developed for *Phaedranassa tunguraguae* (Oleas, 2005)

[2]Primers developed for *Hymenocallis coronaria* (Markwith & Scalon, 2007)

4.3.3 Polymerase Chain Reaction (PCR).

The Taguchi method (Cobb & Clarkson, 1994) was used for optimising the PCR reactions. All 10µl PCR reactions were performed on either an Applied Biosystems GeneAmp PCR system 9600 or Applied Biosystems 2720 Thermal Cycler. The cycle conditions were a 3 min 95ºC denaturation step, followed by 35 cycles of denaturation for 15s at 95ºC, 15 s at the appropriate annealing temperature (Table 4.2) and a 45 s elongation step at 72ºC. The 35 cycles were followed by a 4 min and 45 s step at 72ºC and a holding step at 4ºC.

The forward primers were labelled with different fluorescent labels at the 5' prime end, to allow for multiplexing. The SSR reaction products were evaluated for polymorphisms on 2% agarose gels, which were stained with 2 µl ethidium bromide per gel.

4.3.4 Fragment scoring and cluster analysis.

The SSR gel images were scanned using the Gel Doc 2000 Bio-Rad system and analyzed with Quantity One Software v. 4.0.1 (Bio-Rad Laboratories). Fragments were sized using Genemapper (Applied Biosystems).

A mixture of diluted PCR amplified products, Hi-Di Formamide and LIZ 500 internal size standard was run on an ABI 3130 DNA analyser for fragment analysis. The software Genemarker (SoftGenetics LLC, 2004-2005) was used to detect and analyse microsatellites.

4.4 RESULTS & DISCUSSION

The primer sets were used on a panel of 20 specimens of *Clivia caulescens*, to analyse the possibility of amplifying homologous and polymorphic loci. As seen in Figure 4.1 some primers had one or more than one fragment.

An obvious problem when evaluating cross-species amplification is to determine whether or not a homologous locus has been detected (Primmer *et al*., 1996). According to Primmer *et al*. (1996), all cases should be considered when one (homozygote) or two (heterozygote) bands of similar size to the control samples were observed, as a successful detection of a homologous locus.

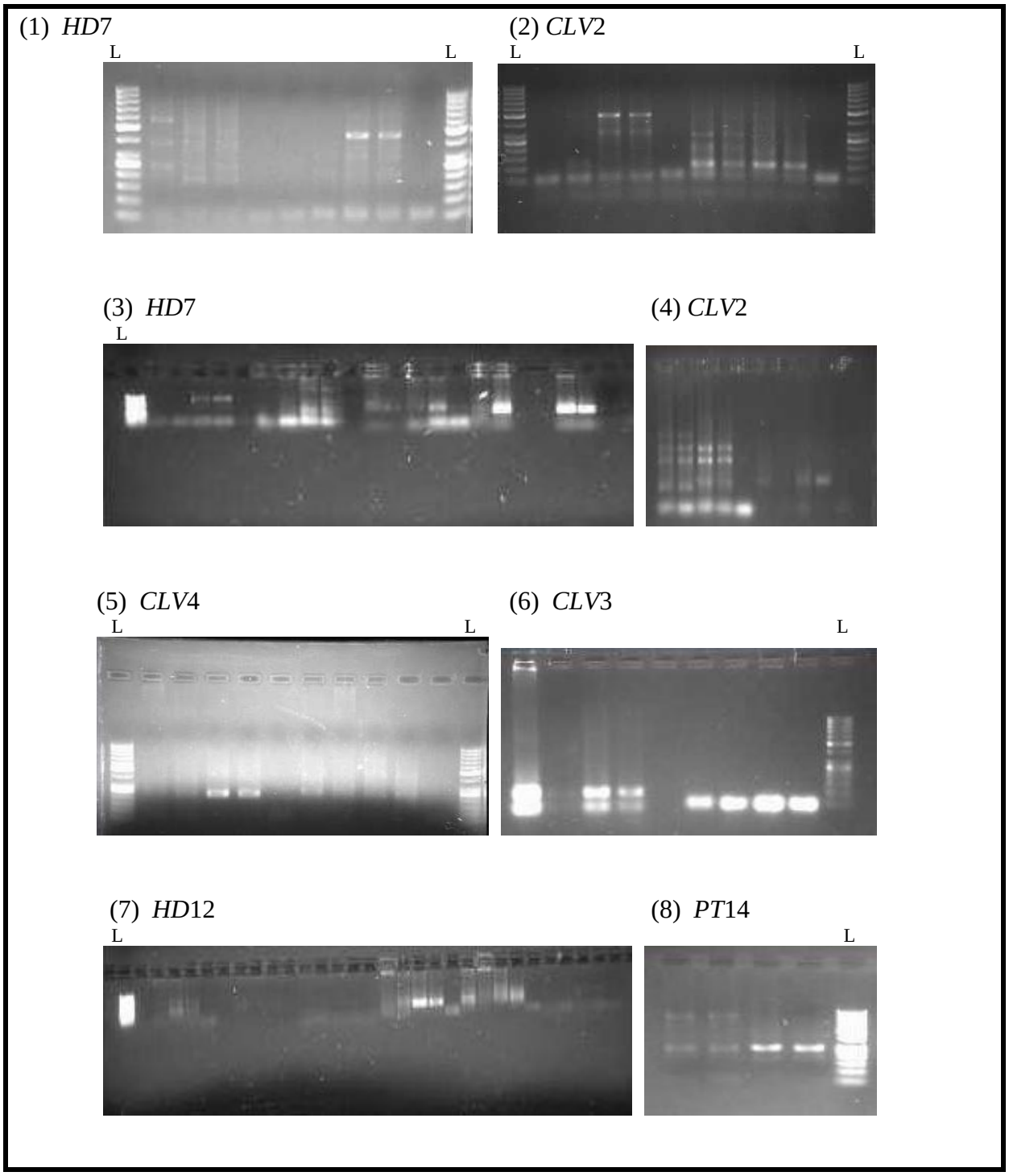


Figure 4.1: Profiles of the different *C. caulescens* populations as revealed by SSRs. Multiple bands detected by SSR primers. Image 1 & 3 - *HD*7; Image 2 & 4 - *CLV*2; Image 5 – *CLV*4; Image 6 – *CLV* 3; Image 7 – *HD*12 and Image 8 – *PT*14, respectively. L: DNA ladder marker (10 000-bp) (GeneRuler DNA Ladder Mix, ready to use (Fermentas)).

There were no cross-amplification tests done on *Phaedranassa tunguraguae* but tests for cross-amplification with a congener and a con-familial species were conducted by Markwith & Scalon (2007) and five *Zephyranthes candida* samples were tested with the primers developed for *Hymenocallis coronaria*. Allele sizes were within the range of sizes found for *Hymenocallis coronaria*, therefore not only were these microsatellite markers polymorphic, but may be useful across the genus and outside the genus with other species within the *Amaryllidaceae*. However, in the current study no amplification for the tested microsatellites was achieved in *Clivia caulescens*. This may possibly be attributed to genetic (DNA) differences between *Clivia* and the genus for which the primers were designed, *Hymenocallis*. This hypothesis may be supported by the fact that *Clivia* is endemic to South Africa and *Hymenocallis* to the Americas. Phylogenetically, the two genera segregated into two different subfamilies (Chase *et al*., 2009).

The different evolutionary pathways are also supported by chromosomal differences and variation in genome sizes (2C-value) (Zonneveld, 2001). All species of *Clivia* studied have 22 chromosomes, 2n = 22 (Gouws, 1949; Ran *et al*., 1999; Murray *et al.*, 2004). Zonneveld & Van Iren (2001) described DNA 2C-values were estimated using flow cytometry with the fluochrome propidium iodide (PI), a stain for DNA. Taxa with similar chromosome numbers have been verified to vary in genome size (Zonneveld & Van Iren, 2001; Zonneveld, 2001). The genome sizes for *Clivia miniata* and *C. caulescens* is respectively, 39.2 pg and 38.7 pg (Zonneveld, 2001) and no genome size is available for *Clivia xnimbicola*. The nuclear 2C value of *Phaedranassa tunguraguae* and *Hymenocallis coronia* has not been measured. Thus it is not possible to compare *Clivia* with the genera for the developed primers.

Another reason for failure to obtain amplification of the correct microsatellite region during this study may be of a procedural nature. Although all attempts were made to follow standard procedures and to optimize them, differences in standard operating procedures may have negatively influenced the results. We could deduce from the electropherograms that not any peaks within the theoretical allele size were obtained. The peaks obtained cannot with certainty be said to be microsatellites. A possibility is null alleles. A common cause of heterozygote deficit is amplification failure of certain alleles at a single locus, namely null alleles (Selkoe & Toonen, 2006). Null alleles fail to amplify in a PCR, either because the PCR conditions are not ideal or the primer-binding region contains mutations that inhibit binding (Selkoe & Toonen, 2006). A possible reason for the cross-species markers not being

as successful as one would have wished is the possibility of mutations in the flanking regions of the primers, and thus not binding to the DNA strand where it was supposed to. The success of cross-species markers rely on the nucleotide sequences of the flanking regions being conserved (Selkoe & Toonen, 2006). For confirmation that null alleles did occur, the reactions were repeated several times at different annealing temperatures and information was obtained from studies focused on the optimization of the use of cross species markers (Brownie *et al*., 1997 and Niens *et al*., 2005) and were integrated in our protocol.

When all the primers were amplified with the annealing temperature of 56ºC, stutter peaks were produced and therefore the annealing temperature was reduced. Such modifications in especially annealing temperature, in the PCR protocols are recommended for the successful transfer of microsatellite primer pairs between species (Rosetto, 2001).

Regarding the use of the *Clivia miniata* primers developed by Swanevelder *et al*. (2003), we did get amplification for some of the *Clivia caulescens* specimens, but no microsatellites could be located in the electropherograms. Some problems encountered were that primer sets CLV2 and CLV3 bound unspecific, resulting in numerous fragments being produced. Furthermore, due to insufficient flanking sequences, only four primer sets could be developed. Primer set CLV3 and CLV4 produced a fragment (band) that was longer than their designed length. According to Swanevelder (2003) two primer sets namely CLV2 and CLV4 showed polymorphisms between the *C. miniata* samples from different populations, which make them ideal for population studies of *C. miniata*. Swanevelder proposed that although the other two primer sets did not reveal any polymorphisms between different populations that it could still be of use for other species of the genus, *Clivia*, but we were unable to amplify any microsatellites for *C. caulescens*.

Due to time constraint this aspect of the study could not be completed.

4.6 CONCLUSION

In conclusion there could be a number of reasons why cross-species amplification was not as effective as hoped. It could possibly be because of ancient polyploidy of *Clivia*, however the genetic map of *C. caulescens* is not available and the inherited traits through the previous lineages are unknown. Microsatellites' utility can be reduced by mutations which alter the

allelic identity and priming sites (Colson & Goldstein, 1999; Noor *et al*., 2001). The mutational processes of microsatellites can be complex (Schlotterer, 2000; Beck *et al*., 2003; Ellergen, 2004).

Alternatively the microsatellites species-specific nature could have a negative effect on obtaining results, although other researchers (as mentioned in the introduction) have been able to employ cross-species markers successfully. Glen & Schabble (2005) reported that a given pair of microsatellite primers rarely works across broad taxonomic groups, and so primers are usually developed anew for each species. There is also a considerable failure rate in the isolation of new markers for marine invertebrates (Cruz *et al*., 2005), lepidopterans (Meglecz *et al*., 2004) and birds (Primmer *et al*., 1997).

The next step would be to attempt designing primers specifically for *Clivia caulescens*.

4.7 REFERENCES

ALDRICH, P. R., JAGTRAP, M., MICHLER, C. H. & ROMERO-SEVERSON, J. 2003. Amplification of North American red oak microsatellite markers in European white oaks and Chinese chestnut. *Silvae Genetics* **52**: 176-179.

BECK, N., DOUBLE, M. C. & COCKBURN, A. 2003. Microsatellite evolution at two hypervariable loci revealed by extensive avian pedigrees. *Molecular Biology and Evolution* **20**: 54-61.

BRANDHAM P. E., & DURODIE, J. 1981. Chromosome variation involving ditelosomy in *Phaedranassa dubia* (Amaryllidaceae). *Kew Bulletin* **36**: 213-215.

BROWNIE, J., SHAWCROSS, S., THEAKER, J., WHITCOMBE, D., FERRIE, R., NEWTON, C. & LITTLE, S. 1997. The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Research* **16**: 3235-3241.

CHASE, M. W., REVEAL, J. L. & FAY, M. F. 2009. A subfamilial classification for the expanded asparagalean families Amaryllidaceae, Asparagaceae and Xanthorrhoeaceae. *Botanical Journal of Linnean Society* **161**: 132-136

COBB, B. D. & CLARKSON, S. A. 1994. A simple method for optimizing the polymerase chain reaction using taguchi methods. *Nucleic Acids Research* **22**: 3801-3805.

COLSON I., MACDONALD, D. S & GOLDSTEIN, D. B. 1999. Microsatellite markers for interspecific mapping of *Drosophila simulans* and *D. Sechellia. Molecular Ecology* **8** (11): 1951-1955.

CRUZ, F., PEREZ, M. & PRESA, P. 2005. Distribution and abundance of microsatellites in the genome of bivalves. *Gene* **346**: 241-247.

ELLERGEN, H. 2004. Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics* **5**: 435-445.

FLORY, W. S. 1976. The distribution and chromosome numbers and types of various species and taxa of *Hymenocallis. Nucleus* **19**: 204-227.

GLENN, T. C. & SCHABLE, N. A. 2005. Isolating microsatellite DNA loci. *In* Zimmer, E. A. & Roalson, E. [eds.], *Molecular Evolution: Producing the Biochemical Data, Part B,* 202-222, Academic Press, San Diego, USA.

GOUWS, J. B. 1949. Karyology of some South African *Amaryllidaceae*. *Plant Life Herbertia* **5**: 54-80.

GUPTA, P. K & VARSHNEY, R. K. 2000. The development and use of microsatellite markers for genetic analysis and plant breeding with the emphasis on bread wheat. *Euphytica* **113**: 163-185.

LA COUR, L. F. & WELLS, B. 1973. Deformed lateral elements in synaptonemal complexes of *Phaedranassa viridiflora*. *Chromosoma* **41**: 289-296.

MARKWITH, S. H. & SCALON, M. J. 2007. Multi-scale analysis of *Hymenocallis coronaria* (*Amaryllidaceae*) genetic diversity, genetic structure, and gene movement under the influence of unidirectional stream flow. *American Journal of Botany* **94**: 151-160.

MEGLECZ, E., PETENIAN, F., DANCHIN, E., D'ACIER, A. C., RASPLUS, J. Y. & FAURE, E. 2004. High similarity between flanking regions of different microsatellites detected with each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas arinia*. *Molecular Evolution* **13**: 1693-1700.

MURRAY B. G, RAN, Y., DE LANGE E. P. J., HAMMETT K. R. W., TRUTER J. T. & SWANEVELDER Z. H. 2004. A new species of *Clivia* (Amaryllidaceae) endemic to the Pondoland Centre of Endemism, South Africa. *Botanical Journal of the Linnean Society* **146**: 369-374.

NIENS, M., SPIJKER, G. T., DIEPSTRA, A. & TE MEERMAN, G. J. 2005. A factorial experiment for optimizing the PCR conditions in routine genotyping. *Biotechnology and Applied Biochemistry* **42**: 157-62.

NOOR, M. A. F., KLIMAN, K. M. & MACHADO, C. A. 2001. Evolutionary history of microsatellites in the Obscura group of Dorsophila. *Molecular Biology and Evolution* **18**: 551-556.

OLEAS, N. A. 2005. Isolation and characterization of eight microsatellite loci from *Phaedranassa tunguragae* (Amaryllidaceae). *Molecular Ecology Notes* **5**: 791-793.

PEAKALL, R., GILMORE, S., KEYS, W., MORGANTE, M. & RAFALSKI, A. 1998. Cross-species amplification of Soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for transferability of SSRs in plants. *Molecular Biology and Evolution* **15**: 1275-1287.

POWELL, W., MACHRAY, G. C. & PROVAN, J. 1996. Polymorphism revealed by simple sequence repeates. *Trends in Plant Scientific* **1**: 215-222.

PRIMMER, C. R., MOLLER, A. P. & ELLERGEN, H. 1996. A wide-range survey of cross-species microsatellite amplification in birds. *Molecular Ecology Notes* **5**: 365-378.

PRIMMER, C. R. RAUDSEPP, T., CHOWDHARY, B. P., MOLLER, A. R. & ELLERGEN, H. 1997. Low frequency of microsatellites amplification in birds. *Molecular Ecology* **5**: 365-378.

RAN, Y., HAMMETT, K. R. W. & MURRAY, B. G. 1999. Karyotype analysis of the genus *Clivia* by Giemsa and fluorochrome banding and *in situ* hybridization. *Euphytica* **106**: 167-174.

ROEDER M. S., PLASCHKE, J., KONIG, S. U., BORNER, A., SORRELLS, M. E., TANKSLEY, S. D. & GANAL, M. W. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics* **246**: 327-333.

ROGSTAD, S. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon* **41**: 701-708.

ROSETTO, M. 2001. Sourcing of SSR markers from related plant species. *In* Henry, R. J. [ed]. *Plant genotyping: DNA fingerprinting of plants*, 210-224. CAB International, Oxford.

SELKOE, K. A. & TOONEN, R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**: 615-629.

SHLÖTTERER, C. 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**: 365-371.

SoftGenetics LLC 2004-2005, *Operation Manual for GeneMarker*, *State College PA 16803, www.softgentics.com*

SWANEVELDER, Z. H. 2003. Diversity and population structure of *Clivia miniata* Lindl. (Amaryllidaceae): Evidence from molecular genetics and ecology. MSc Dissertation, University of Pretoria.

YASHODHA, R. GHOSH, M., SUMATHI, R. & GURUMURTHI, K. 2005. Cross-species amplification of eucalyptus SSR markers in *Casuarinaceae*. *Acta Botanica Croatica* **64**: 115-120.

ZANE, L., BARGELLONI, L. & PATARNELLO, T. 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology Notes* **11**: 1-16.

ZONNEVELD, B. J. M. 2001. Nuclear DNA content of all species of *Helleborus* (Ranunculacea) discriminates between species and sectional divisions. *Plant Systematics and Evolution* **229**: 125-130.

ZONNEVELD, B. J. M. & VAN IREN, F. 2001. Genome size and pollen viability as taxonomic criteria: Application to the genus *Hosta*. *Plant Biology* **3**: 176-185.

ZUCCHI, M. I., BRONDANI, R. P. V., PINHEIRO, J. B., BRONDANI, C., & VENCOVSKY, R. 2002. Transferability of microsatellite markers from *Eucalyptus* spp. to *Eugenia dysenterica* (Myrtaceae family). *Molecular Ecology Notes* **2**: 512.

Chapter 5:

GENERAL DISCUSSION AND CONCLUSION

Suzanne Stegmann

5.1 DISCUSSION

Clivia caulescens grows on the edge of the escarpment from Soutpansberg in the north to Swaziland. Although specimens from different localities appear phenotypically different, the degree of genetic variation within these populations and between different populations has not been studied. In addition very little is known about the mechanism(s) involved in the production of *C. xnimbicola*, a natural hybrid between *C. miniata* and *C. caulescens* growing on the Bearded Man Mountain on the border between Mpumalanga and Swaziland.

Firstly, the aim of this study concerning techniques was whether SNPs or microsatellites can be used in *Clivia* to determine the degree of genetic variation in *C. caulescens*. Secondly, it was investigated if any genetic variation could be determined within and between different populations in this study. Thirdly, it was determined if gene flow occurred or not and fourthly whether molecular markers could be use to determine where a specimens' geographical origin was; what the genetic variation in the *C. xnimbicola*'s population was; what the correlation between the genetic variation of *C. xnimbicola*'s and its putative parents *C. caulescens* and *C. miniata* in the Bearded Man population was. Lastly it was investigated if *C. xnimbicola* is continuously formed by random pollination events.

SNPs through means of direct sequencing of four chloroplast gene regions proved to be successful in determining if there is any genetic variation within a population or between populations, especially when used in combination. A total of 82 basepairs differed and 24 of them were parsimony informative when the combined dataset was used in the *C. caulescens* populations included in this study. The region of *atpH-I* and *matK* also proved to be sufficient in determining genetic variation individually, but it is recommended to use these regions in combination with other gene regions for optimal results. It could also be detected from the results obtained that *rpoC1* did not reveal information. There were a total of 23 parsimonious informatives sites obtained by the combined dataset for the Clivias at the Bearded Man Mountain, which can indicate that, a combination of different regions is the approach to follow to investigate molecular diversity in *Clivia* populations. Three different indels and four SNPs clearly indicate that *C. caulescens* differ from the other species. Sufficient variation within *C. caulescens* exists to attempt the identification of specific specimens/localities.

From all the flow diagrams (Appendix F and G) it was detected that there is definitely gene flow between the different *C. caulescens* populations. Mariepskop was regarded as a

population on its own as this population was more different to the most of all the other populations. SNPs were shared by specimens from different populations which indicate that gene flow did occur between the populations in this study e.g. SNP 5 from the *atpH-I* region were shared by three of the Bearded Man specimens (*Spies 8567*, *Spies 8569*, *Spies 8571*); four of the God's Window specimens (*Spies 8480*, *Spies 8481*, *Spies 8482*, *Spies 8483*) and a Wonderview specimen; indicating that gene flow occurred between these populations. Gene flow was also indicated by the use of SNPs at Bearded Man Mountain, e.g. SNP 3 from the *atpH-I* region was shared between two *C. caulescens* specimens, three *C. miniata* specimens and one of *C. xnimbicola*.

From the Network Tree it was clear that the Mariepskop specimens group together. It appeared as if this group developed from the Northern Group (Soutpansberg, Magoebaskloof and Wolkberg). This was quite interesting since it is geographically much closer to God's Window. Geographical and climatic changes through time should be studied to understand this phenomenon. The Southern group (God's Window, Bearded Man and Swaziland) is geologically isolated from the northern range and this isolation is evident in these results.

Regarding the variation within the different populations of *Clivia caulescens* that have been examined, it was detected that Mariepskop population displayed the most variation within the population and God's Window population displayed the least variation within the population. The region of *atpH-I* was instrumental in detecting any mean genetic distances within the species found at the Bearded Man Mountain. The region of *rpoC1* however did not reveal any interspecific variation for the subpoplutions. *Clivia caulescens* displayed the most interspecific variation from all the species there. Although the region of *rpoC1* did not reveal any interspecific variation for either the *C. caulescens* or *C. miniata* species, it did reveal interspecific variation for their putative natural hybrid, *C. xnimbicola*.

In some instances molecular markers can assist with the identification of *C. caulescens* at different populations. Different combinations of SNPs were obtained for each specimen included in this study regarding the *C. caulescens* populations. Most of the specimens had different combinations of SNPs at the Bearded Man for *atpH-I* region, except for specimens 469 and 395 that shared the same SNPs; and specimens 139 and 257 that shared the same SNP.

Intraspecific variation was obtained for all the regions and combined dataset, except for the *matK* region for the putative natural hybrid of *C. caulescens* and *C. miniata* at the Bearded

Man Mountain. *Clivia xnimbicola* always clustered with *C. caulescens* and not with *C. miniata* in both the Maximum Parsimony and Minimum Evolutions' cladograms. Both of the putative parents had a mean distance of 0.009 to the natural hybrid. For the combined dataset, the *matK* and *rpoB* region, *C. caulescens* displayed not as much mean genetic distance diversity to *C. xnimbicola* as did *C. miniata*. When considering the phylogenetic analysis together with the statistical analysis, it is supported that the mean distance between *C. miniata* and *C. xnimbicola* is more, because *C. miniata* did not cluster with *C. xnimbicola* like *C. caulescens* did.

The unique SNPs obtained for the Bearded Man Clivias indicate that either the specific parent responsible for the hybridization process was not included in the study, or that new mutations occurred after the speciation event. The variation among *C. xnimbicola* specimens further suggests that this hybrid species was produced several times and does not represent a single hybridization event. The possibility of several different (genetically different) specimens forming the hybrid species seems consequently the most possible explanation for the differences in DNA sequences from different specimens. This also makes the probability of the "'correct parents' not included in the study hypothesis", more plausible.

All attempts during this study to amplify STRs and test allelic diversity in 13 microsatellite loci for 20 specimens failed. Cross-species amplification was not as effective as we hoped. Microsatellites' utility can be reduced by mutations which alter the allelic identity and priming sites and we don't know what traits were inherited through the previous lineages. Microsatellites' species-specific nature could have a negative effect on obtaining results, although other researchers (as mentioned in the introduction of Chapter 4) could employ cross-species markers successfully. Glen & Schabble (2005) reported that a given pair of microsatellite primers rarely works across broad taxonomic groups, and so primers are usually developed anew for each species. The next step would therefore be to attempt the designing of specific primers for *C. caulescens*.

5.2 CONCLUSION

The attempt to determine the genetic variation, using microsatellites,between and within the different populations of *C. caulescens* and the Clivia species at the Bearded Man Mountain was unsuccessful. It is suggested that specific microsatellite primers for *C. caulescens* should be designed in the future.

In conclusion, interspecific and intraspecific variation does not occur for most of the different *C. caulescens* populations as well as for the clivias found at the Bearded Man Mountain. The sequencing of a combination of different regions is the best approach to follow in order to investigate molecular diversity in *Clivia* populations.

Through the use of SNPs, statistic and phylogenetic analysis it was determined that gene flow does exist between the different *C. caulescens* populations in this study, as well as for the Bearded Man Mountain clivias.

The population at Mariepskop is genetically most different from the south located populations of *C. caulescens*, which is interesting because as mentioned before, Mariepskop is closer located to the southern populations than towards the northern populations.

Both the phylogenetic and statistical analysis (genetic distances) supported that the mean distance between *C. miniata* and *C. xnimbicola* is higher, because *C. miniata* does not cluster with *C. xnimbicola* like *C. caulescens* did. Therefore it can be concluded that *C. caulescens* is the motherplant of the natural hybrid and not as previously suggested *C. miniata*. The rationale why many *Clivia* enthusiasts may have considered that *C. miniata* is the motherplant is because of its flared trumped shaped flowers which form a large umbel and is much more approachable for pollination than that of *C. caulescens*' pendulous and tubular flowers. Furthermore it is speculated that maybe insects are responsible for the pollination of *C. xnimbicola* and not just pollen dispersal. However, the only way to get the final answer about this whole hybridization process will be to study numerous samples from the area to determine the degree of introgression.

5.3 REFERENCES

GLENN, T. C. & SCHABLE, N. A. 2005. Isolating microsatellite DNA loci. *In* Zimmer, E. A. & Roalson, E. [eds.], *Molecular Evolution: Producing the Biochemical Data, Part B*, 202-222, Academic Press, San Diego, USA.

Chapter 6:

SUMMARY/ OPSOMMING

Suzanne Stegmann

6.1 SUMMARY

At present, the genus *Clivia* consists of six species, including *Clivia nobilis* Lindl., *C. miniata* (Lindl.) Regel, *C. gardenii* Hook., *C. caulescens* RA Dyer, *C. mirabilis* Rourke and *C. robusta* Murray, Ran, De Lange, Hemmet, Truter & Swanevelder. Many of the species and cultivars are extensively grown worldwide, making this group of considerable horticultural importance.

This study mostly focused on *Clivia caulescens* with a natural habitat on the escarpment from Limpopo to Swaziland through Mpumalanga.

The overlapping distribution between *C. miniata* and *C. caulescens* resulted in the formation of a natural hybrid between these species at the Bearded Man Mountain. The occurrences of natural hybrids between the various species are rarely recorded.

In an attempt to find out if genetic erosion is currently a threat to the various *C. caulescens* populations and Bearded Man Mountain clivias, this study was conducted to establish if genetic variation is present.

Genetic variation refers to the variation in the genetic material of a population, and includes the nuclear, mitochondrial, ribosomal DNA as well as the DNA of other organelles. The relative genetic diversity among individuals or populations can be determined using morphological and molecular markers.

Five chloroplast DNA regions, i.e. *atpH-I*, *matK*, *rpoB*, *rpoC* and *trnL-F*, were used in an attempt to study the molecular diversity of *C. caulecsens*. This study concentrated on Single Nucleotide Polymorphisms (SNPs) from these regions and microsatellites to study genetic variation.

The aim of this study was to determine the genetic variation between and within the different populations of *C. caulescens*, to determine whether gene flow occur between the different populations and to determine which of the DNA regions included in the study can contribute to the identification of plants from a specific geographical area. Regarding the study of Clivias situated at the Bearded Man Mountain, the main objectives were to estimate genetic diversity and determine the genetic relationship among the different species of *Clivia* (*C. miniata*, *C. caulescens* and *C. xnimbicola*) from this area.

Of the initial five regions that were sequenced, *trnL-F* amplification failed repeatedly, and this region was therefore excluded from all analyses. The other four regions showed variation between the different populations of *C. caulescens* and for the Bearded Man Mountain clivias, except the *rpoC1* region.

When the results of the phylogenetics and statistical analysis (genetic distances) were combined, it was detected that most Bearded Man Mountain specimens and God's Window specimens clustered together in the cladograms and in the mean distances tables. Intraspecific variation was present in all the regions and combined dataset.

All attempts during this study to amplify STRs and test allelic diversity in 13 microsatellite loci for 20 specimens failed. Cross-species amplification was not as effective as hoped. Microsatellites' species-specific nature could have a negative effect on obtaining results, although other researchers (as mentioned in the introduction of Chapter 4) could employ cross species markers successfully. Glen & Schabble (2005) reported that a given pair of microsatellite primers rarely works across broad taxonomic groups, so primers are usually developed anew for each species. The next step would therefore be to attempt the designing of specific primers for *C. caulescens*.

KEY WORDS: *atpH-I*, Genetic variation, Hybridization, Introgression, *matK*, Microsatellite, Polymorphism, *rpoB*, *rpoC1*, Single Nucleotide Polymorphisms (SNPs).

6.2 OPSOMMING

Tans is daar ses spesies in die genus *Clivia*, naamlik *C. nobilis* Lindl., *C. miniata* (Lindl.) Regel, *C. gardenii* Hook., *C. caulescens* RA Dyer, *C. mirabilis* Rourke en *C. robusta* Murray, Ran, De Lange, Hemmet, Truter & Swanevelder. Al hierdie spesies en kultivars word wêreldwyd gekweek en is dus van groot tuinboukundige belang (Truter *et al.*, 2007).

Hierdie studie het gefokus op *C. caulescens*, met 'n geografiese verspreiding op die platorand van Limpopo na Swaziland, deur Mpumalanga. 'n Natuurlike baster tussen *C. caulescens* en *C. miniata* kom voor as gevolg van hulle simpatriese verspreidings op die "Bearded Man" berg. Die voorkoms van natuurlike basters tussen verskeie spesies is seldsaam en min voorbeelde is beskryf.

‘n Poging is aangewend om te bepaal of genetiese erosie in die verskillende populasies van *C. caulescens* wat bestudeer is, voorkom. Genetiese variasie verwys na die variasie in genetiese materiaal van ‘n populasie, en sluit kern, mitochondriale, ribosomale genome en genome van ander organelle in. Die relatiewe genetiese diversiteit onder individue of populasies kan bepaal word deur die gebruik van morfologiese of molekulêre merkers.

Vyf chloroplas DNS gebiede naamlik *atpH-I, matK, rpoB, rpoC* en *trnL-F*, is gebruik in ‘n poging om die molekulêre diversiteit van *C. caulescens* te bepaal. Die studie het gefokus op enkel nukleotied polimorfismes van die onderskeie gebiede en mikrosatelliete om genetiese variasie te ondersoek.

Die doel van die studie was om die genetiese variasie tussen en binne die verskillende populasies van *C. caulescens* te bepaal, asook of geenvloei bestaan tussen die verskillende populasies en om te bepaal watter van hierdie DNS gebiede kan bydrae tot die identifisering van plante van ‘n spesifieke geografiese gebied. Ten opsigte van die “Bearded Man” berg plante was die hoof mikpunte om genetiese variasie te bepaal en die genetiese verhouding van die verskillende spesies teenoor mekaar te ondersoek.

Van die aanvanklike gebiede waarvan die nukleotiedvolgordes bepaal is, het *trnL-F* se amplifisering herhaaldelik misluk en hierdie gebied is uitgelaat by alle analises. Die ander vier gebiede het ge-amplifiseer en verskille tussen al die populasies van *C. caulescens* en die spesies van “Bearded Man” berg aangedui. Die uitsondering was die *rpoC1* gebied wat geen verskille getoon het nie.

Die gekombineerdie filogenetiese en statistiese resultate (genetiese afstande) dui daarop dat die eksemplare van “Bearded Man” berg met die van God’s Window saam groepeer in die kladogramme en in die hoof afstandstabelle. Intraspesifieke variasie word deur al die gebiede en die gekombineerde datastel aangetoon.

Alle pogings gedurende die studie om 13 mikrosatelliete vir 20 eksemplare te amplifiseer, het herhaaldelik misluk. Kruis-spesie amplifisering was nie so effektief soos vooraf gereken nie. Mikrosatelliete se spesie-spesifieke aard kon ‘n negatiewe effek gehad het op die verkryging van resultate. Hoewel ander navorsers suksesvol was, is daar ook diegene wat nie sukses behaal het as gevolg van die moontlikheid dat sekere mikrosatellietmerkers nie oor groot taksonomiese groepe werksaam is nie en dat daar gewoonlik nuwe merkers spesifiek vir elke spesie ontwerp moet word. Die volgende stap sou dus wees om te poog om nuwe merkers,

spesifiek vir *C. caulescens*, te ontwikkel indien navorsers verder deur mikrosatelliete genetiese variasie in Clivias wil bepaal.

SLEUTELWOORDE: *atpH-I*, Enkele Nukleotied Polimorfismes ("SNPs"), Genetiese variasie, Introgressie, *matK*, Mikrosatelliet, Polimorfismes, *rpoB*, *rpoC1*, Verbastering.

APPENDIX A: Nanodrop readings of the *Clivia caulescens* specimens used in this study.

Voucher Number	Sample ID	ng/µl	A260	A280	260/280	260/230
Spies 8565	316	200.76	4.015	2.026	1.98	1.81
Spies 8567	318	75.06	1.501	0.751	2.00	1.73
Spies 8569	320	262.98	5.260	4.101	1.28	0.54
Spies 8571	322	188.57	3.771	2.647	1.42	0.71
Spies 8701	448	56.31	1.126	0.484	2.33	-7.63
Spies 8480	211	214,27	4. 285	2.207	1.95	1.39
Spies 8481	212	114,24	2. 285	1. 243	1.84	0.99
Spies 8482	213	472.12	9.442	5.048	1.87	1.28
Spies 8483	214	114,78	2 .296	1. 177	1. 95	1.34
Spies 8479	210	177.55	3.551	1.878	1.89	1.01
Spies 8893	639	99.02	1.980	1.172	1.69	0.61
Spies 8494	225	483.97	9. 679	6. 612	1.46	0.92
Spies 8495	226	566.71	11. 334	7 .966	1.42	0.85
Spies 8496	227	255.90	5.118	2.711	1.89	1.13
Spies 8644	391	111.86	2.237	1.162	1.93	1.64
Spies 8784	532	90.00	1.800	0.892	2.02	0.38
Spies 8757	505	472.12	9.442	5.048	1.87	1.28
Spies 8640	387	106.11	2.122	1.168	1.82	1.43
Spies 8487	218	86.20	1.719	0.924	1.86	0.69
Spies 8596	347	115.29	2.306	1.152	2.00	1.26

APPENDIX B: Gel documents of amplification products from all the populations represented in this study of *Clivia caulescens*. (GeneRuler DNA Ladder Mix - three different reference bands (3000, 1000 and 500bp)).

L 218 387 211 212 213 214 225 226

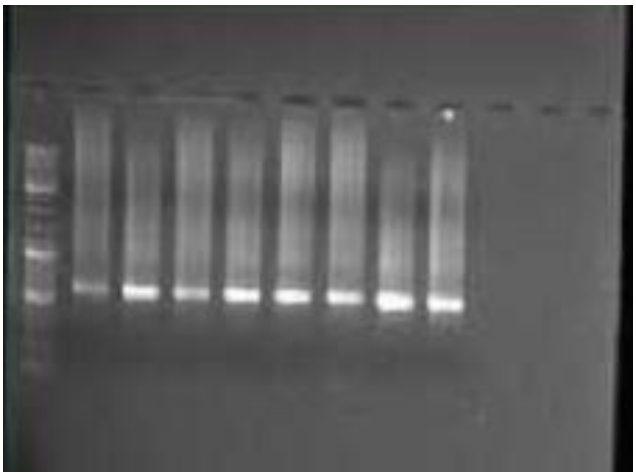


Fig.B.1: The photograph illustrates the fragments obtained after pre-sequencing ***atpH-I***. The numbers indicate the specimens listed in specimen list. L - ladder

L 227 639 347 391 210 448 505 532

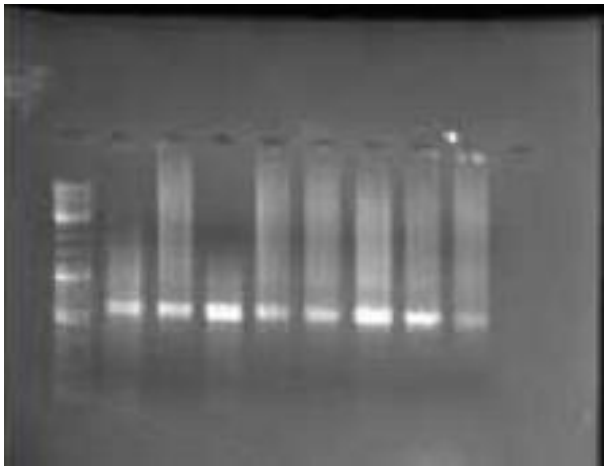


Fig.B.2: The photograph illustrates the fragments obtained after pre-sequencing ***atpH-I***. The numbers indicate the specimens listed in specimen list. L – ladder.

L 318 316 322 320 NA NA NA NA

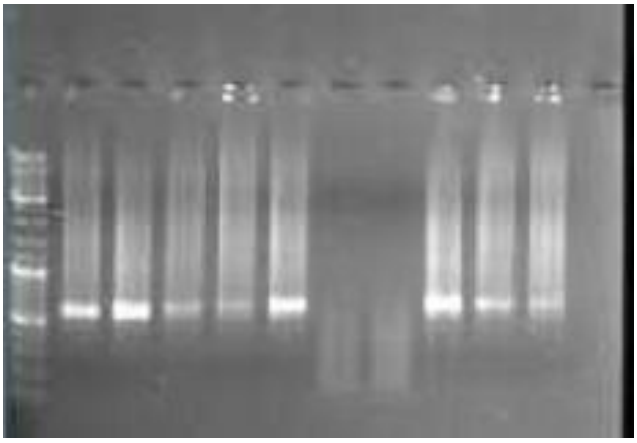


Fig.B.3: The photograph illustrates the fragments obtained after pre-sequencing ***atpH-I***. The numbers indicate the specimens listed in specimen list. L - ladder

L 225 226 227 218 347 639

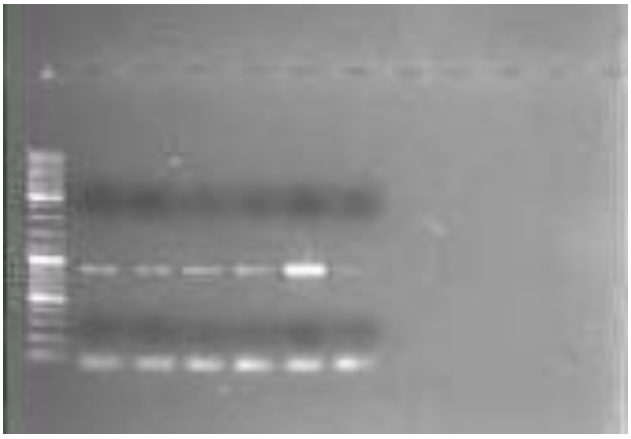


Fig.B.4: The photograph illustrates the fragments obtained after pre-sequencing ***matK***. The numbers indicate the specimens listed in specimen list. L – ladder.

*NA – Not applicable

L 210 505 387 391 532 213 212 214 211 NA 448

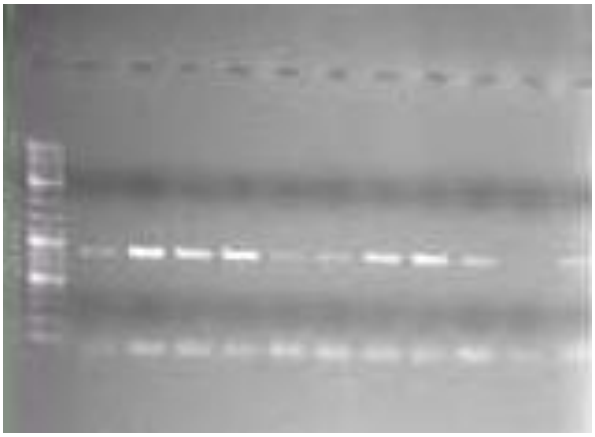


Fig.B.5: The photograph illustrates the fragments obtained after pre-sequencing ***matK***. The numbers indicate the specimens listed in specimen list. L - ladder

L 318 NA 316 322 320 NA NA NA NA

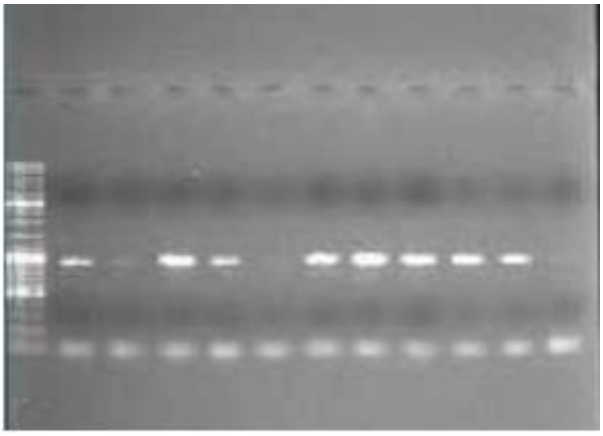


Fig.B.6: The photograph illustrates the fragments obtained after pre-sequencing ***matK***. The numbers indicate the specimens listed in specimen list. L – ladder.

L 213 212 214 211 225 226 387 211 347 639 227

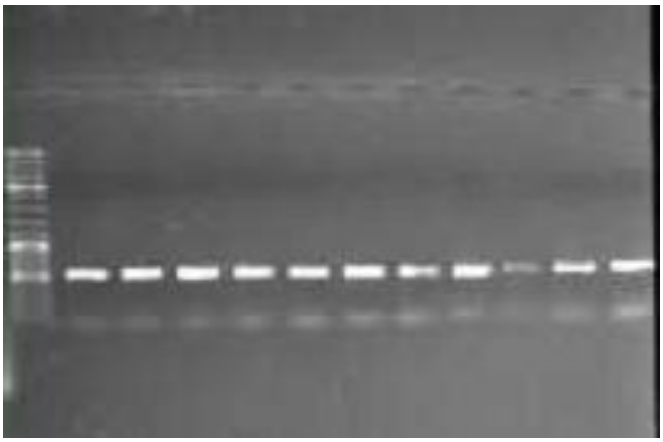


Fig.B.7: The photograph illustrates the fragments obtained after pre-sequencing ***rpoB***. The numbers indicate the specimens listed in specimen list. L - ladder

L 391 532 505 218

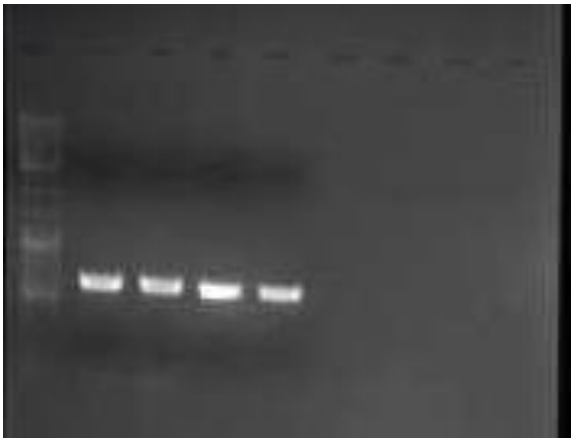


Fig.B.8: The photograph illustrates the fragments obtained after pre-sequencing ***rpoB***. The numbers indicate the specimens listed in specimen list. L – ladder.

*NA – Not applicable

L 316 318 320 322 210 448 NA NA NA NA

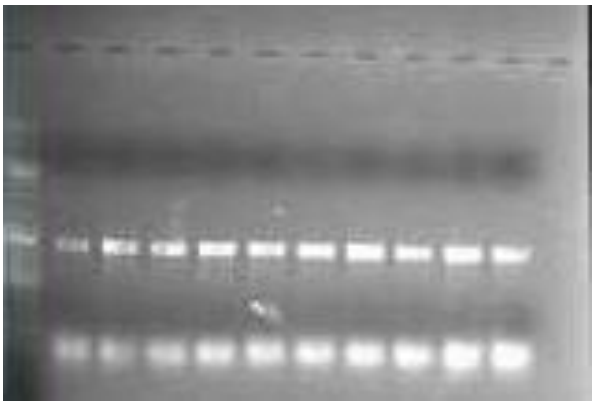


Fig.B.9: The photograph illustrates the fragments obtained after pre-sequencing ***rpoB***. The numbers indicate the specimens listed in specimen list. L - ladder

L 316 318 320 322 210 505 387 391 448 532

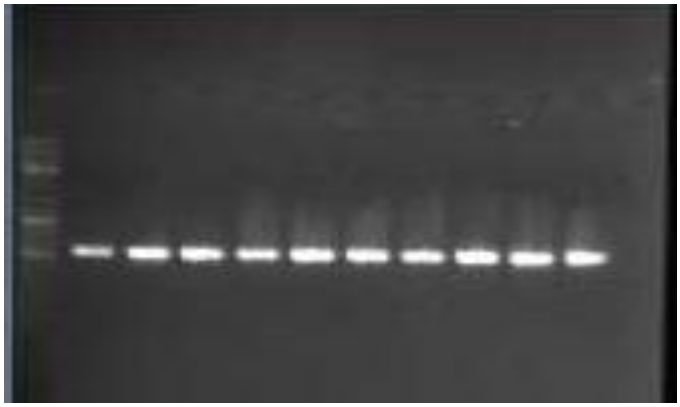


Fig.B.10: The photograph illustrates the fragments obtained after pre-sequencing ***rpoC1***. The numbers indicate the specimens listed in specimen list. L – ladder.

L 347 218 211 NA 213 225 212 214 226 227 639

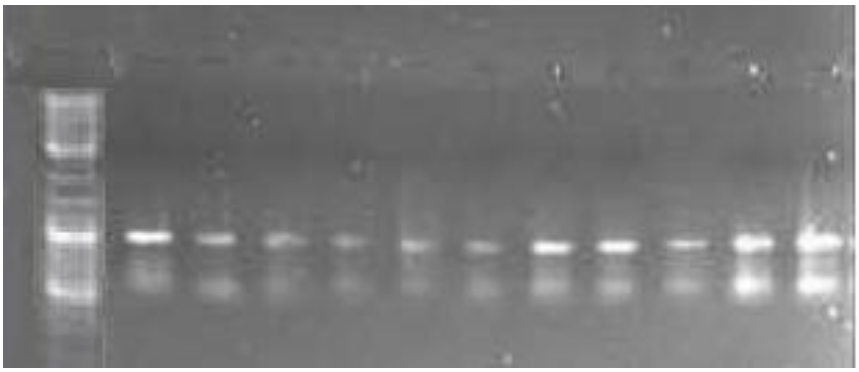


Fig.B.11: The photograph illustrates the fragments obtained after pre-sequencing ***rpoC1***. The numbers indicate the specimens listed in specimen list. L - ladder

*NA – Not applicable

218 387 211 212 213 639 391 532 505 320 322 210

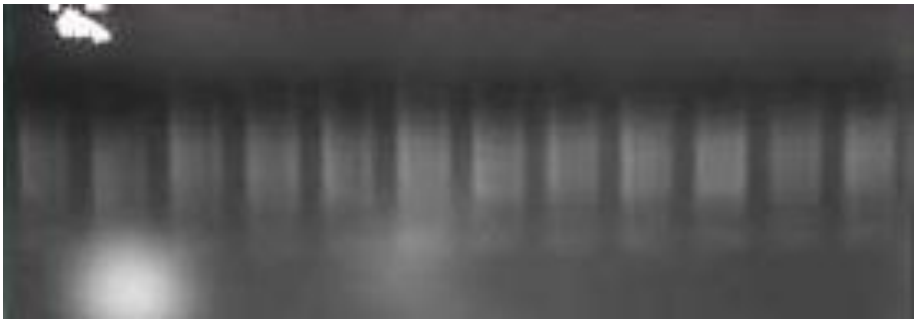


Fig.B.12: The photograph illustrates the fragments obtained after pre-sequencing ***trnL-F***. The numbers indicate the specimens listed in specimen list.

448 NA NA NA NA NA NA NA

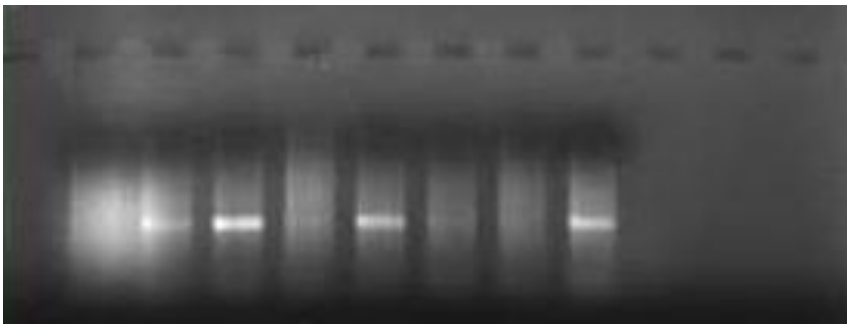


Fig.B.13: The photograph illustrates the fragments obtained after pre-sequencing ***trnL-F***. The numbers indicate the specimens listed in specimen list.

214 225 226 227 347 391 212 316 318

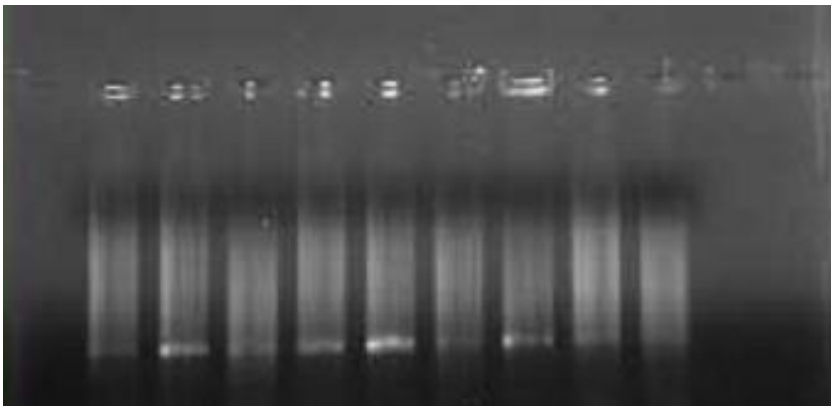


Fig.B.14: The photograph illustrates the fragments obtained after pre-sequencing ***trnL-F***. The numbers indicate the specimens listed in specimen list.

*NA – Not applicable

APPENDIX C: Geneious illustrations of the sequences obtained.

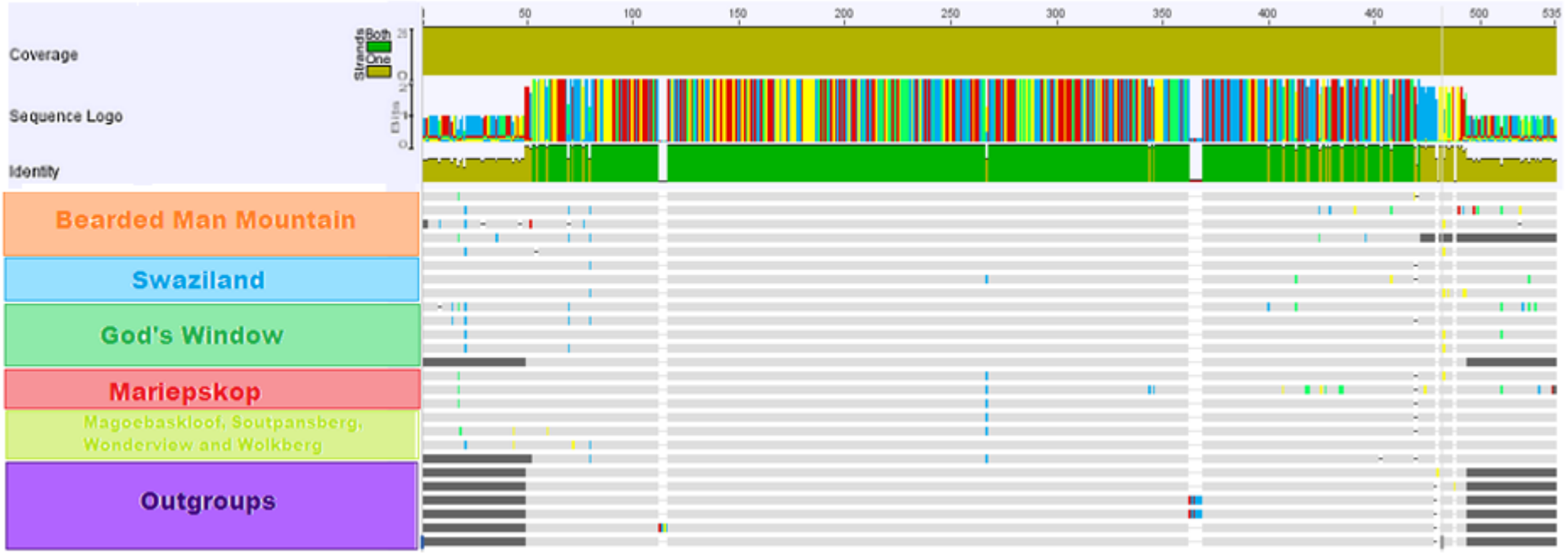


Figure C.1: Illustration of the ***atpH-I*** regions' GENEIOUS output with the conservative sites unmarked.

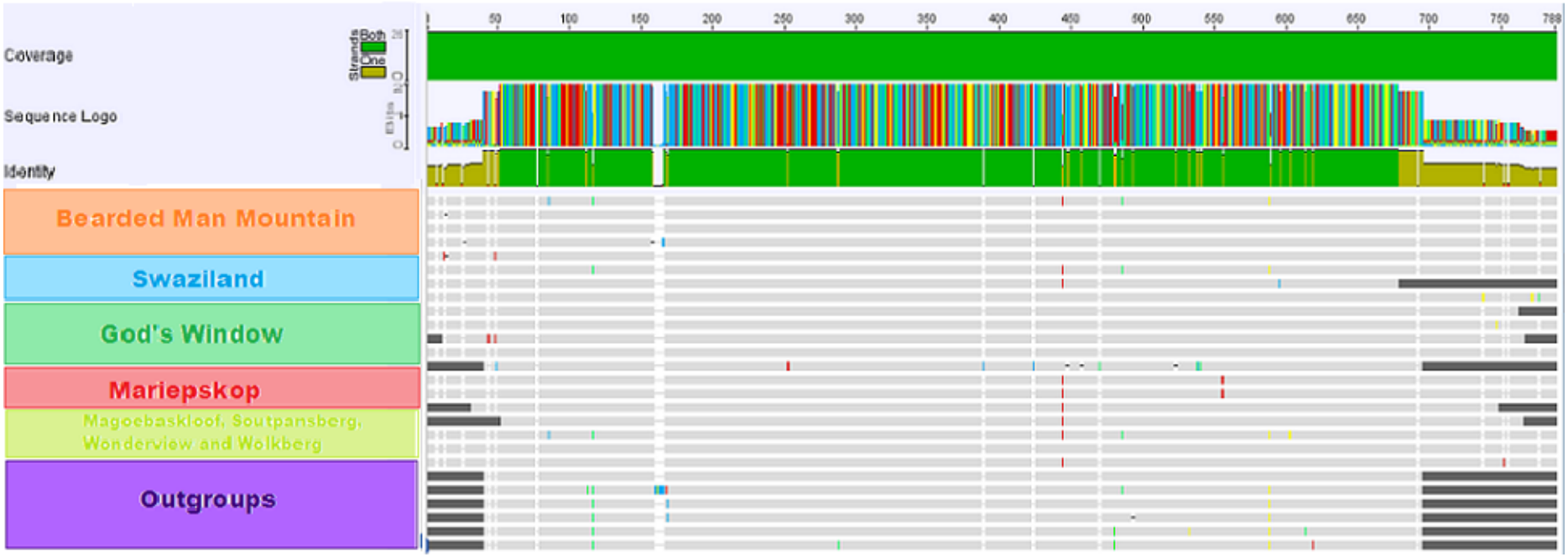


Figure C.2: Illustration of the ***matK*** regions' GENEIOUS output with the conservative sites unmarked.

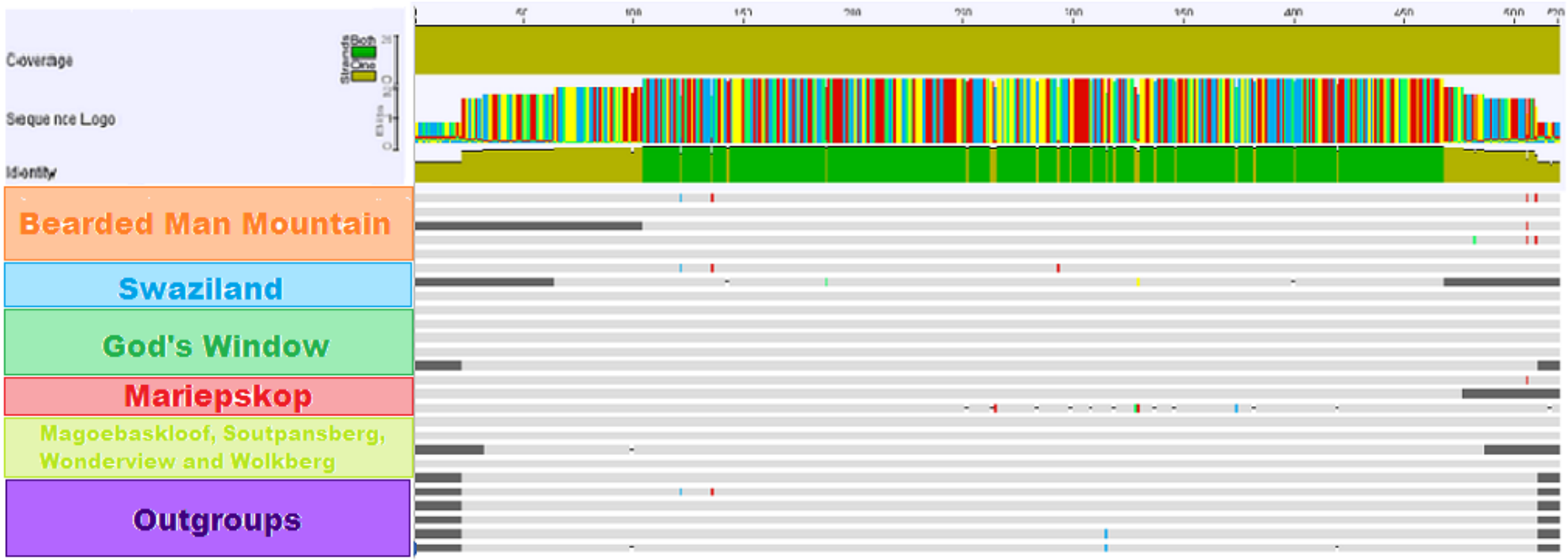


Figure C.3: Illustration of the ***rpoB*** regions' GENEIOUS output with the conservative sites unmarked.

The *rpoC* regions' GENEIOUS output revealed no relevant information, therefore it was not included here.

APPENDIX D: Cladograms obtained from the other gene regions (*atpH-I*; *matK*; *rpoB* and *rpoC1*) by making use of Maximum Parsimony and Minimum Evolution methods.

Figure D.1: Cladogram conducted through **Minimum Evolution** method for the *atpH-I*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

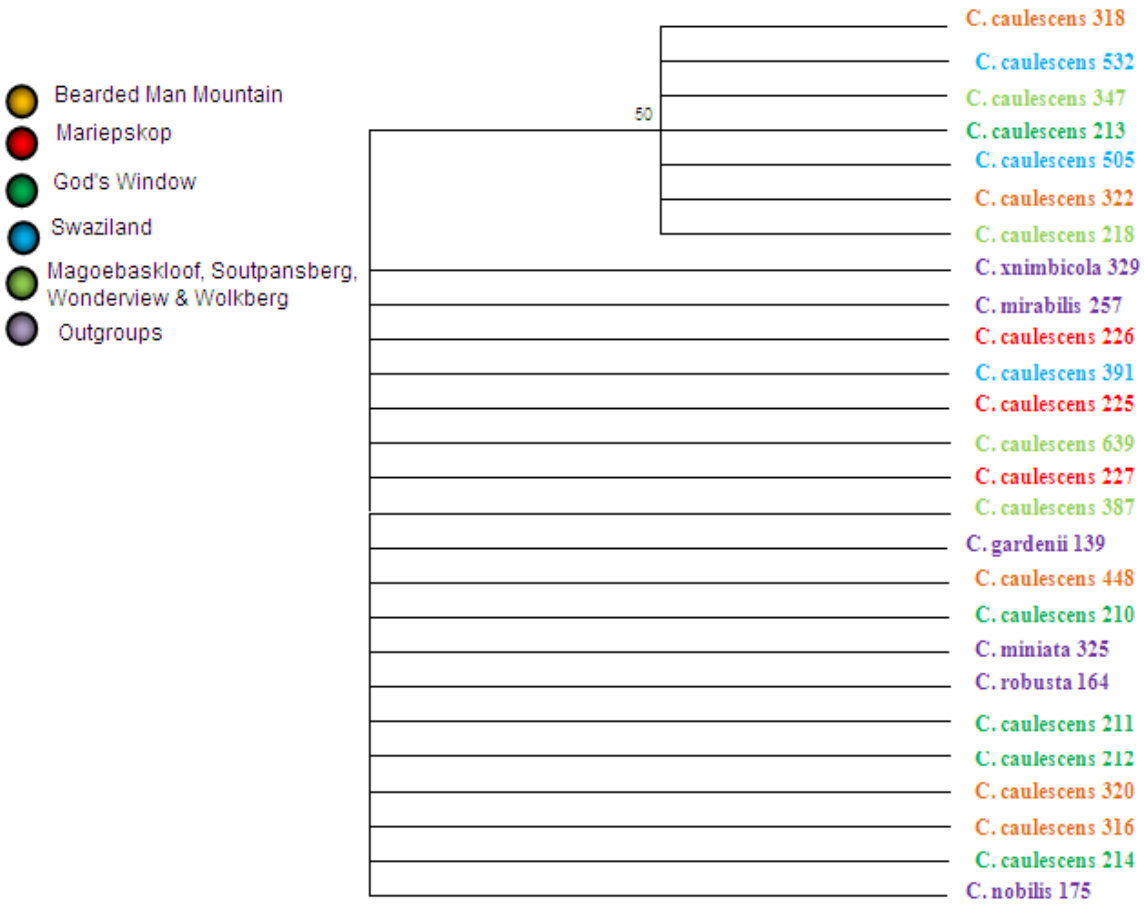


Figure D.2: Cladogram conducted through **Maximum Parsimony** method for the *atpH-I*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

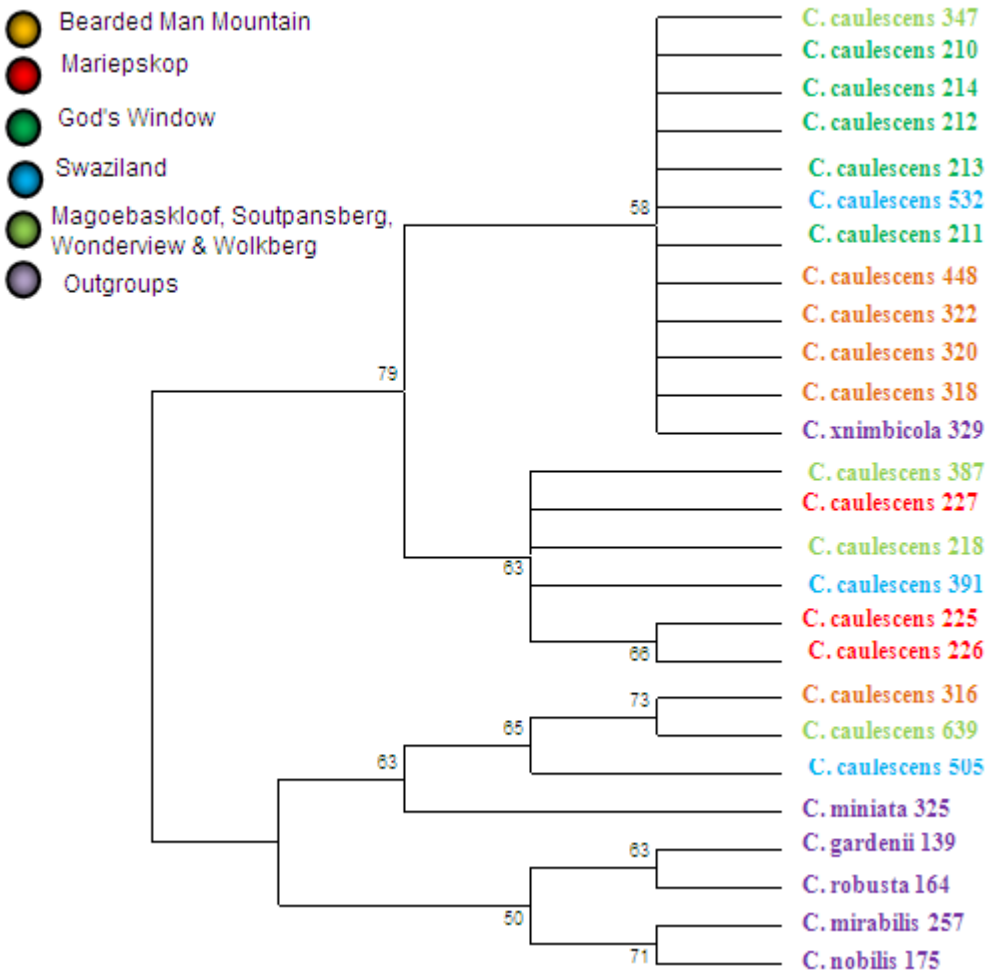


Figure D.3: Cladogram conducted through **Minimum Evolution** method for the *matK*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

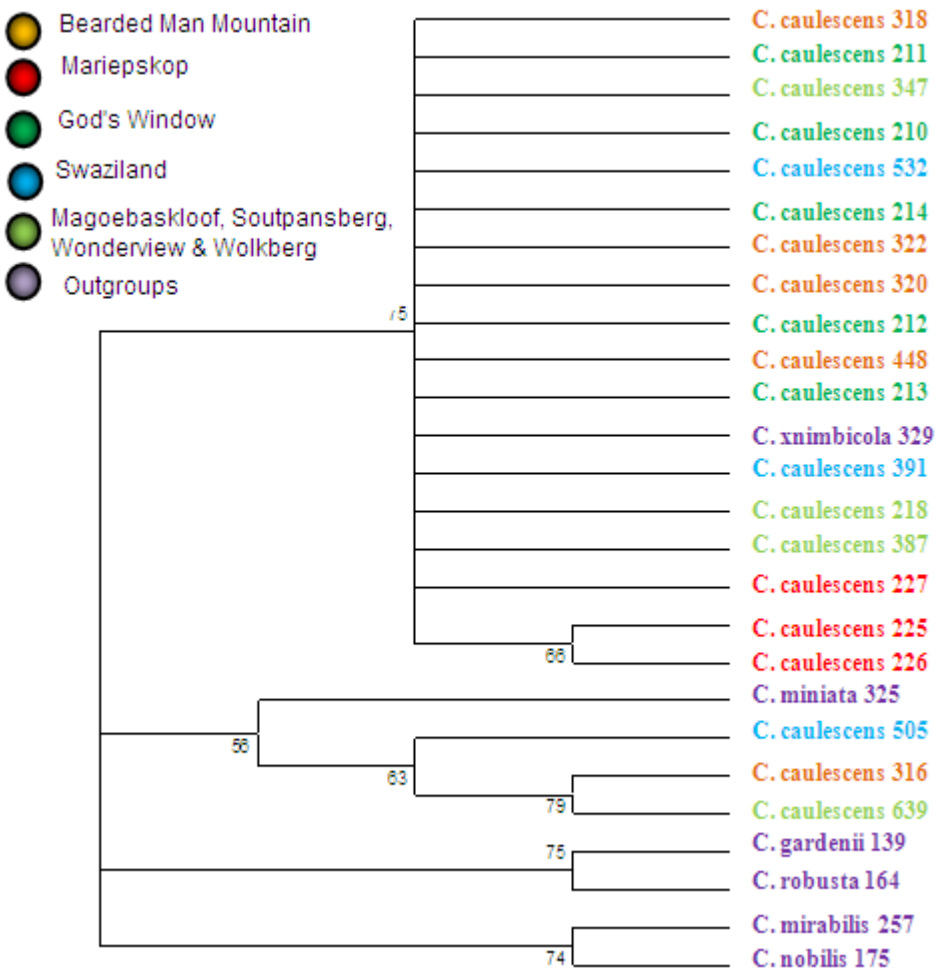


Figure D.4: Cladogram conducted through **Maximum Parsimony** method for the *matK*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

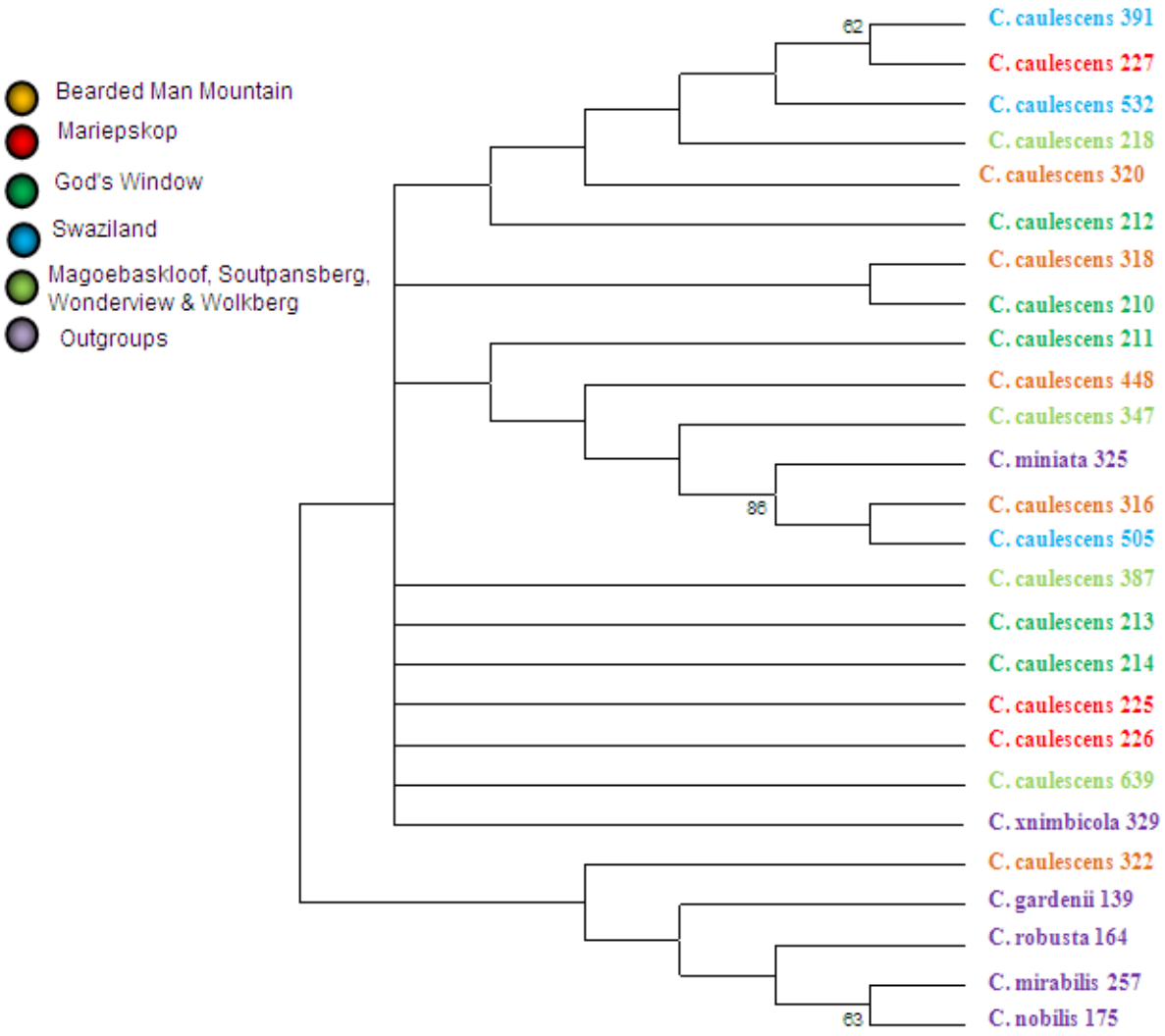


Figure D.5: Cladogram conducted through **Minimum Evolution** method for the *rpoB*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

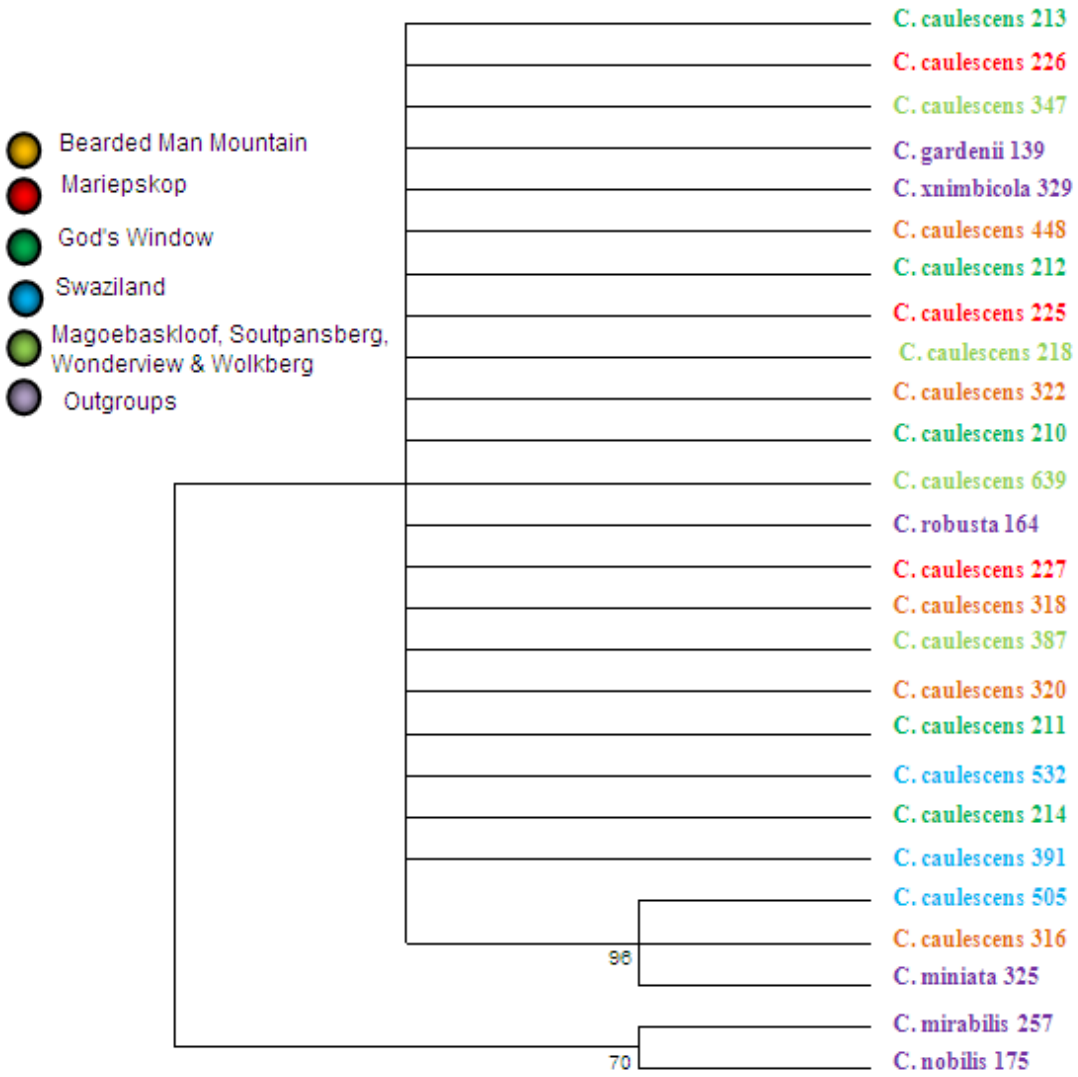


Figure D.6: Cladogram conducted through **Maximum Parsimony** method for the *rpoB*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*;

448 – *Spies 8701*; **211** – *Spies 8480*; **212** – *Spies 8481*; **213** – *Spies 8482*;

214 – *Spies 8483*; **210** – *Spies 8479*; **639** – *Spies 8892*; **225** – *Spies 8494*;

226 – *Spies 8495*; **227** – *Spies 8496*; **387** – *Spies 8640*; **391** – *Spies 8644*;

532 – *Spies 8757*; **505** – *Spies 8757*; **218** – *Spies 8487*; **347** – *Spies 8596*;

139 – *Spies 8418*; **325** – *Spies 8574*; **257** – *Spies 8267*; **175** – *Spies 8254*

164 – *Spies 8440*; **329** – *Spies 8578*

APPENDIX E: Mean genetic distances between groups for the different gene regions

Table E.1: Mean genetic distance **between** groups based on *matK* gene region.

(1 – Bearded Man Mountain; 2 – Swaziland; 3 – Soutpansberg; 4 – God's Window; 5 – Mariepskop; 6 – Magoebaskloof; 7 – Wonderview; 8 – Wolkberg)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Bearded Man Mountain								
Swaziland	0.005							
Soutpansberg	0.004	0.003						
God's Window	0.002	0.003	0.002					
Mariepskop	0.005	0.004	0.001	0.003				
Magoebaskloof	0.010	0.008	0.008	0.010	0.009			
Wonderview	0.002	0.003	0.002	0.000	0.003	0.010		
Wolkberg	0.004	0.003	0.000	0.002	0.001	0.008	0.002	
Outgroups	0.007	0.007	0.007	0.006	0.009	0.010	0.006	0.007

Table E.2: Mean genetic distance **between** groups based on *rpoB* gene region.

(1 – Bearded Man Mountain; 2 – Swaziland; 3 – Soutpansberg; 4 – God's Window; 5 – Mariepskop; 6 – Magoebaskloof; 7 – Wonderview; 8 – Wolkberg)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Bearded Man Mountain								
Swaziland	0.005							
Soutpansberg	0.001	0.005						
God's Window	0.001	0.005	0.000					
Mariepskop	0.005	0.008	0.004	0.004				
Magoebaskloof	0.001	0.005	0.000	0.000	0.004			
Wonderview	0.001	0.005	0.000	0.000	0.004	0.000		
Wolkberg	0.001	0.005	0.000	0.000	0.004	0.000	0.000	
Outgroups	0.003	0.006	0.002	0.002	0.006	0.002	0.002	0.002

Table E.3: Mean genetic distance **between** groups based on the Combinded set gene region.

(1 – Bearded Man Mountain; 2 – Swaziland; 3 – Soutpansberg; 4 – God's Window; 5 – Mariepskop; 6 – Magoebaskloof; 7 – Wonderview; 8 – Wolkberg)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Bearded Man Mountain								
Swaziland	0.005							
Soutpansberg	0.003	0.003						
God's Window	0.002	0.003	0.002					
Mariepskop	0.006	0.006	0.003	0.005				
Magoebaskloof	0.006	0.006	0.004	0.005	0.007			
Wonderview	0.003	0.004	0.002	0.001	0.006	0.006		
Wolkberg	0.003	0.003	0.001	0.002	0.004	0.004	0.002	
Outgroups	0.004	0.005	0.004	0.003	0.007	0.005	0.004	0.004

APPENDIX F: Flow diagrams illustrating the mean distances between populations *C. caulescens*.

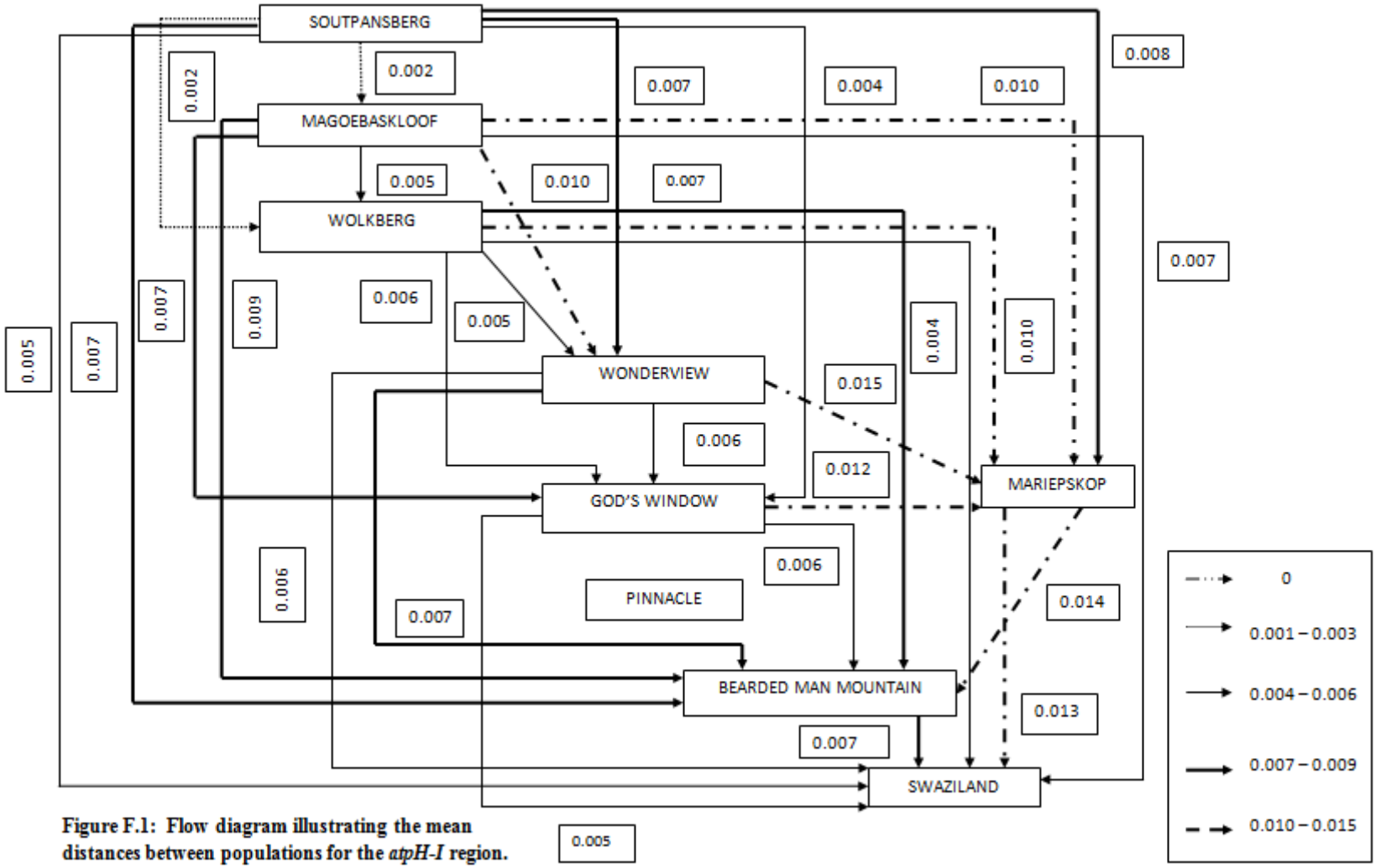


Figure F.1: Flow diagram illustrating the mean distances between populations for the *atpH-I* region.

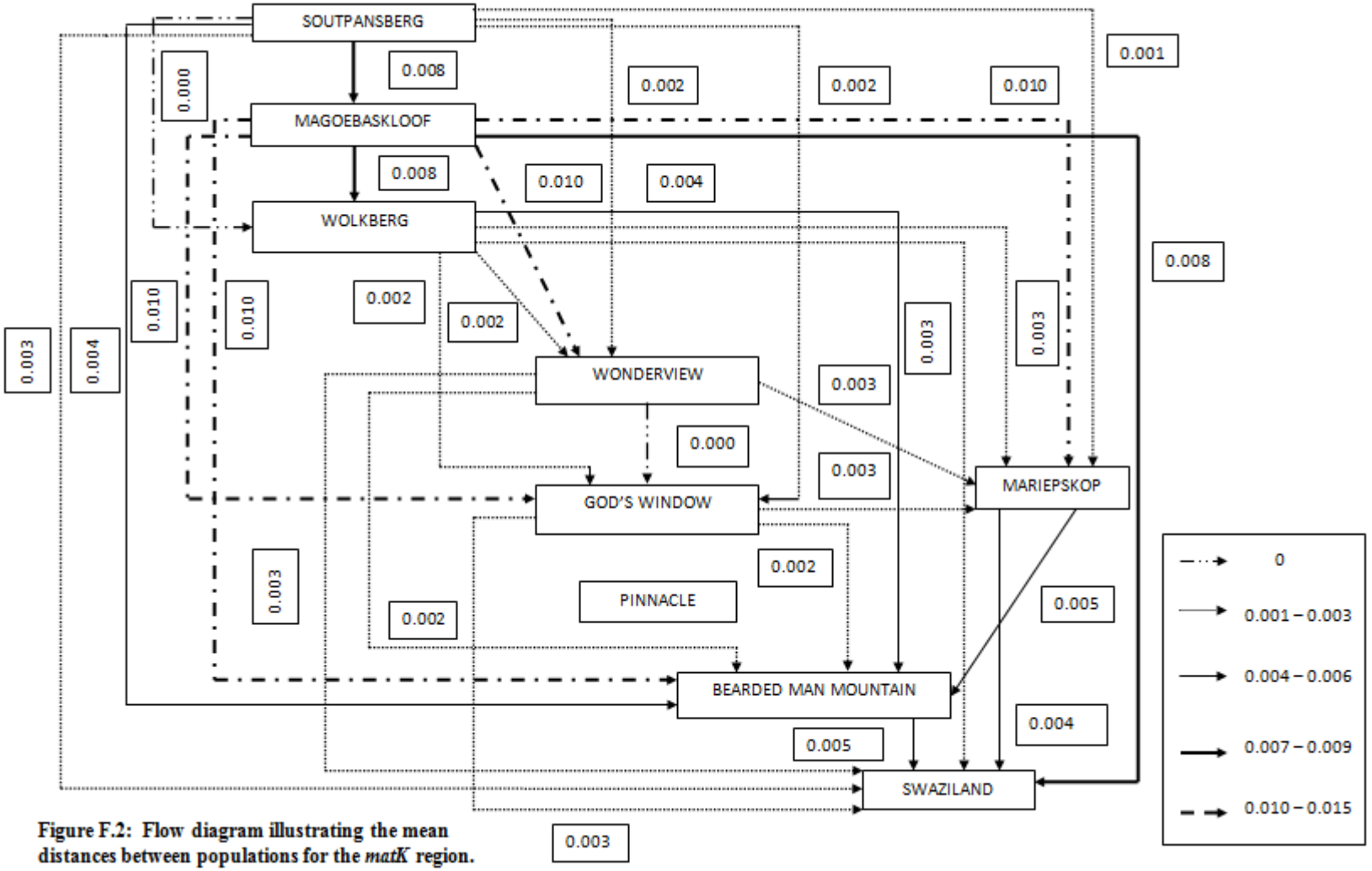


Figure F.2: Flow diagram illustrating the mean distances between populations for the *matK* region.

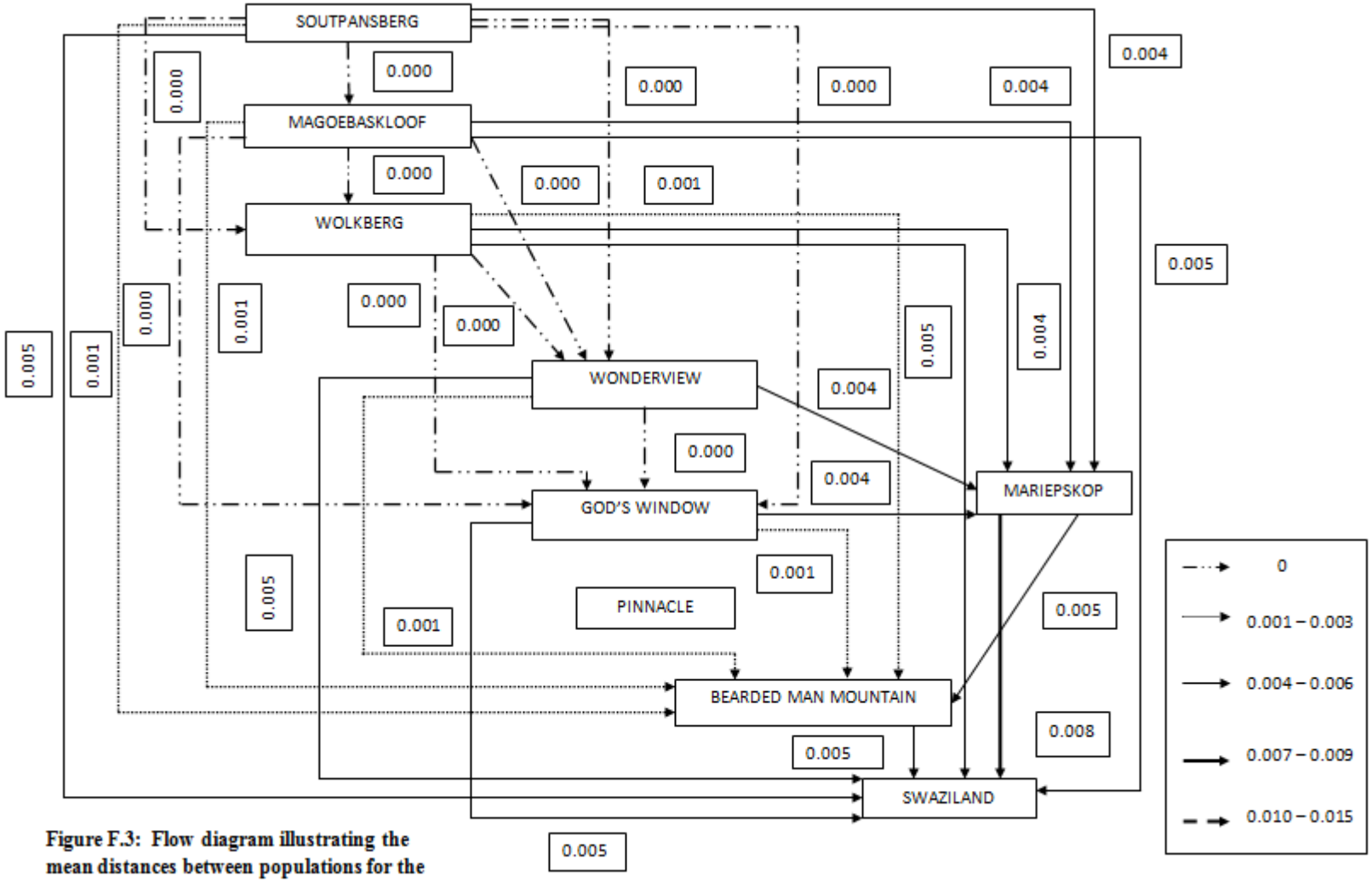


Figure F.3: Flow diagram illustrating the mean distances between populations for the *rpoB* region.

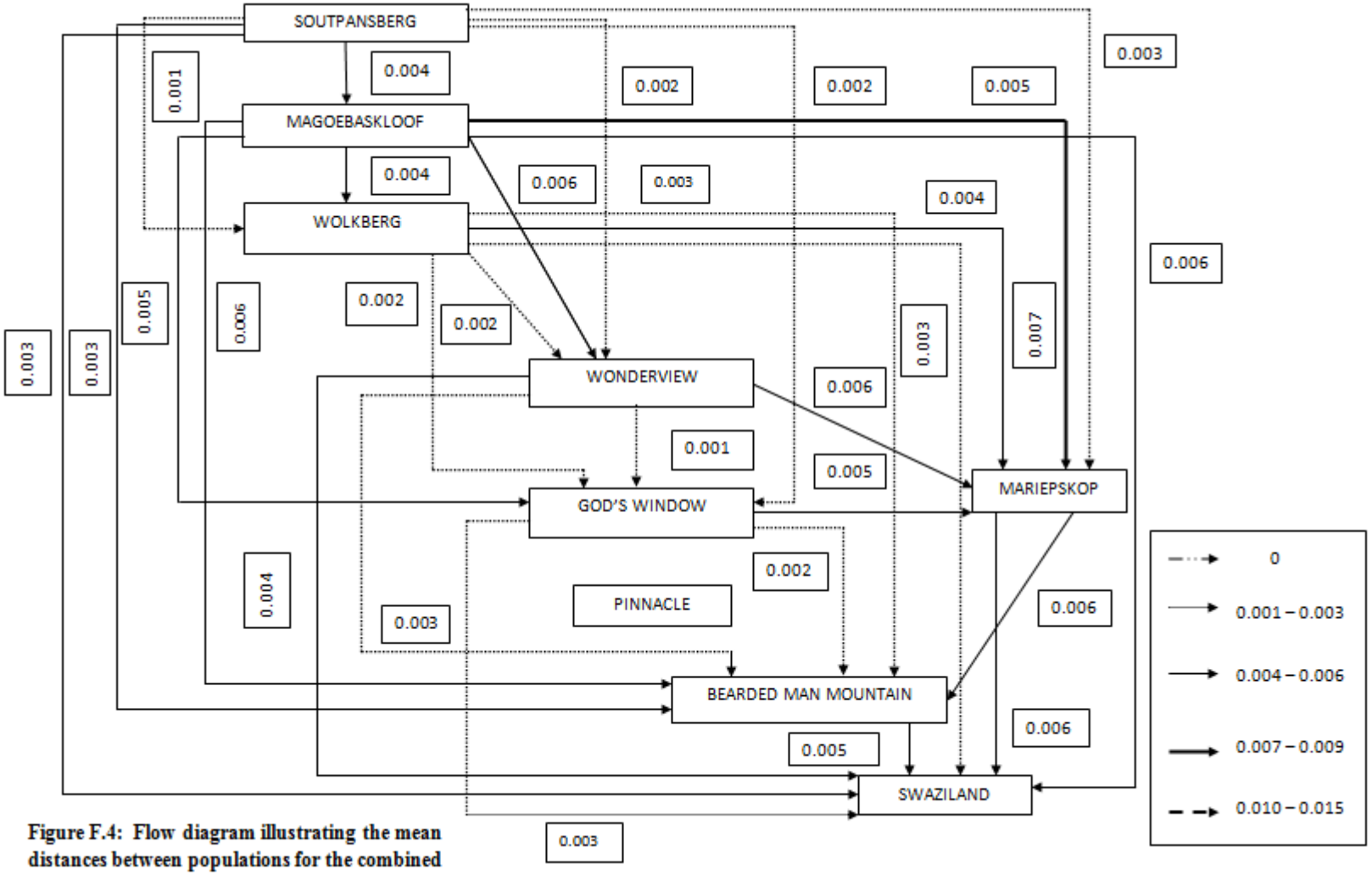


Figure F.4: Flow diagram illustrating the mean distances between populations for the combined dataset region.

APPENDIX G: Flow diagrams of mean distances between locations of *C. caulescens* specimens included in this study

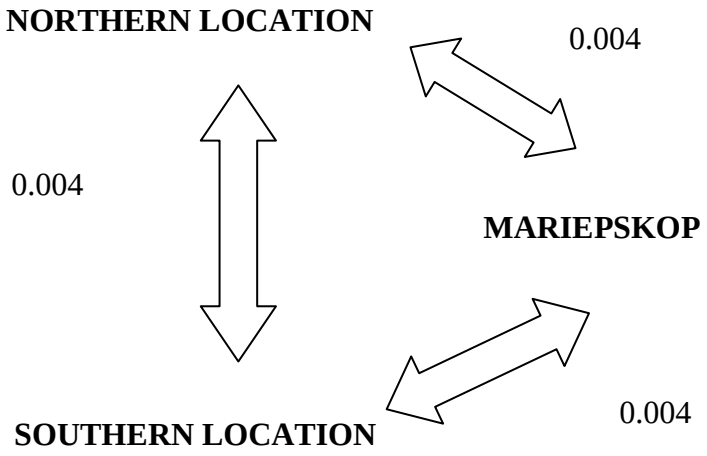


Figure G.1: Flow diagram of mean distances between locations for ***matK*** region.

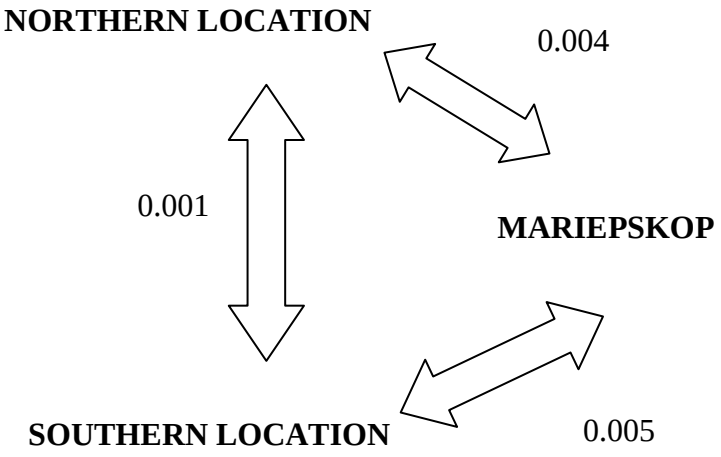


Figure G.2: Flow diagram of mean distances between locations for ***rpoB*** region.

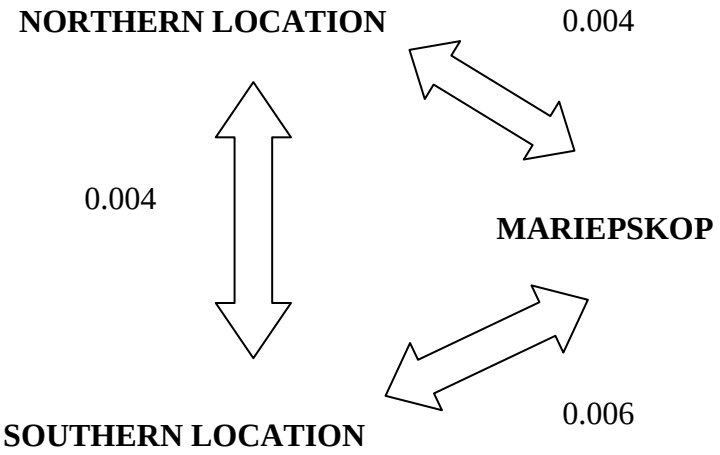


Figure G.3: Flow diagram of mean distances between locations for **Combined dataset.**

APPENDIX H: Nanodrop readings of the Clivias from Bearded Man Mountain specimens used in this study.

Voucher Number	Sample ID	ng/µl	A260	A280	260/280	260/230
Spies 8565	C. caulescens 316	200.76	4.015	2.026	1.98	1.81
Spies 8567	C. caulescens 318	75.06	1.501	0.751	2.00	1.73
Spies 8569	C. caulescens 320	262.98	5.260	4.101	1.28	0.54
Spies 8571	C. caulescens 322	188.57	3.771	2.647	1.42	0.71
Spies 8701	C. caulescens 448	56.31	1.126	0.484	2.33	-7.63
Spies 8674	C. miniata 421	381.76	7.635	4.915	1.55	0.70
Spies 8656	C. miniata 403	170.48	3.410	1.750	1.95	1.71
Spies 8558	C. miniata 309	103.04	2.061	1.199	1.72	0.74
Spies 8572	C. miniata 323	135.17	2.703	1.369	1.97	1.81
Spies 8721	C. miniata 469	501.74	10.035	5.018	2.00	1.79
Spies 8641	C. xnimbicola 388	492.98	9.860	5.255	1.88	1.51
Spies8654	C. xnimbicola 401	162.50	3.250	1.640	1.98	1.85
Spies 8655	C. xnimbicola 402	129.85	2.597	1.604	1.62	1.22
Spies 8648	C. xnimbicola 395	100.17	2.003	0.980	2.04	1.85
Spies 8578	C. xnimbicola 329	140.57	2.890	1.622	1.78	1.52

APPENDIX I: Gel documents of amplification products from the Bearded Man Mountain. (GeneRuler DNA Ladder Mix - three different reference bands (3000, 1000 and 500bp)).

L NA 318 322 395 NA 388 401 402 316 NA

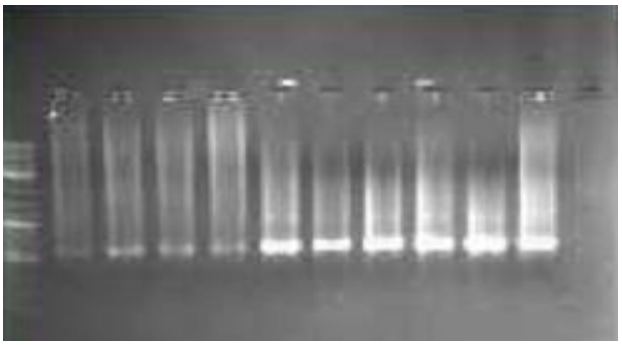


Fig.I.1: The photograph illustrates the fragments obtained after pre-sequencing ***atpH-I***. The numbers indicate the specimens listed in specimen list. L - ladder

L 320 421 403 402 309 448 323 469

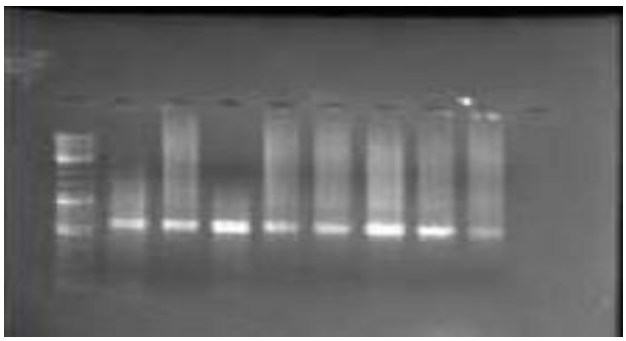


Fig.I.2: The photograph illustrates the fragments obtained after pre-sequencing ***atpH-I***. The numbers indicate the specimens listed in specimen list. L – ladder.

L 318 320 316 322 NA 323 309 388 401 402 NA

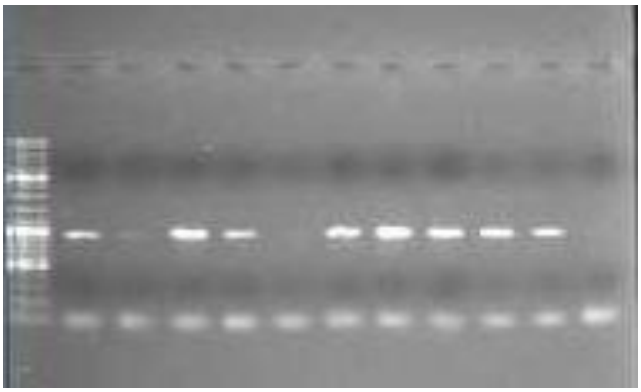


Fig.I.3: The photograph illustrates the fragments obtained after pre-sequencing ***matK***. The numbers indicate the specimens listed in specimen list. L - ladder

L 322 448 421 403 469 395 469 NA NA NA

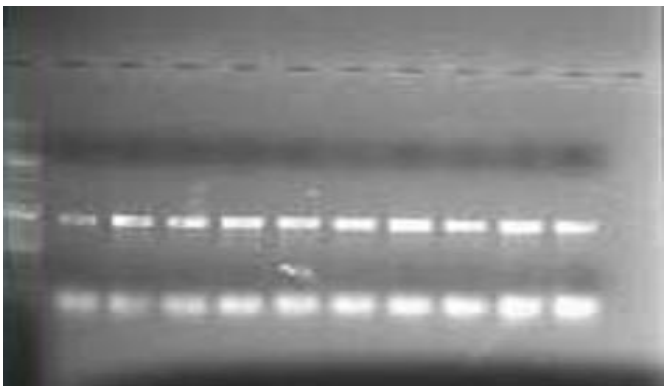


Fig.I.4: The photograph illustrates the fragments obtained after pre-sequencing ***matK***. The numbers indicate the specimens listed in specimen list. L – ladder.

*NA – Not applicable

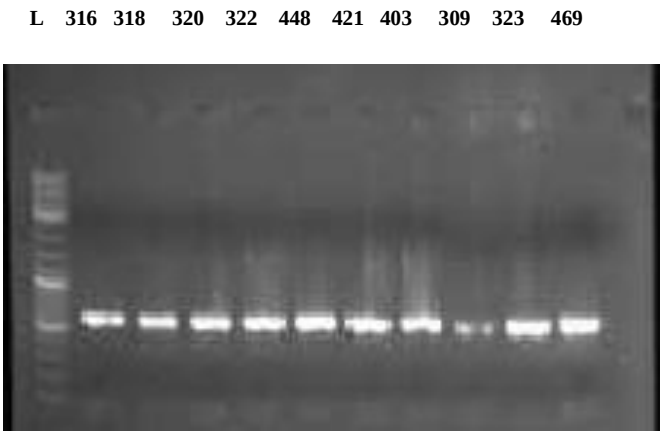


Fig.I.5: The photograph illustrates the fragments obtained after pre-sequencing ***rpoB***. The numbers indicate the specimens listed in specimen list. L - ladder

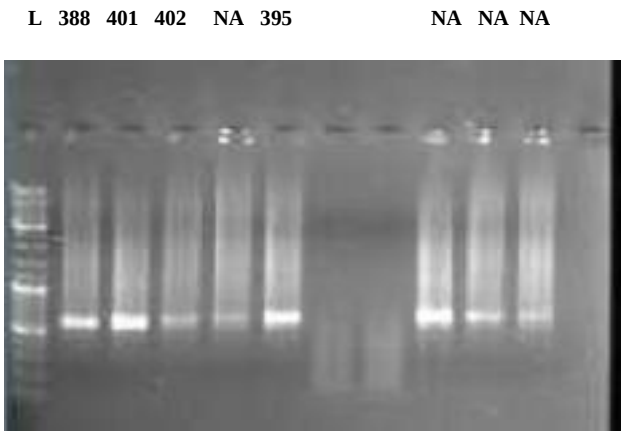


Fig.I.6: The photograph illustrates the fragments obtained after pre-sequencing ***rpoB***. The numbers indicate the specimens listed in specimen list. L – ladder.

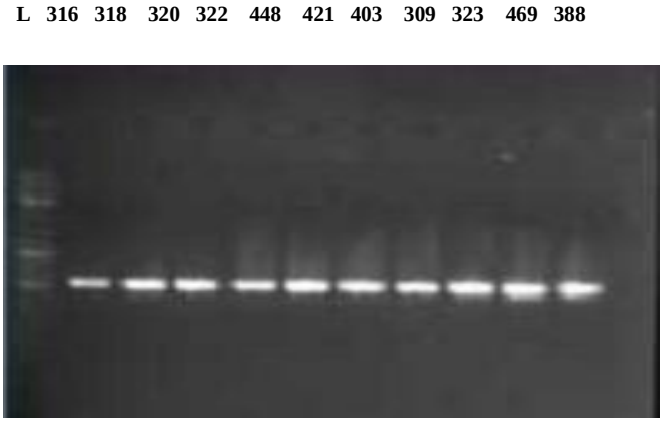


Fig.I.7: The photograph illustrates the fragments obtained after pre-sequencing ***rpoC***. The numbers indicate the specimens listed in specimen list. L - ladder

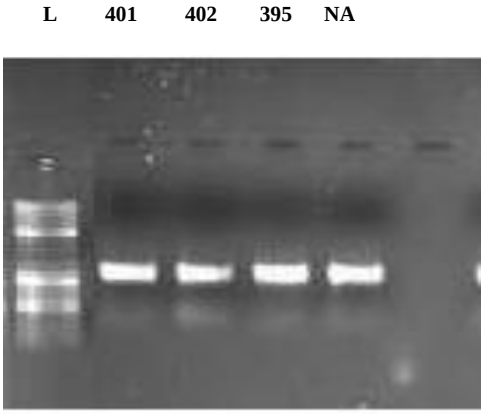


Fig.I.8: The photograph illustrates fragments obtained after pre-sequencing ***rpoC***. The numbers indicate the specimens listed in specimen list. L – ladder.

*NA – Not applicable

APPENDIX J: Illustration of Geneious output for the different gene regions used in this study.

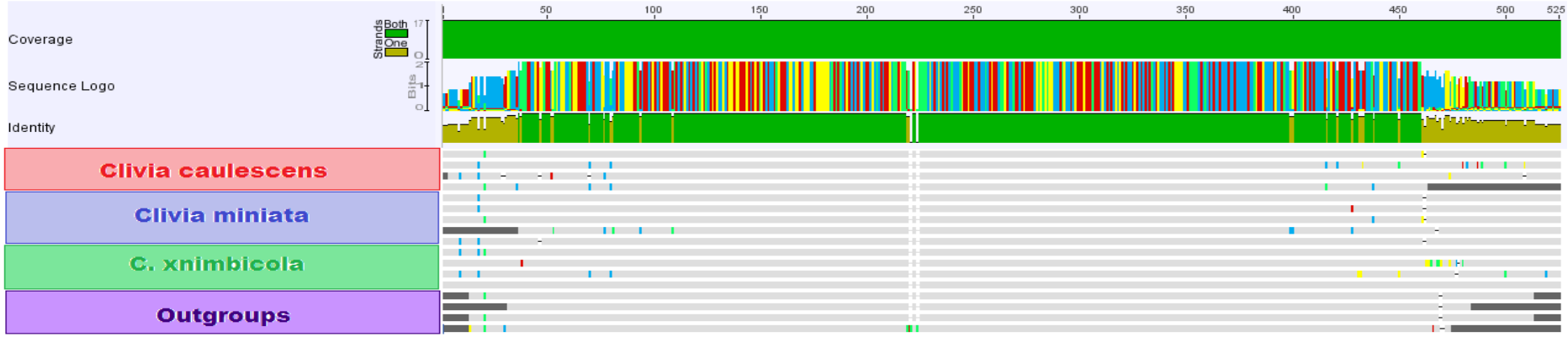


Figure J.1: Illustration of the ***atpH-I*** regions' GENEIOUS output with conservative sites unmarked.

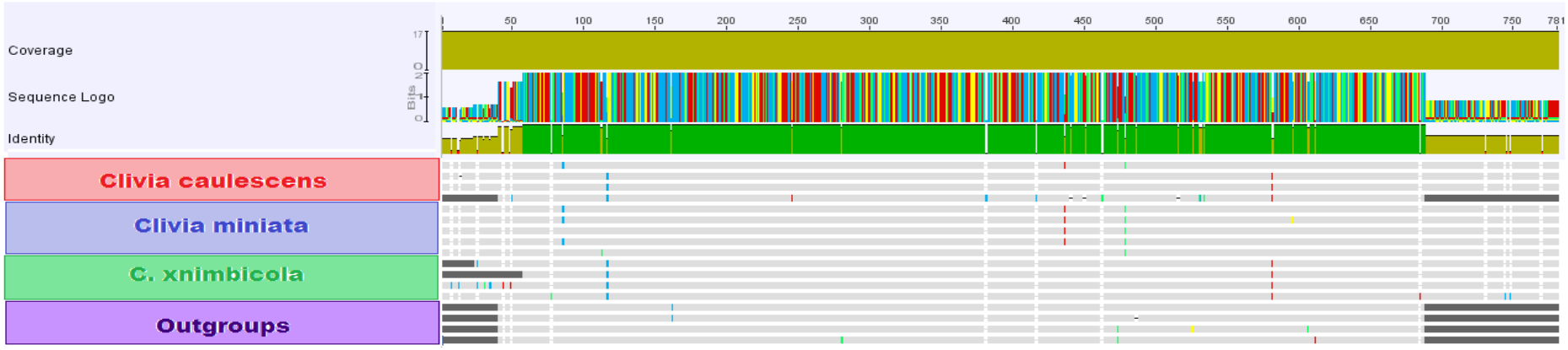


Figure J.2: Illustration of the ***matK*** regions' GENEIOUS output with conservative sites unmarked.

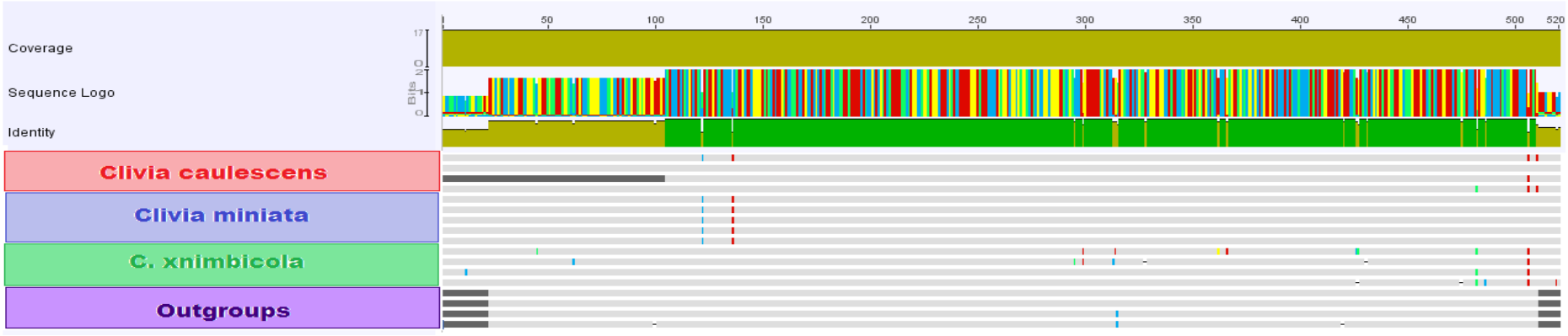


Figure J.3: Illustration of the ***rpoB*** regions' GENEIOUS output with conservative sites unmarked.

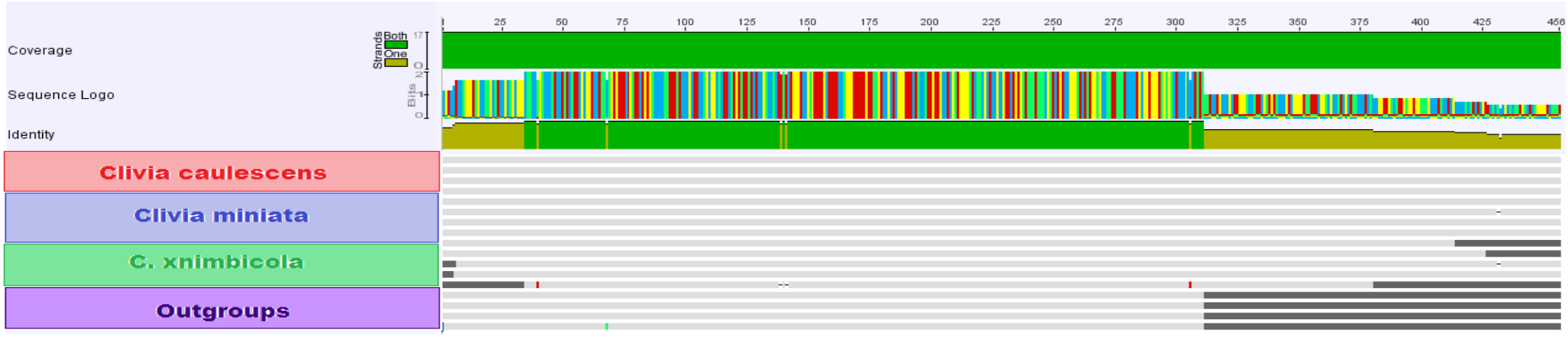


Figure J.4: Illustration of the ***rpoC*** regions' GENEIOUS output with conservative sites unmarked.

APPENDIX K: Illustration of SNPs obtained through the different gene regions (*atpH-I*, *matK* and *rpoB*)

K.1: Table of the SNPs obtained through gene region *atpH-I*

atpH-I	1	2	3	4	5	6	7	8	9	10
316				C						
318			T							T
320	T		T					A		
322				C		T				T
421			T							
403			T							
309				C						
323									C	
469	T		T							
388	T		T	C						
401							A			
402										T
395	T		T							
139				C						
164										
257				C						
175		G		C	T					

atpH-I	11	12	13	14	15	16	17	18	19	20
316										
318		T								
320	T									
322		T								
421										
403										
309										
323			C	T	C					T
469										
388										
401										
402		T								
395										
139										
164										
257										
175						C	A	C	C	

atpH-I	21	22	23	24	25	26	27	28	29	30
316										G
318		T	T				G		C	
320										
322		C						T		
421										
403				A						
309								T		G
323	T			T						
469										
388										
401										
402					G	G			G	
395										
139										
164										
257										
175										

atpH-I	31	32	33	34	35	36	37	38	39	40
316										
318									A	T
320							G			
322										
421										
403										
309										
323										
469										
388										
401	G	G	C	C	C	G	G	T	C	
402										
395										
139										
164										
257										
175										

atpH-I	41	42	43	44	45	**TOTAL SNPs**
316						**4, 30**
318	A	C	C	G		**3, 10, 12, 22, 23, 27, 29, 39, 40, 41, 42, 43,44**
320						**1, 3, 8, 11, 37**
322						**4, 6, 10, 12, 22, 28**
421						**3**
403						**3, 24**
309						**4, 28, 30**
323						**9, 13, 14, 15, 20, 21, 24**
469						**1, 3**
388						**1, 3, 4**
401			C		T	**7, 31, 32, 33, 34, 35, 36, 37, 38, 39,45**
402						**10, 12, 25, 26, 29**
395						**1, 3**
139						**4**
164						
257						**4**
175						**2, 4, 5, 16, 17, 18, 19**

K.2: Table of the SNPs obtained through gene region *matK*.

matK	1	2	3	4	5	6	7	8	9	10
316										
318										
320										
322									T	
421										
403										
309										
323										
469										
388			T							
401										
402	T	T	T	C	T	A	A			
395								A		C
139										
164										
257										
175										

matK	11	12	13	14	15	16	17	18	19	20
316	T								A	
318			T							
320			T							
322			T		A		T	T		C
421	T								A	
403	T								A	
309									A	
323	T								A	
469		C								
388			T							
401			T							
402			T							
395										
139				T						
164				T						
257										
175						C				

matK	21	22	23	24	25	26	27	28	29	30
316		C								
318							A			
320							A			
322				C	T	C	A			
421		C								
403		C						G		
309		C								
323		C								
469										
388							A			
401							A			
402							A			
395							A			
139										
164										
257	C		G						C	
175	C									A

matK	31	32	33	TOTAL SNPs
316				**11, 19, 22**
318				**13, 27**
320				**13, 27**
322				**9, 13, 15, 17, 18, 20, 24, 25, 26, 27**
421				**11, 19, 22**
403				**11, 19, 22, 28**
309				**19, 22**
323				**11, 19, 22**
469				**12**
388				**3, 13, 27**
401				**13, 27**
402				**1, 2, 3, 4, 5, 6, 7, 13, 27**
395	A	T	T	**8, 10, 27, 31, 32, 33**
139				**14**
164				**14**
257				**21, 23, 29**
175				**16, 21, 30**

K.3: Table of the SNPs obtained through gene region *rpoB*

rpoB	1	2	3	4	5	6	7	8	9	10
316				T	A					
318										
320										
322										
421				T	A					
403				T	A					
309				T	A					
323				T	A					
469				T	A					
388		C					A		A	
401			T			C	A	T		
402	T									
395										
139										
164										
257										T
175										T

rpoB	11	12	13	14	15	16	17	18	19	TOTAL SNPs
316							A	A		4, 5, 17, 18
318										
320							A			17
322					C		A	A		15, 17, 18
421										4, 5
403										4, 6
309										4, 7
323										4, 8
469										4, 9
388	G	A	T	C	C		A			2, 7, 9, 11, 12, 13, 14, 15, 17
401							A			3, 6, 7, 8, 17
402					C		A			1, 15, 17
395					C	T	A		A	15, 16, 17, 19
139										
164										
257										10
175										10

APPENDIX L: Cladograms obtained from the other gene regions (*atpH-I*; *matK*; *rpoB* and *rpoC1*) by making use of Maximum Parsimony and Minimum Evolution methods for the Bearded Man Mountain specimens.

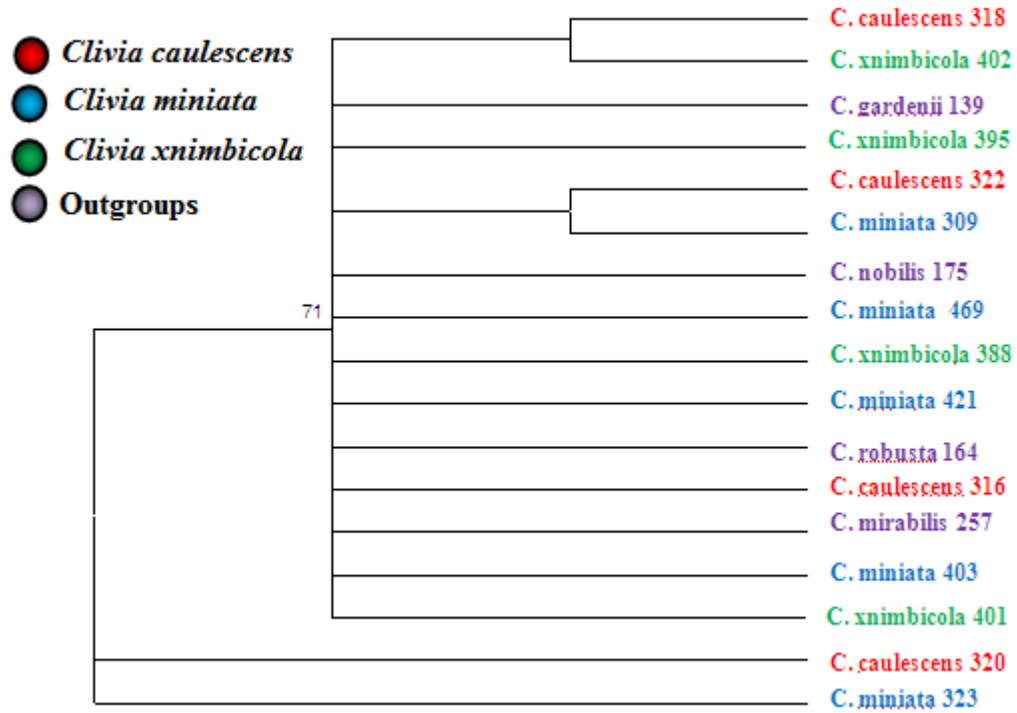


Figure L.1: Cladogram conducted through **Maximum Parsimony** method for *atpH-I*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

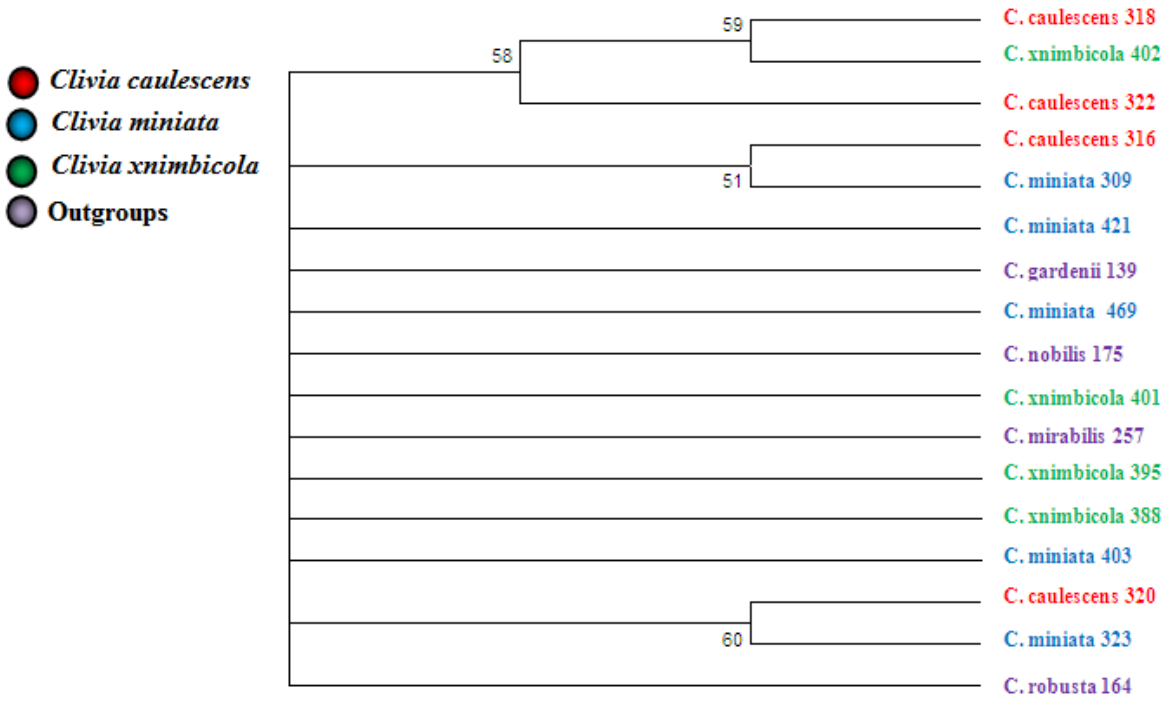


Figure L.2: Cladogram conducted through **Minimum Evolution** method for *atpH-I*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

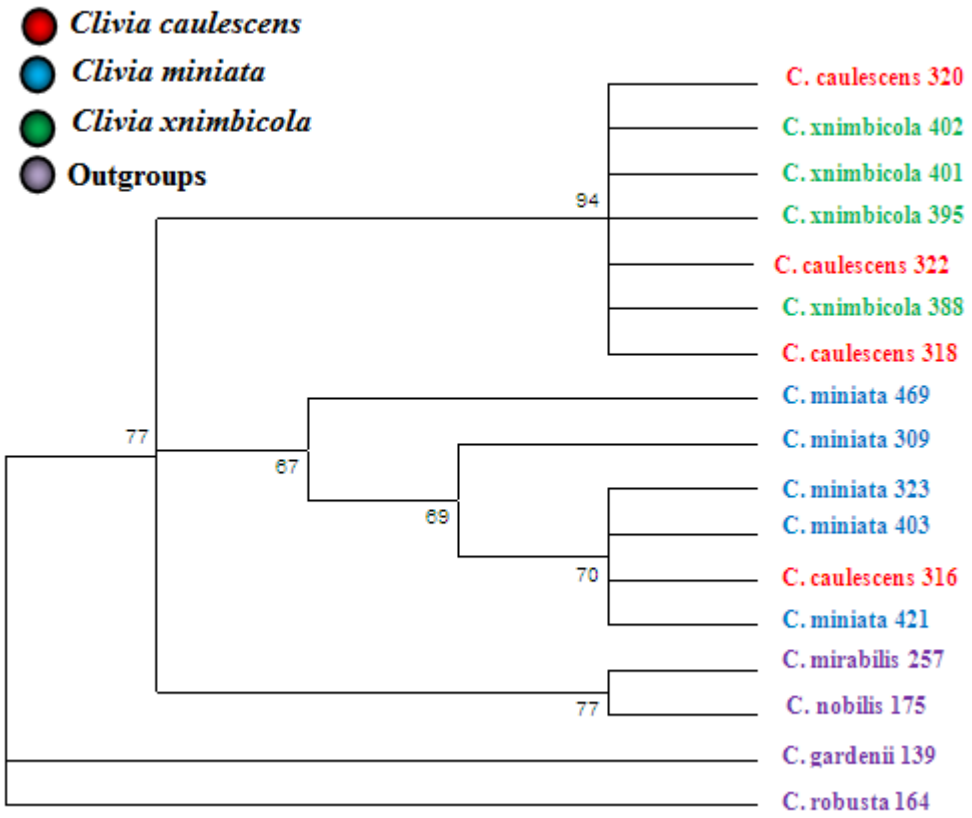


Figure L.3: Cladogram conducted through **Maximum Parsimony** method for *matK*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

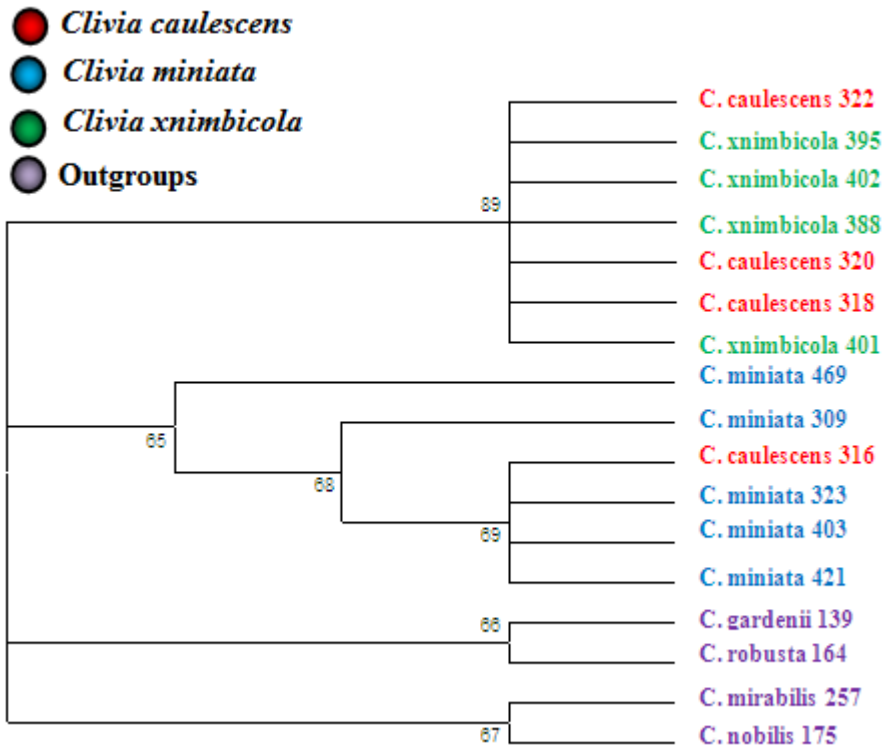


Figure L.4: Cladogram conducted through **Minimum Evolution** method for *matK*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

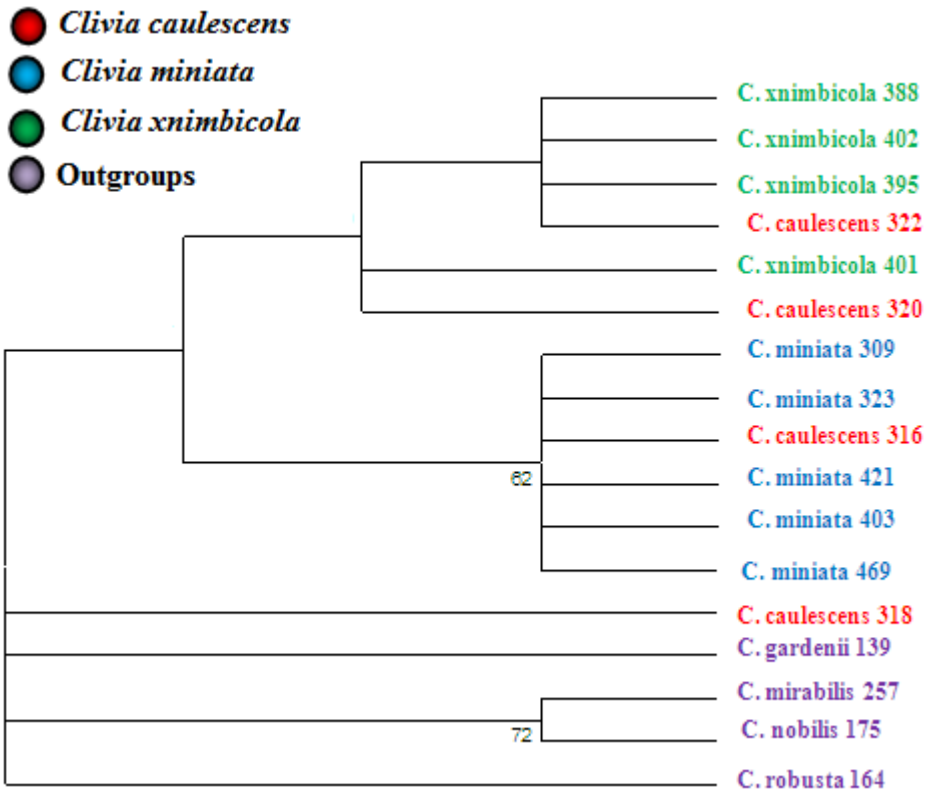


Figure L.5: Cladogram conducted through **Maximum Parsimony** method for *rpoB*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

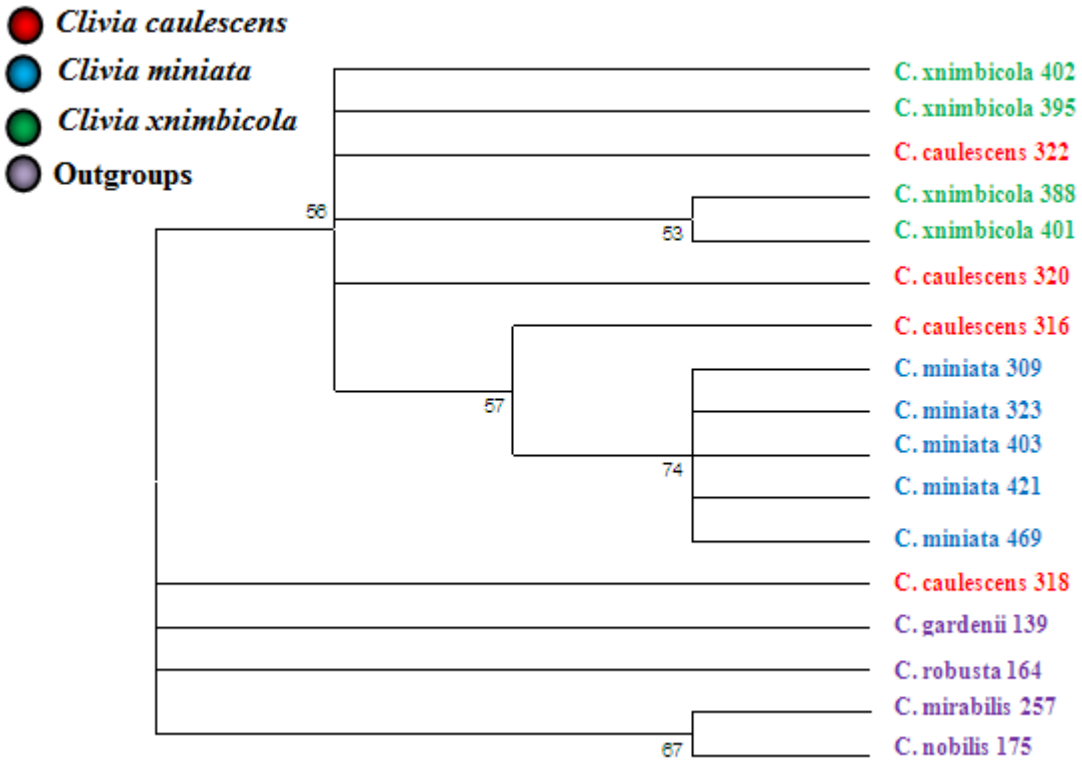


Figure L.6: Cladogram conducted through **Minimum Evolution** method for *rpoB*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

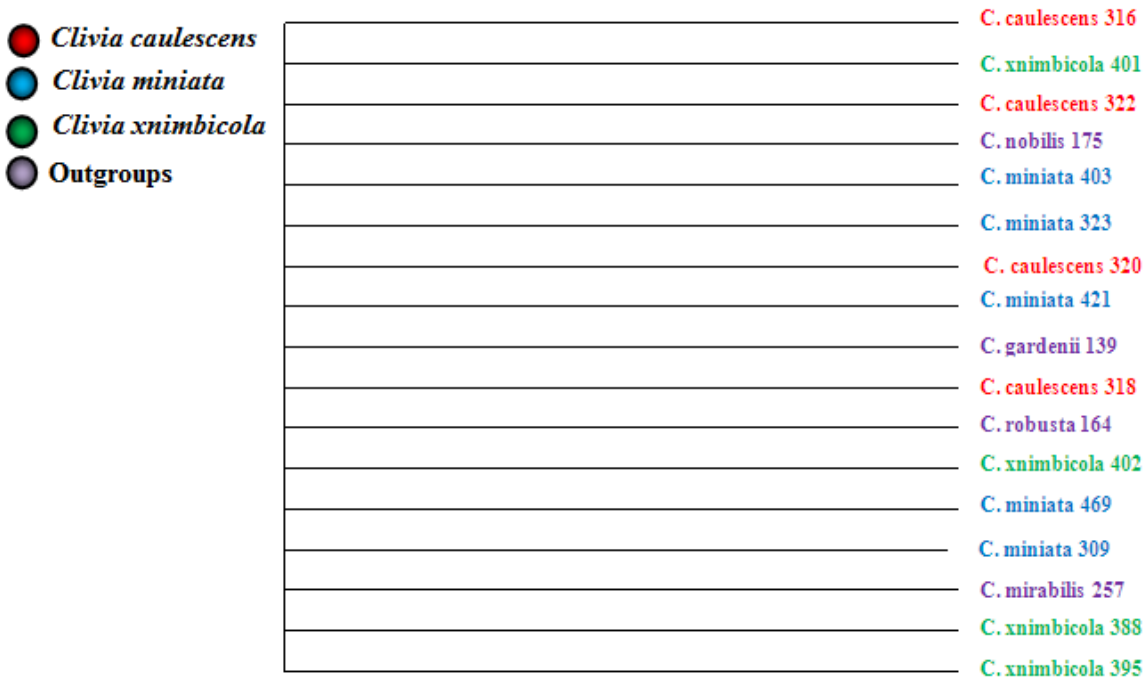


Figure L.7: Cladogram conducted through **Maximum Parsimony** and **Minimum Evolution** method for *rpoC*.

316 – *Spies 8565*; **318** – *Spies 8567*; **320** – *Spies 8569*; **322** – *Spies 8571*; **421** – *Spies 8674*; **403** – *Spies 8656*; **309** – *Spies 8558*; **323** – *Spies 8572*; **469** – *Spies 8721*; **388** – *Spies 8641*; **401** – *Spies 8654*; **402** - *Spies 8655*; **395** – *Spies 8648*; **164** – *Spies 8440*; **139** – *Spies 8418*; **175** – *Spies 8254*; **257** – *Spies 8267*

APPENDIX M: Mean genetic distances between species at Bearded Man Mountain

Table M.1: Mean genetic distances between groups for ***matK*** region.

(1 – *C. caulescens*; 2 – *C. miniata*; 3 – *C. xnimbicola*)

	1	**2**	**3**
C. caulescens			
C. miniata	**0.008**		
C. xnimbicola	**0.004**	**0.008**	
Outgroups	**0.008**	**0.008**	**0.006**

Table M.2: Mean genetic distances between groups for ***rpoB*** region.

(1 – *C. caulescens*; 2 – *C. miniata*; 3 – *C. xnimbicola*)

	1	**2**	**3**
C. caulescens			
C. miniata	**0.008**		
C. xnimbicola	**0.010**	**0.015**	
Outgroups	**0.006**	**0.006**	**0.011**

Table M.3: Mean genetic distances between groups for ***rpoC*** region.

(1 – *C. caulescens*; 2 – *C. miniata*; 3 – *C. xnimbicola*)

	1	**2**	**3**
C. caulescens			
C. miniata	**0.000**		
C. xnimbicola	**0.002**	**0.002**	
Outgroups	**0.001**	**0.001**	**0.003**

Table M.4: Mean genetic distances between groups for **combined dataset.**

(1 – *C. caulescens*; 2 – *C. miniata*; 3 – *C. xnimbicola*)

	1	**2**	**3**
C. caulescens			
C. miniata	**0.007**		
C. xnimbicola	**0.006**	**0.009**	
Outgroups	**0.006**	**0.006**	**0.006**

APPENDIX N: Flow diagram of the mean distances between the different species at the Bearded Man Mountain.

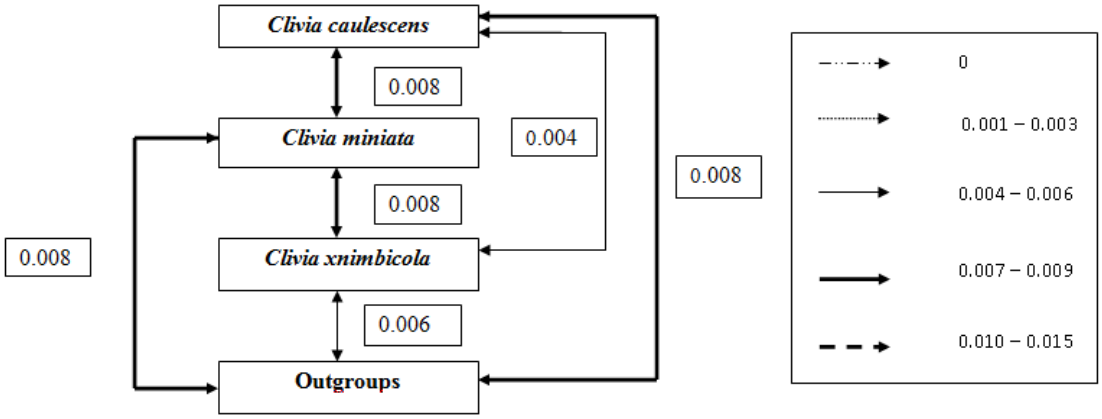


Figure N.1: Flow diagram to illustrate the mean distances between the different species at the Bearded Man Mountain for *matK* region.

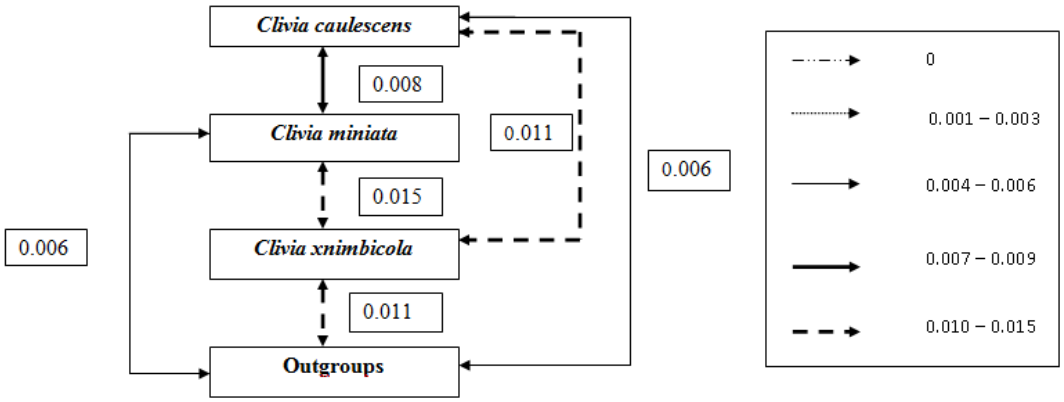


Figure N.2: Flow diagram to illustrate the mean distances between the different species at the Bearded Man Mountain for *rpoB* region.

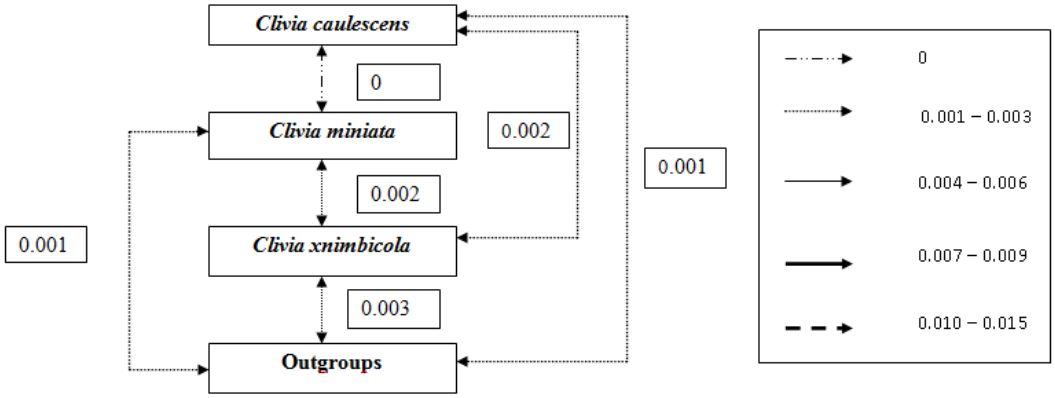


Figure N.3: Flow diagram to illustrate the mean distances between the different species at the Bearded Man Mountain for *rpoC* region.

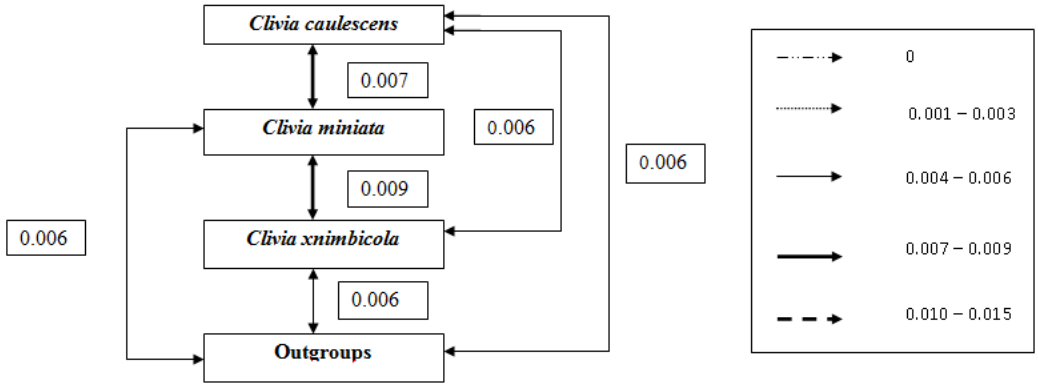


Figure N.4: Flow diagram to illustrate the mean distances between the different species at the Bearded Man Mountain for the combined dataset.

APPENDIX O: Electropherograms

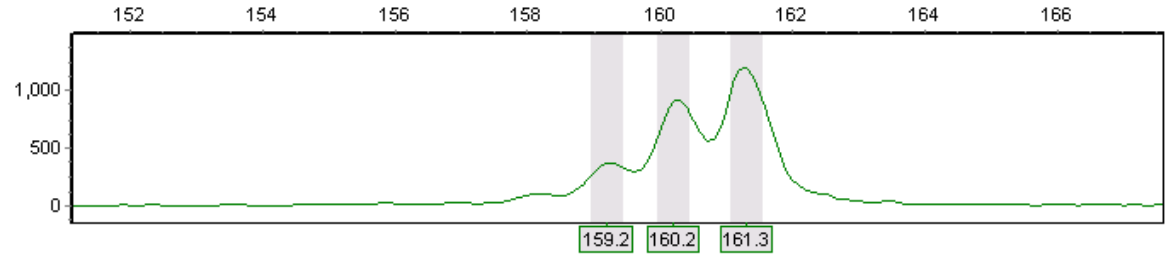
Fragment analysis

SoftGenetics — **Allele Report** — 6/3/2009 9:31:21 AM

GeneMarker V1.60 — Page 1

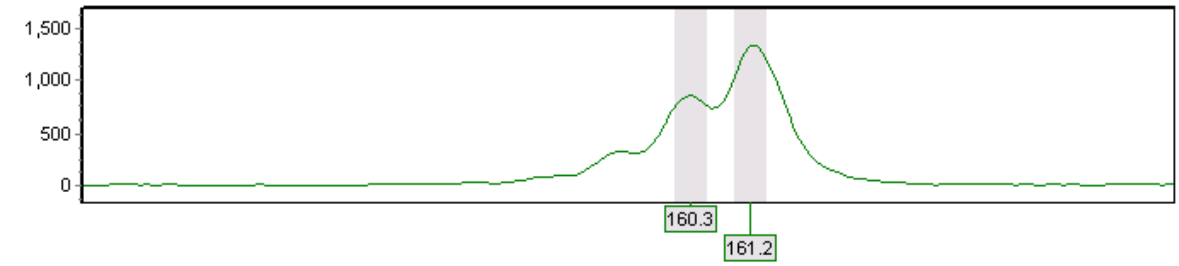
Sample 1:

Dye: Green - 19 peaks - A_2009-06-01.fsa



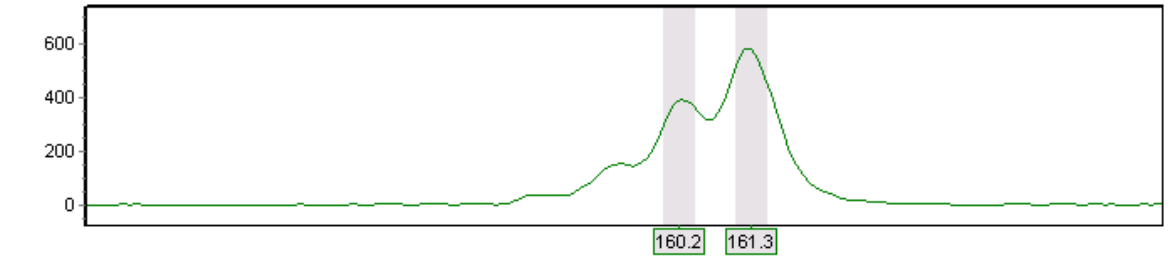
Sample 2:

Dye: Green - 14 peaks - B_2009-06-01.fsa



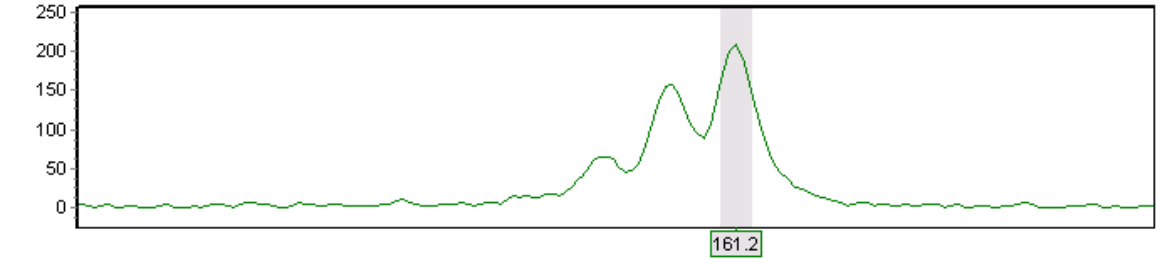
Sample 3:

Dye: Green - 8 peaks - C_2009-06-01.fsa



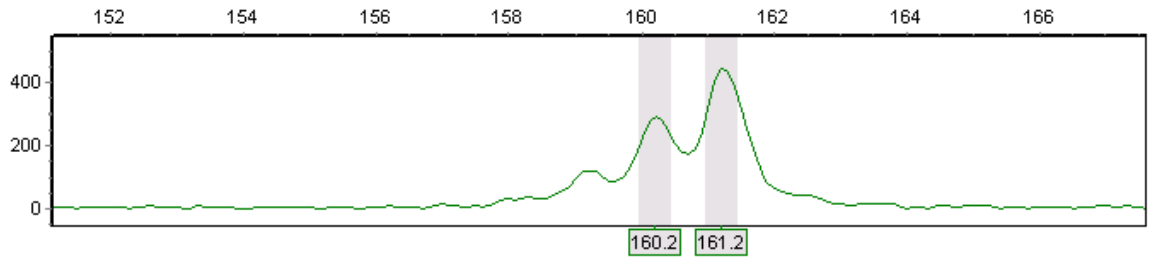
Sample 4:

Dye: Green - 10 peaks - D_2009-06-01.fsa



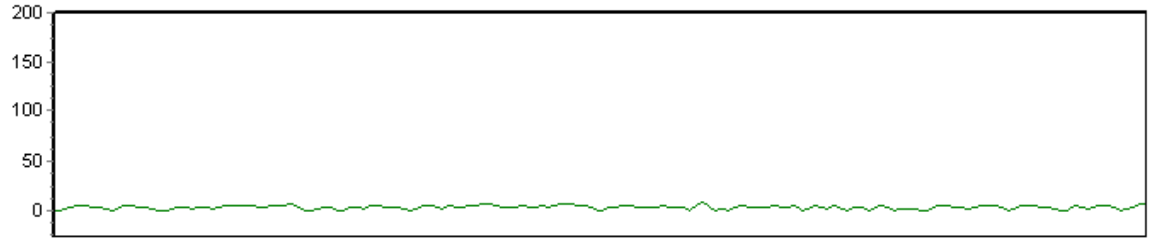
Sample 5:

Dye: Green - 11 peaks - E_2009-06-01.fsa



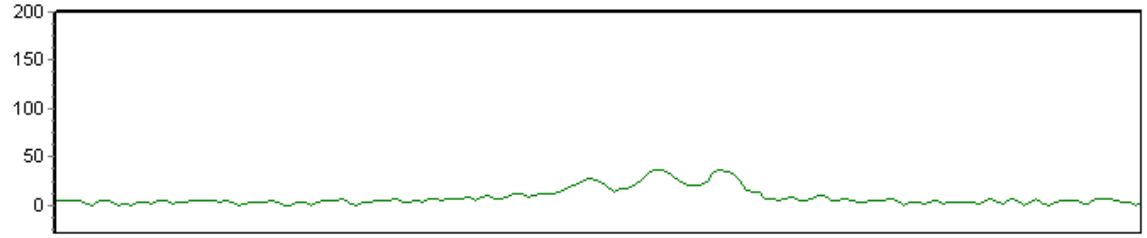
Sample 6:

Dye: Green - 3 peaks - F_2009-06-01.fsa



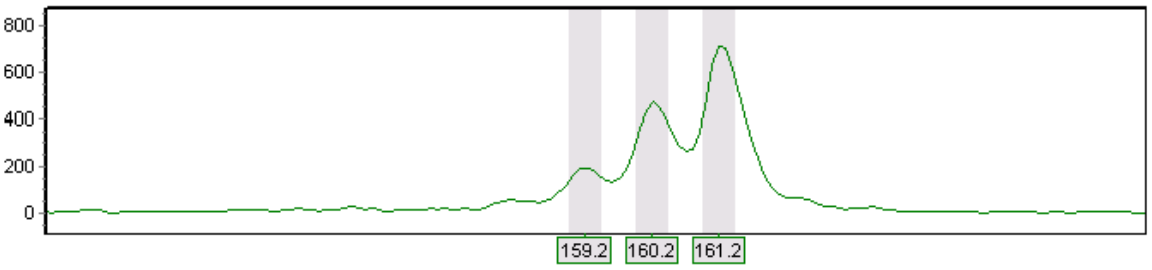
Sample 7:

Dye: Green - 3 peaks - G_2009-06-01.fsa



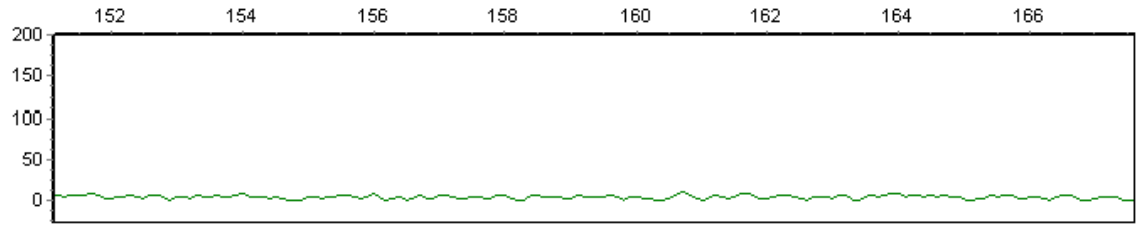
Sample 8:

Dye: Green - 14 peaks - H_2009-06-01.fsa



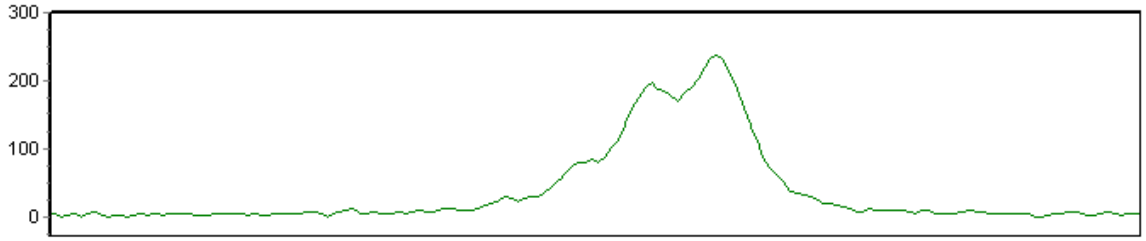
Sample 9:

Dye: Green - 3 peaks - I_2009-06-01.fsa



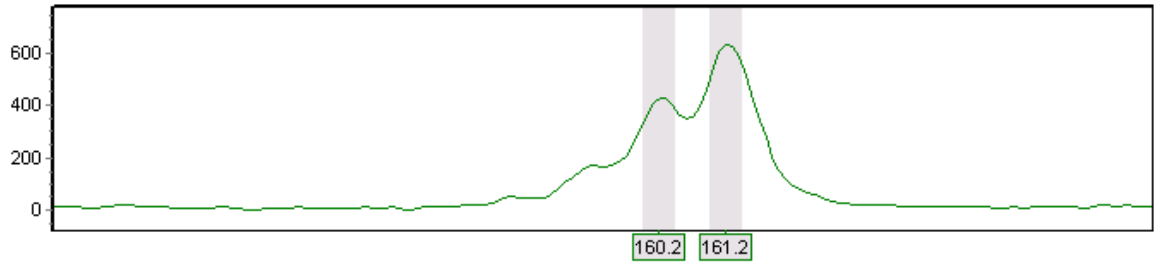
Sample 10:

Dye: Green - 7 peaks - J_2009-06-01.fsa



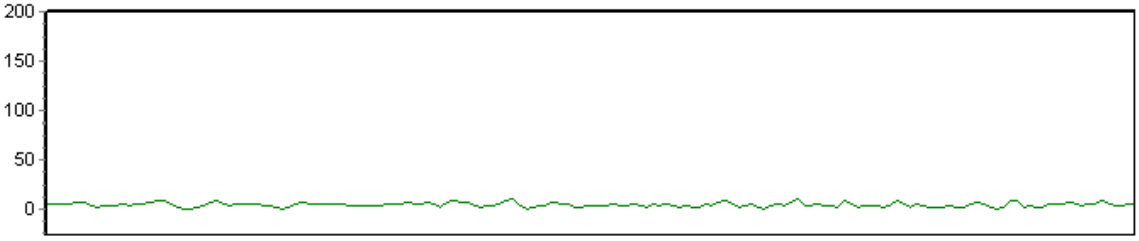
Sample 11:

Dye: Green - 10 peaks - K_2009-06-01.fsa



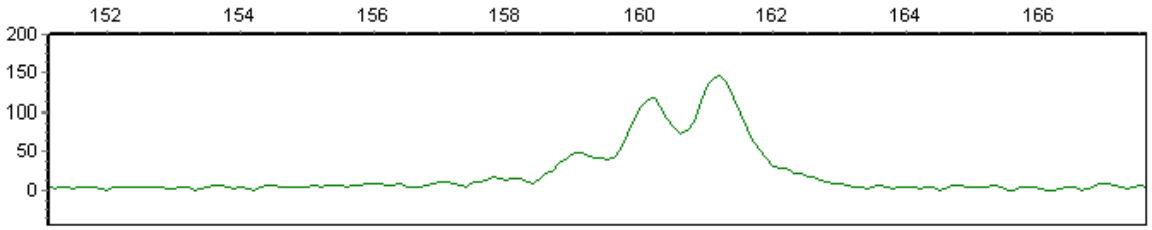
Sample 12:

Dye: Green - 3 peaks - L_2009-06-01.fsa



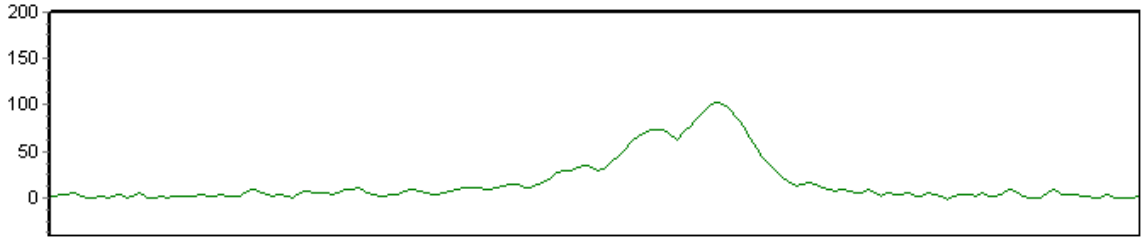
Sample 13:

Dye: Green - 7 peaks - M_2009-06-01.fsa



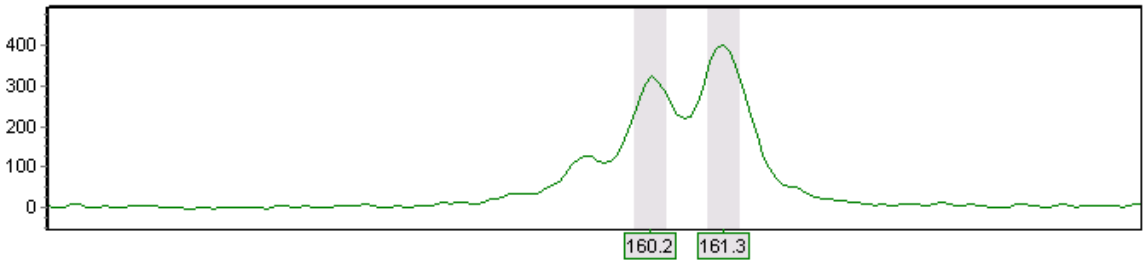
Sample 14:

Dye: Green - 6 peaks - N_2009-06-01.fsa



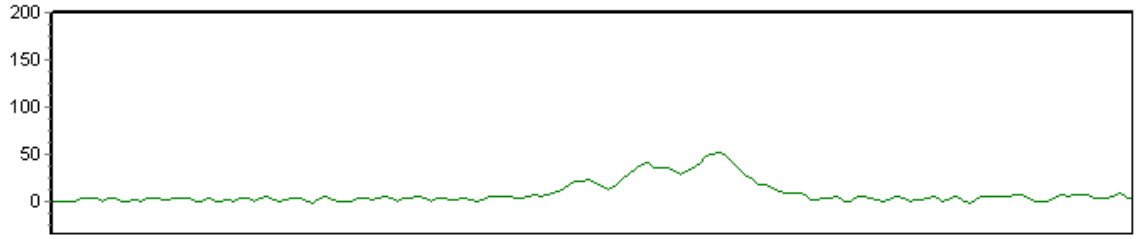
Sample 15:

Dye: Green - 9 peaks - O_2009-06-01.fsa



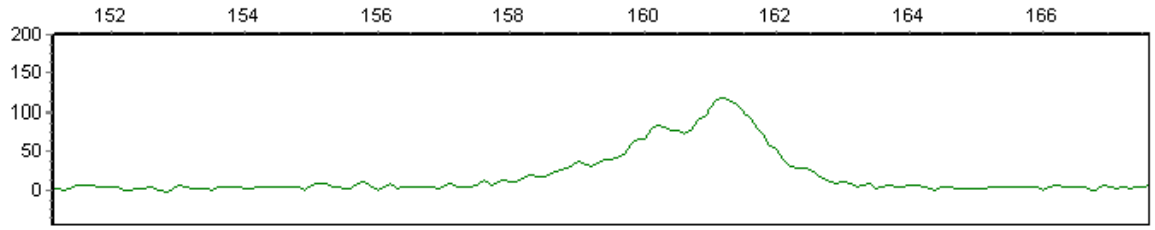
Sample 16:

Dye: Green - 9 peaks - P_2009-06-01.fsa



Sample 17:

Dye: Green - 6 peaks - Q_2009-06-01.fsa



Sample 18:

Dye: Green - 7 peaks - R_2009-06-01.fsa