Molecular characterisation of the commercially important *Agathosma* species

Lizex H.H. Hüsselmann

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science (Plant Biotechnology) at the University of Stellenbosch

Study Leaders: J-H Groenewald

Dr L.L. Dreyer

April 2006

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any other university for a degree.

L.H.H. Hüsselmann

Date signed

Summary

The development of a reliable and reproducible method for the genetic characterisation and identification of the commercially important *Agathosma* species was investigated. Previous research attempts aimed at developing a reliable and reproducible method of identifying these *Agathosma* species failed, mostly because these studies were based on phenotypic traits and these methods were therefore influenced by environmental factors. In this study amplified fragment length polymorphisms (AFLPs) were successfully used to quantify the genetic variation between the *Agathosma* species and as a result three distinct groups could be identified. The data obtained were elaborated with the Dice genetic similarity coefficient, and analysed using different clustering methods and Principle Coordinate Analysis (PCoA). Cluster analysis of the genotypes revealed an overall genetic similarity between the populations of between 0.85 and 0.99. The AFLP-based dendrogram divided the populations into three major groups: (1) the *A. serratifolia* and *A. crenulata* populations, (2) the putative hybrid, *A. betulina* X *A crenulata* populations, and (3) the *A. betulina* populations, confirming that this technique can be used to identify species. The question of hybridisation was also clarified by the results of the PCoA, confirming that the putative hybrid is not genetically intermediately spread between the *A. crenulata* and *A. betulina* populations, and that it is genetically very similar to *A. betulina.* The putative hybrid can therefore rather be viewed as a genetically distinct ecological variant of *A. betulina*.

As the AFLP technique cannot be directly applied in large-scale, routine investigations due to its high cost and complicated technology, the development of polymerase chain reaction (PCR)-based molecular markers, able to accurately identify the species, was undertaken. Due to the superior quality of *A. betulina* oil, the development of such markers is especially critical for this species. Several species-specific AFLP markers were identified, converted to sequence characterised amplified regions (SCARs) and ultimately single nucleotide polymorphisms (SNPs) were characterised. The developed SCARs were unable to distinguish between the species. The conversion of AFLP fragments to SCARs is problematic due to multiple fragments being amplified with the AFLP fragment of interest. The diagnostic feature of the SNP-based markers was not sensitive enough, since this technique could not distinguish between the *A. betulina* and *A. crenulata* and/or the putative hybrid populations. The SNPs that were characterised were found not to be species-specific; they were only specific to the particular clone.

Although a quick and robust marker specific for *A. betulina* has not yet been developed, this study sets the stage for future genetic studies on *Agathosma* species. Such a marker, or set of markers, would be an invaluable contribution to a blooming buchu oil industry.

Opsomming

Die ontwikkeling van ‘n betroubare en herhaalbare metode vir die genetiese karakterisering en identifisering van die kommersiëel belangrike *Agathosma* spesies is ondersoek. Vorige navorsingspogings gemik op die ontwikkeling van so ‘n betroubare en herhaalbare identifiserings-metode van *Agathosma* spesies het misluk siende dat die studies gebaseer was op fenotipiese kenmerke en gevolglik beïnvloed was deur omgewingsfaktore. In die studie is geamplifiseerde fragment lengte polimorfismes (AFLPs) suksesvol gebruik om genetiese variasie tussen die *Agathosma* spesies te kwantifiseer en gevolglik kon drie unieke groepe geidentifiseer word. Die data verkry is ondersteun deur die Dice geneties- verwante koeffisient en ontleed deur gebruik te maak van verskillende groeperingsmetodes en Beginsel Koördinaat Analiese (BKoA). Groep analiese van die genotipe dui algehele genetiese verwantheid tussen die populasies van tussen 0.85 en 0.99 aan. Die AFLP- gebaseerde dendrogram verdeel die populasies in drie hoof groepe: (1) die *A. serratifolia* en *A. crenulata* populasie, (2) die skynbare hybried, *A. betulina* X *A crenulata* populasie, en (3) die *A. betulina* populasie, bevestigend daarvan dat die tegniek gebruik kan word vir spesies identifisering. Die kwessie van hibridisering is ook uitgeklaar deur die resultate van die BKoA, bevestigend daarvan dat die skynbare hibried nie geneties intermediêr versprei is tussen die *A. betulina* en *A. crenulata* populasies nie en dat dit geneties nou verwant is aan *A. betulina*. Die skynbare hibried kan dan gevolglik as ‘n geneties unieke ekologiese variant van *A. betulina* beskou word.

Aangesien die AFLP tegniek nie direk toegepas kan word in grootskaalse roetine ondersoeke nie as gevolg van hoë kostes en gekompliseerde tegnologie, is die ontwikkeling van polemerase ketting reaksie (PKR)-gebaseerde molekulêre merkers wat instaat is om spesies akkuraat te kan identifiseer, onderneem. As gevolg van die superieure kwaliteit van *A. betulina* olie is die ontwikkeling van sulke merkers veral krities vir die spesies. Verskeie spesies-spesifieke AFLP merkers is geidentifiseer, omgeskakel na volgorde gekaraktiseerde geamplifiseerde areas (VKAA) en uiteindelik is enkel nukleotied polymorfismes (ENPs) gekarakteriseer. Die ontwikkelde VKAAs was nie in staat om tussen die spesies te onderskei nie. Die omskakeling van AFLPs na VKAAs is problematies as gevolg van veelvuldige fragmente wat ook saam met die verlangde fragment geamplifiseer word. Die diagnostiese kenmerk van die ENP-gebaseerde merkers was nie sensitief genoeg nie, aangesien die tegnieke nie tussen die *A. betulina* en *A. crenulata* en/of die skynbare hibried populasies kon onderskei nie. Die ENPs wat gekarakteriseer is, is bevind om nie spesies-spesifiek te wees nie en was slegs spesiek vir daardie spesifieke kloon.

Alhoewel ‘n vinnige en robuuste merker spesifiek vir *A. betulina* nog nie ontwikkel is nie, stel die studie ‘n platvorm daar vir toekomstige genetiese studies op *Agathosma* spesies. So ‘n merker of stel merkers sal van onskatbare waarde wees vir die ontluikende boegoe olie bedryf.

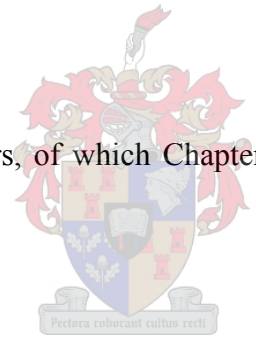
Dedicated to

My late father, Ernst Jacobus Hüsselmann, who has been a great inspiration to me.

PREFACE

The experimental work in this study was carried out in the Institute for Plant Biotechnology, University of Stellenbosch, under the supervision of Mr J-H Groenewald and Dr L. L. Dreyer.

These studies are the original work of the author. Where use was made of the work of others, it has been duly acknowledge in the text.



This thesis is a compilation of chapters, of which Chapters 3 and 4 are presented in the style of publication.

Chapter One	General Introduction
Chapter Two	Literature Review
Chapter Three	Molecular Characterisation
Chapter Four	Species Identification
Chapter Five	General Discussion and Conclusions

Acknowledgements

This study was performed in the Institute for Plant Biotechnology (IPB) in the Department of Natural Science, University of Stellenbosch. I would like to acknowledge the National Research Foundation and the U S Agency for International Development for their financial support.

I would like to thank my Lord and Saviour, Jesus Christ, for strength and wisdom during the course of this study. My sincere thanks is expressed to Jan-Hendrik Groenewald for his advice on the project. I would also like to thank Dr L Dreyer for her input and advice. Thanks go to my colleagues Dr M Venter and F. N. Hiten, and other colleagues at the IPB for their support and assistance throughout this project. I am indebted to Louisa Blomerus at Elsenburg Agricultural Research Council for her assistance and to the local buchu farmers for providing sample material. I would also like to extend my sincere appreciation to my mother, sisters, brothers, and especially my wife, for their prayers, encouragement and support during the past three years.

Table of Contents

List of Figures and Tables

Chapter Two

Chapter Three

Chapter Four

List of Abbreviations

A	adenine
AFLP	amplified fragment length polymorphism
APS	ammonium persulphate
bp	base pair
C	cytosine
CFR	Cape Floristic Region
CO_2	carbon dioxide
CsCl	cesium chloride
CTAB	cetyltrimethylammonium-bromide
cv	cultivar
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
E	east
EDTA	ethylene diamine tetra acetic acid
G	guanine
gDNA	genomic DNA
GLC	gas-liquid chromatography
GS	genetic similarity
ha	hectare
IPB	Institute for Plant Biotechnology
Kb	kilobase
kg	kilogram
LB	Luria broth
LR-iPCR	long-range inverse PCR
M	molar
MAS	marker-assisted selection
Mg^{+2}	magnesium cation
MgCl	magnesium chloride
$mg.ml^{-1}$	milligram per milliliter
ml	milliliter
mM	millimolar
m/v	mass per volume
MXCOMP	matrix comparison
NaCl	sodium chloride

ng	nanogram
PAGE	polyacrylamide gel electrophoresis
pBS	pBluescript
PCoA	principle coordinate analysis
PCR	polymerase chain reaction
PNK	T4-polynucleotide kinase
QTL	quantitative trait loci
RAPD	random amplified polymorphic DNA
rpm	revolutions per minute
RFLP	restriction fragment length polymorphism
S	south
SAHN	sequential, agglomerative, hierarchical and nested clustering
SAMPL	selective amplification of microsatellite polymorphic loci
SCARs	sequence characterized amplified regions
SCMV	sugar-cane mosaic virus
SSR	simple sequence repeat
T	thymidine
TBE	tris, boric acid, EDTA
TEMED	N,N,N,N-tetramethylethylenediamine
Tm	melting temparture
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloric acid
µCi	microcurie
µg	microgram
µl	microliter
µM	micromolar
UPGMA	unweighted pair-group method of averages
US	University of Stellenbosch
UV	ultraviolet
V	volts
v/v	volume per volume
w/v	weight per volume

Chapter One

Introduction

Buchu, an indigenous medicinal plant belonging to the family Rutaceae, is distributed throughout the Western Cape Province of South Africa. These species are endemic to the high-altitude regions of the western part of the Cape Floristic Region (CFR) (Goldblatt and Manning, 2000), where a Mediterranean climate prevails. They roughly extend from Clanwilliam in the north, to Tulbagh and southward to Paarl and Riversdale (Blommaert and Bartel, 1976). The three most important commercial species, *Agathosma betulina* (Berg.) Pillans, *A. crenulata* (L.) Pillans and *A. serratifolia* (Curt.) Spreeth were initially wild-harvested for their use in medicine (Spreeth, 1976). However, the latter became less important and only *A. betulina* and *A. crenulata* were recognised as "true buchus" by industry (Roberts, 1990). *A. crenulata* and *A. betulina* have mainly been cultivated for their use in medicine, with a large component of the essential oil also being used in the flavour and fragrance industries (Turpie *et al.*, 2003). The export industry is estimated to be worth R100 million per annum, of which approximately 50% is currently harvested from natural veld (Rust, 2003).

Due to renewed interest in buchu, increased exploitation of the wild buchu species jeopardises the future survival of these species. This necessitates the cultivation of buchu in order to supply the growing demand for oil and to protect the country's indigenous genetic material. Apparent hybridisation between the two true buchu species (*A. betulina* and *A. crenulata*) has taken place over the years, and this has complicated the identification of these two species. Initial identification was based on leaf shape, but because of many intermediate leaf shapes this method of classification is not very reliable. Leaf shape had, however, become an important criterion for evaluating buchu, and overseas buyers traditionally preferred the round-leaf species (*A. betulina*) due to its higher buchu camphor (diosphenol) content (Blommaert and Bartel, 1976; Webber *et al*., 1999).

Proper means of classification are therefore required to identify species with the desired oil profile. In order to certify the authenticity of a species, various morphological, anatomical, taxonomical and chemo-taxonomical studies have been conducted to classify the two commercially important species *A. betulina* and *A. crenulata*. The classification system followed potentially has a huge impact on the buchu industry, since wrong classification can be disastrous and lead to extensive financial losses. A reliable and reproducible method of classification is therefore imperative.

Despite the problems caused by apparent hybridisation, morphology is still the most commonly used method of classification. Spreeth (1976) found that this method is only reliable for *A. betulina* and *A. crenulata* in their natural habitat in the absence of apparent hybridisation. In regions where apparent hybridisation had taken place, it was no longer certain, since the taxonomical delimitation of the two species was based solely on leaf shape (Blommaert and Bartel, 1976). Thus this method of identification is today neither consistent nor trustworthy.

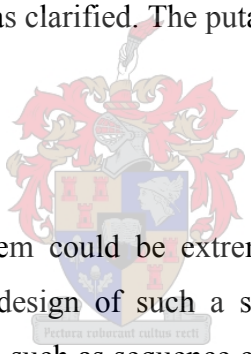
A chemotaxonomic classification method proposed by Endenburg (1972) revealed that in cases where morphological characteristics fail to give a clear distinction, the buchu species could be distinguished on the chemical composition of their oil. Analytical studies done by Blommaert and Bartel (1976), including gas-liquid chromatography (GLC) and ultraviolet (UV) spectrophotometry analyses of buchu oils, showed that the presence of a relatively large percentage of ρ-diosphenol and diosphenol in *A. betulina* leaf oil is the only valid criterion for botanically separating the two species *A. betulina* and *A. crenulata*. Posthumus *et al*. (1996) carried out a study to distinguish among the three taxa *A. betulina*, *A. crenulata* and a putative *A. betulina* X *A. crenulata* hybrid, based on the chemical composition of the oil, and confirmed that the three taxa could be distinguished by their monoterpene chemistry. Furthermore, a comparative analysis done on *A. crenulata* and *A. betulina* oil found *A. betulina* to have the highest menthonethiol content, confirming that *A. betulina* has better oil than *A. crenulata* (Posthumus *et al*., 1996). Although, these studies could classify the species based on their oil composition, they were unstable and not reproducible due to the influence of various experimental factors such as time of harvest, distillation methods and growing conditions.

To date, no method of classification of buchu has thus proven to be reliable or reproducible. The development of molecular biology techniques, such as DNA-based markers, has provided a new opportunity for genetic characterisation, allowing the direct comparison of different genetic material without environmental influences. The concept of genetic polymorphism is fundamental to all current methods of determining genetic identity and relatedness (Potter *et al*., 2002). Various molecular marker techniques, such as random amplified polymorphic DNAs (RAPDs), simple sequence repeats (SSRs), selective amplification of microsatellite polymorphic loci (SAMPL) and amplified fragment length polymorphisms (AFLPs) are available to detect diversity at the DNA level. One of these techniques, AFLPs, has proven to be valuable in genotype characterisation in

many crop species (Vos *et al*., 1995). No genetic studies such as genotype characterisation have to date been conducted on *Agathosma* species.

The aim of the present study was therefore to: (1) establish a molecular marker system to identify commercially cultivated species, (2) estimate genetic diversity within commercially important *Agathosma* species and (3) to develop a robust molecular marker to routinely screen large numbers of *Agathosma* species.

The AFLP technique (Chapter 3) was successfully applied and data generated were used to assess genetic diversity of the commercially important *Agathosma* species. The data generated clearly group these three species into three distinct clusters, i.e. *A. betulina, A. crenulata* and the putative *A. betulina* X *A. crenulata*. In addition, the status and genetic composition of the putative hybrid between *A. betulina* and *A. crenulata* was clarified. The putative hybrid is most likely only a genetic or ecological variant of *A. betulina*.



Furthermore, a molecular marker system could be extremely useful to validate the identity of commercially cultivated species. The design of such a system was attempted in Chapter 4 by converting AFLPs to molecular markers such as sequence characterised amplified regions (SCARs) and single nucleotide polymorphisms (SNPs) that may be used to routinely screen large numbers of the commercially important species. However, the markers developed were not specific enough to distinguish between the species. Thus further research needs to be done to develop a quick and robust marker to screen for the most sought-after buchu.

The concluding Chapter 5 summarises the main outcomes of the present study including (1) the successful establishment of a molecular marker technique for the commercially important *Agathosma* species, (2) that there is substantial genetic variation among the studied *Agathosma* species and (3) the first attempts towards the development of a molecular marker for the routine screening of these species.

References

Blommaert, K.L.J. and Bartel, E. 1976. Chemotaxonomic aspects of the Buchu species *Agathosma betulina* (Pillans) and *Agathosma crenulata* (Pillans) from local plantings. *Journal of South African Botany* 42 (2): 121-126.

Endenburg, H.G. 1972. A study of the yield and composition of the essential oils from *Agathosma crenulata* and *A. betulina*. MSc-Thesis. University of Stellenbosch.

Goldblatt, P. and Manning, J. 2000. Cape plants. A conspectus of the Cape flora of South Africa *Strelitzia* 9. National Botanical Institute, Pretoria & Missouri Botanical Garden, Missouri, pp 68-73.

Potter, D., Gao, F., Aiello, G., Leslie, C. and McGranahan, G. 2002. Intersimple sequence repeat markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia*) cultivars. *Journal of American Society for Horticultural Science* 127: 75-81.

Posthumus, M.A., van Beek, T.A., Collins, N.F. and Graven, E.H. 1996. Chemical composition of the essential oil of *Agathosma betulina*, *A. crenulata* and an *A. betulina* x *crenulata* Hybrid (Buchu). *Journal of Essential Oil Research* 8: 223-228.

Roberts, M. 1990. Indigenous Healing Plants. Southern Book Publishers, Halfway House, South Afica, pp. 42-43.

Rust, K. 2003. The beauty of buchu. *Food Review* 65: 55-56.

Spreeth, A.D. 1976. A revision of the commercially important *Agathosma* species. *Journal of South African Botany* 42(2): 109-119.

Turpie, J.K., Heydenrych, B.J. and Lamberth, S.J. 2003. Economic value of terrestrial and marine biodiversity in the Cape Floristic Region: implications for defining effective and socially optimal conservation strategies. *Biological Conservation* 112: 233-251.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. *Nucleic Acids Research*. 23(21): 4407-4414.

Webber, L.N., Magwe, M.L. and van Staden, J. 1999. Alternative uses for some invader plants: turning liabilities into assets. *South African Journal of Science* 95: 329-331.

Chapter Two

Towards the molecular characterisation of *Agathosma* species

2.1. Introduction

This review focuses on the molecular characterisation of three commercially important South African *Agathosma* taxa, namely *Agathosma betulina* (round-leaf buchu), *A. crenulata* (oval-leaf buchu) and their putative hybrid, *A. betulina* X *A. crenulata*. Buchu has such a huge market potential, both in and outside South Africa that it has become one of the most sought after indigenous crops. However, the genotype affects the quality of the oil to such an extent that it has become imperative to distinguish between the commercially important species in order to propagate and sell the right buchu species. Since the discovery of the commercial value of these species, various studies have been directed at the correct classification of the two species. Linnaeus conducted phenotypic (morphology and anatomy based) studies on buchu as early as 1755 (Bean, 1993). Fluck *et al.* (1961) and Kaiser *et al.* (1975) performed chemotypic and analytical studies on buchu oil. Blommaert and Bartel (1976) and Posthumus *et al.* (1996) conducted further analytical studies on buchu oil to distinguish between the species *A. betulina*, *A. crenulata* and *A. betulina* X *A. crenulata*. These studies were, however, neither reliable nor reproducible, and thus did not aid the persistent need for correct species classification.

Genotypic studies, which are reliable and reproducible, thus seem superior in their ability to solve the problem of classification, because the markers generated are not influenced by the environment and can be scored at all stages of plant growth. No molecular studies have thus far focussed on the genetic characterisation of these *Agathosma* species, despite the availability of various molecular marker tools. Molecular markers commonly used in practice include random amplified polymorphic DNAs, microsatellites or simple sequence repeats, selective amplification of microsatellite polymorphic loci and amplified fragment length polymorphisms. The choice of marker is often dictated by the intended application, convenience and the costs involved. For this study, amplified fragment length polymorphism was the marker of choice for the molecular characterisation of the commercially important *Agathosma* species.

2.2. Historical background

Buchu has come a long way since its discovery by the San and Khoi, two tribes native to southern Africa. They used the dried and powdered buchu leaves mixed with sheep's fat to perfume their bodies. This mixture also served as a deterrent to insects. They used the name buchu or "boegoe" for several strongly scented plants, which led to considerable confusion. Today the group of aromatic herbs and shrubs known as buchus are restricted to members of the genus *Agathosma* (previously *Barosma*) in the Rutaceae (Simpson, 1998).

There were initially three commercially important species (Spreeth, 1976), i.e. *Agathosma betulina* (Berg.) Pillans (round-leaf buchu), *A. crenulata* (L.) Pillans (oval-leaf buchu) and *A. serratifolia* (Curt.) Spreeth (long-leaf buchu). The latter became less marketable because plant material was mixed with *Empleurum serrulatum* Ait by exporters (Spreeth, 1976) in order to increase its bulk. This damaged the desirability of *A. serratifolia* and resulted in the species becoming less popular. Today only two commercially grown species, *A. betulina* and *A. crenulata*, are recognised as "true buchu" (Roberts, 1990). *A. betulina* is a fragrant shrub, which seldom exceeds a height of one metre. The leaves are small, round, dotted with oil glands on the abaxial side and have a strong aromatic smell (Roberts, 1990) (Fig. 2.1). The flowers are small, star-shaped, ranging in colour from white to pink (Fig. 2.2). *A. crenulata* is an aromatic woody shrub, which reaches a height of up to two metres. Their leaves are larger, elongated (Fig. 2.1) and the flowers are white or mauve (Simpson, 1998).

Figure 2.1 Foliage leaves of *A. betulina, A. crenulata* X *A. betulina* and *A. crenulata* displaying the distinctly different leaf shapes associated with the two species, and the intermediate leaf morphology of the putative hybrid.

Figure 2.2 The small, star-shaped flowers of buchu range in colour from white to pink and are mostly pollinated by bees.

These species are endemic to the high-altitude regions of the western part of the Cape Floristic Region (CFR) of South Africa (Goldblatt and Manning, 2000), where a mediterranean climate prevails. They roughly extend from Clanwilliam in the North, to Tulbagh, and southward to Paarl and Riversdale respectively (Blommaert and Bartel, 1976) (Fig. 2.3). Since the establishment of buchu plantings in various parts of many cultivated plants has risen (Blommaert and Bartel, 1976). Alleged hybridisation was thought to have taken place in the Western Cape over the years wherever the two species were interplanted. New plantings were established from the seeds collected from these interspecific crosses and that caused a great deal of confusion as to the identity of the new plantings. The result was that leaf shape, which has always been the best character by which to distinguish between the two species, started displaying many intermediate forms (Fig. 2.1). Leaf shape is an important criterion for evaluating buchu, since overseas buyers traditionally prefer the round-leaf species (*A. betulina*) for its higher buchu camphor (diosphenol) content (Blommaert and Bartel, 1976; Webber *et al.*, 1999).

Figure 2.3 Western part of the Cape Floristic Region depicting the geographic distribution of the populations from which Agathosma species were collected. Samples collected from populations in the Paarl region represented cultivated stands, while those from other regions represented wild (natural) stands.

In recent years an international growing demand for buchu oil has led to increased exploitation of wild buchu populations, jeopardizing the species' future survival. Cultivation of buchu was thus initiated to supply in the growing demand and to protect the country's indigenous genetic material.

2.3. Exploitation as a natural resource

Agathosma crenulata and *A. betulina* are collected from Mountain Fynbos, mainly from a region of roughly 600 000 ha in the western part of the CFR. In addition to a small local market for medicinal use and buchu brandy, the buchu industry has a large export component, as the essential oils are desirable in the international flavour and fragrance industries (Turpie *et al.*, 2003). The export industry is estimated to be worth about R100 million per annum, of which approximately 50% is harvested from natural veld (Rust, 2003). This is confirmed by Wesgro, a South African company, that reports that 90% of buchu production is exported, with Europe, in particular Germany, as the most important destination. The demand for natural food ingredients by consumers is the driving force that manufacturers simply have to cater for. As a result of the increased interest in buchu,

over-commercialisation followed directly, since extracts from these plants find application in the medicinal, food, flavour and fragrance industries. In addition, Webber *et al*., 1999 mentions that the use of the essential oil in aromatherapy needs to be exploited futher. Buchu plants also have a small market as ornamental garden plants (Simpson, 1998).

Medicinal properties

Buchu has been used for centuries as a medicine by the San and Khoi tribes who chewed its leaves to cure stomach ailments (Rust, 2003). They introduced these herbs to the European settlers who took them to Europe, where they were used medicinally as a diuretic (Lis-Balchin *et al.*, 2001). Buchu was introduced into the pharmaceutical industry as early as 1821, when Reece and Company imported *A. betulina* from South Africa (Simpson, 1998). Today *A. betulina* and *A. crenulata* species are still grown widely for their leaves and oil. The leaves are dried and used medicinally (Roberts, 1990). Buchu vinegar played an important role in the Crimean and First World Wars as a powerful antiseptic with which to clean wounds (Rust, 2003). Buchu is currently used for stimulating perspiration in rheumatism and gout, to treat cholera, kidney diseases, haematuria, calculus, infections of the bladder, urethra and prostate, and as a digestive tonic (Lis-Balchin *et al.*, 2001). Buchu is one of three South African medicinal plants used in international medicine and *A. betulina,* in particular, is recognised as an official medicine in the 1977 edition of Martindale's Pharmacopoeia (Simpson, 1998). Newall *et al.* (1996) stated that buchu possesses urinary, antiseptic and diuretic properties and can be used for cystitis, urethritis, prostatitis and especially for acute catarrhal cystitis.

Simpson (1998) referred to buchu as "South Africa's amazing remedy". This statement was supported by extensive analyses of the essential oil of buchu species. Diosphenol (barosma camphor) was identified as the active ingredient of the antiseptic and diuretic effects of buchu oil. This could account for the stimulation of perspiration that an infusion causes, as well as for the remarkable flushing action of the kidneys (Roberts, 1990). In contrast, *A. crenulata* contains approximately 50% pulegone, a hepatotoxin (Posthumus *et al.*, 1996) that makes it potentially as toxic as pennyroyal, an abortifacient in folk medicine (Aetna InteliHealth, 2005).

The food-, flavour- and fragrance industries

The major use of buchu oil is in the flavour industry, where it is used to enhance fruit flavours, especially blackcurrant flavours. The minor compound, menthonethiol (3-oxo-ρ-menthane-8-thiol),

and other sulphur components such as diosphenol, are responsible for the characteristic blackcurrant smell and flavour (Posthumus *et al.*, 1996; Fuchs *et al.*, 2001). Buchu stands superior to other agents as it provides a potent berry flavour that can be married to various berry beverages to give added flavour and health value to the product (Rust, 2003). The oil is also used in perfumes such as colognes and chyprP bases, and for fruity notes.

Buchu material is currently sold by farmers to distillers at ± R55/kg wet material (*A. betulina*) and the essential oil (1 kg oil/100 kg wet material) is sold to clients at ± R6 000/kg (D. Malherbe, personal communications, Afriplex). To make ends meet, distillers therefore have to add value to the product that they market. Buchu oil is obtained by steam distillation from leaves and the main components of this oil are limonene, menthone, isomenthone, pulegone and bifunctional diosphenols (Posthumus *et al.*, 1996). These components are fractionated from the oil, adding value to buchu oil. South African companies, such as Afriplex, have geared themselves up to add value to buchu oil by investing in super-critical CO_2 extraction technology to further process and fractionate it into many individual components. They have already developed a blackcurrant flavour that has become a flagship product for the company, ultimately adding value to a South African resource. This company is the first of its kind in South Africa to make their own buchu formulae to satisfy overseas and local demand (Rust, 2003).

Aromatherapy

A market that has not been fully exploited is the use of the oil by aromatherapists and homeopaths in homeopathic medicines (Webber *et al.*, 1999). A number of richly scented buchus make the "most beautiful addition to pot-pourris" (Roberts, 1990). These buchu oils should be used sparingly, as the scent is quite overpowering, especially that of *A. serpyllacea* Licht. ex Roem. & Schult (lemon-scented), *A. dielsiana* Schltr. ex Dummer and *A. cerefolium* (Vent.) Bartl. & H.L.Wendl. These buchu species are also being tested for use in cosmetics, soaps, food colouring and perfumes (Roberts, 1990). Some species have an agent that blocks ultraviolet light, and are therefore actively researched for their potential future inclusion in the manufacture of cosmetics.

Ornamental garden plants

In addition to the utility of different buchu species in the industries discussed above, some species are also grown as ornamental garden plants. These plants, which flower in the winter and spring, have dainty pink, mauve or white, star-shaped flowers (Fig. 2.2). Their bright green leaves, fresh aromatic smell throughout the year and attractive winter displays are assets to any garden (Simpson 1998).

Given the huge market potential of various *Agathosma* species, it is crucial that proper means of classification are in place to classify species with the desired oil profiles. Analytical studies of buchu oil have been ongoing since these plants became so sought after both nationally and internationally. Endenburg (1972) found that the oil of *A. crenulata* consisted mainly of 1-pulegone and contained very little diosphenol, whereas that of *A. betulina* contained a relatively low proportion of pulegone, with diosphenol and ρ-diosphenol as its major constituents. Posthumus *et al.* (1996) also found that *A. betulina* and *A. crenulata* differ in their oil composition; *A. betulina* is characterised by a high content of (iso) mentone (31%), ρ-diosphenol (4 %) and the cis-and trans-3-oxo- ρ -menthane-8-thiol (3%); while *A. crenulata* oil contains a high percentage of pulegone (54%) and higher quantities of trans-3-oxo-ρ-menthane-8-thiol. According to Webber *et al.* (1999), good quality oil has a low diosphenol content and high menthonethiol content. The oil from *A. betulina* has the highest menthonethiol content, confirming that *A. betulina* has better oil than *A. crenulata.* With this comparative analysis, *A. betulina* oil became the sought-after product by international flavour and fragrance houses. A demand for essential oil, high in diosphenol and menthonethiol, prompted the need for a consistent and reproducible method of classification of buchu species to deliver essential oil of high quality and standard.

In order to address this need, various morphological, anatomical, taxonomical and chemotaxonomical studies have since been conducted to distinguish between the two commercially important species *A. betulina* and *A. crenulata*. The method of classification that is followed potentially has a huge impact on the buchu industry, since wrong classification can be very costly and damaging.

2.4. Morphology, anatomy and taxonomy

A morphology-based taxonomic study by Spreeth (1976) on the commercially important *Agathosma* species, *A. betulina* and *A. crenulata,* revealed both significant differences and similarities. In order to define the species, detailed morphological studies of the stems, leaves, flowers and fruits were supplemented by anatomical studies of the leaves, stems and roots of both species. Leaf shape proved to be the only consistent morphological difference between the species, with *A. betulina* having a smaller, rounder leaf with a recurved apex and *A. crenulata* a larger, oval-shaped leaf with no recurved apex. Clear differences in leaf morphology and chemical composition of the oil lead to the subdivision of the taxon *A. crenulata s.l.* into the taxa *A. crenulata s.s.* and *A. serratifolia* (Spreeth, 1976).

Visual classification of buchu based on leaf shape has been the most widely employed method to distinguish between the two species (Pillans, 1950), but caused problems in the case of apparent hybridisation (Blommaert and Bartel, 1976). The taxonomic classification of the two buchu species is almost solely based on leaf shape, which may be valid for species in their natural habitat, but becomes problematic where the ranges of the two species overlap (Fig. 2.1) and they are thought to hybridise. The problems experienced with apparent hybridisation also extend to the leaf-shape based classification of species in cultivated lands. The morphology-based method of classification thus proved problematic.

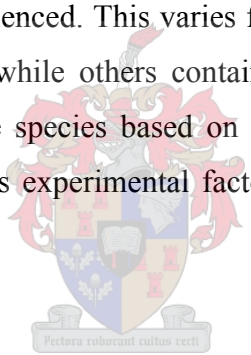
Chemotaxonomy and oil profiles

A chemotaxonomic study by Endenburg (1972) revealed that the buchu species could be distinguished on the basis of the chemical composition of their oil. At the time of their study, very little was known about the consistent differences in chemical composition of the oil of *A. betulina* and *A. crenulata.* Similarly, nothing was yet known about chemical variation present within a single species.

Spreeth (1976) also conducted a biochemical investigation to distinguish between the species *A. betulina* and *A. crenulata* in terms of amino acids, sugars and organic acids. His results revealed no differences among the species with respect to amino acids and sugars, but substantiated the fact that the three species differ in terms of citric acid contents. It was these differences that led him to

propose that *A. crenulata* should be subdivided into the taxa *A. crenulata* and *A. serratifolia*. This provided the first biochemical proof that *A. crenulata* and *A. serratifolia*, (that were lumped as *A. crenulata* by Pillans (1950)) are in fact separate species. Analytical studies by Blommaert and Bartel (1976), including gas-liquid chromatography (GLC) and ultraviolet (UV) spectrophotometry of buchu oils, confirmed that the presence of a large percentage of ρ-diosphenol and diosphenol in *A. betulina* leaf oil is the only valid criterion for botanically separating the two species *A. betulina* and *A. crenulata*. A similar chemotaxonomic study was done by Posthumus *et al.* (1996) in order to distinguish between the three taxa *A. betulina*, *A. crenulata* and an *A. betulina* X *A. crenulata* hybrid. They found that the three taxa could also be distinguished by their monoterpene chemistry. The key characteristic was pulegone and 8-mercapto-ρ-menthan-3-one isomer ratio.

However, Webber *et al.* (1999) stated that the percentage of these constituents in the oil varies from batch to batch and from area to area. The oil content of the plants depends on the growth stage and growing conditions that they had experienced. This varies from 0.3% to 0.9% on a dry mass basis. Some species contain no diosphenol, while others contain more than 90% of this essential oil. Hence, these studies could classify the species based on their oil composition, but were neither reliable nor reproducible due to various experimental factors such as time of harvest, distillation methods and growing conditions.



Various morphological, anatomical, taxonomic and chemo-taxonomic studies have thus been undertaken in an attempt to clarify the uncertainty around the classification of the two commercially important species. As outlined above, the results of all of these approaches were, however, influenced by environmental (time of harvest) and experimental factors (method of distillation). DNA fingerprinting techniques offer reliability and reproducibility in the quest for a consistent method of classification. There are various molecular marker techniques (i.e. RAPDs, SSRs, SAMPL and AFLPs) available to detect diversity at the DNA level. With the aid of DNA fingerprinting techniques, a study that can reveal genetic variation within species of *Agathosma* is now feasible. The genetic profiles can be used to quantify diversity of the commercially important *Agathosma* species and thus solve the problems of classification.

2.5. Molecular phylogenetics

Genetic markers are specific locations on a chromosome, which serve as landmarks for genome analysis (Kumar, 1999). According to Kumar (1999), are there basically two types of genetic markers: morphological and molecular markers. The inheritance of morphological markers can be monitored visually without specialized biochemical or molecular techniques. Molecular markers, in contrast, can reveal polymorphisms at the protein level (biochemical markers) or at the DNA level (DNA markers). Paul *et al*. (1997) report that the relatively low levels of polymorphisms detectable limit morphological markers whereas molecular markers overcome this problem. A DNA marker can be any DNA fragment that can be used as a marker of genetic variation within and among individuals and taxa. One might for example use a particular DNA marker as a diagnostic trait, as a tool for management of a breeding programme, as an aid to systematic analyses, or in a wide variety of ways in basic evolutionary biology. DNA markers have become available for both basic and applied studies (Gupta *et al.*, 1999). One of the most extensive uses of DNA markers has been the development of detailed genetic and physical chromosome maps in many different organisms (Reiter, 2001). An important application of DNA markers in plant systems involves the improvement in efficiency of conventional plant breeding by carrying out indirect selection through DNA markers linked to the traits of interest. This can be done for both simple and quantitative trait loci (QTL). These markers are not influenced by the environment and can be scored at all stages of plant growth. Each of the available marker systems has advantages and disadvantages, and the choice of a marker system is largely dictated by the intended application, convenience and the costs involved.

The identification and use of DNA markers fall into one of three basic technique categories that use either hybridisation or the polymerase chain reaction (PCR): (1) hybridisation-based (non-PCR) techniques, (2) arbitrarily-primed PCR and other PCR-based multi-locus profiling techniques, and (3) targeted sequence and single locus PCR (Karp and Edwards, 1997). DNA markers of interest within this study are PCR-based DNA markers, specifically the sequence-arbitrary method. In the sequence arbitrary method, the PCR is performed using two sequence-dependent oligonucleotide primers. Sequence information of the desired fragment is needed to design primers, which facilitates the successful amplification of a specific DNA fragment. PCR-based markers include: (1) sequence-arbitrary methods such as random amplified polymorphic DNA's, selective amplification of microsatellite polymorphic loci, amplified fragment length polymorphisms and (2) sequence dependent method such as microsatellites or simple sequence repeats.

Random amplified polymorphic DNAs (RAPDs)

RAPD markers are produced by PCR using short oligonucleotide primers of randomly chosen sequence. Different RAPD patterns arise when genomic regions vary in terms of the presence/absence of complementary primer annealing sites. The primers are typically 10 bp long (Williams *et al*., 1990, Welsh and McClelland, 1990) and no specific knowledge of a particular DNA sequence is required to choose or produce a primer. Primers are used singly, not in combination with a second primer, as is the case for standard PCR. Due to this, amplified fragments are those regions of the genome that are flanked by "inward-oriented" sequences complementary to the primer. Allelic variation consists of the presence or absence of particular amplification products, which can be separated on agarose gels stained with ethidium bromide. The RAPD process typically reveals several polymorphic genetic segments per primer within populations; other segments may appear as monomorphic bands within or across populations (Parker *et al*., 1998). The degree of variability observed for many primers confirms that the technique is useful for addressing various biological questions, including individual identification, paternity analysis, strain identification, phylogenetic analysis, construction of genetic linkage maps, gene tagging, the identification of cultivars and the assessment of genetic variation in populations (Gupta *et al*., 1999). These applications have led to the development of species-specific, genome-specific and chromosome-specific markers and, more importantly, to the development of molecular markers for identification and selection of the desired genotypes in segregating populations during breeding programmes (Chen *et al.*, 1998).

RAPD markers are rarely inherited as codominant alleles. Loss of a priming site results in complete absence of the enclosed amplified segment, and not simply a shift in mobility on the gel. In heterozygotes, therefore, differences may appear only as differences in band intensity, which is not usually a reliable phenotype for PCR analysis. As a consequence, information on the parental origin of alleles may be inaccessible for RAPD markers, as compared to codominant markers such as RFLPs or allozymes (Parker *et al*., 1998). Because of their short length, RAPD markers may produce some artifactual amplification products, and careful control of DNA quality and amplification conditions is necessary to ensure reproducible banding patterns (Parker *et al*., 1998).

RAPD technology has proven useful in many crop plants, but it has been put to limited use because of the low level of polymorphisms detected within some species of self-pollinating plants and due to the lack of reproducibility of results (Gupta *et al.*, 1999). However, the lack of reproducibility was solved by Besse *et al*. (2004) by repeating amplifications twice with different DNA

concentrations in separate analyses, and scoring only strong and reproducible fragments. Consistent patterns were obtained. RAPD markers were used in genetic diversity studies of maize (Pejic *et al*., 1998), cultivated vanilla (Besse *et al*., 2004), wild rice (Wu *et al*., 2004), olives (Belaj *et al*., 2003) and *Pistacia* species (Golan-Goldhirsh *et al*., 2004). RAPD markers have also been employed in finding molecular evidence of hybridisation between palm species (González-Perez *et al*., 2004).

Microsatellites or simple sequence repeats (SSRs)

Microsatellites are known by a number of acronyms such as simple tandem repeats (STRs) and simple sequence repeats (SSRs). SSRs are currently the marker system of choice for marker-based genetic analysis and marker-assisted plant breeding (Reiter, 2001). SSRs are ubiquitous sets of tandemly repeated DNA motifs. The repeat regions are composed of perfectly repeated di-, tri-, tetra-, or multi- nucleotide sequence (Reiter, 2001). Compound repeats are composed of two or more repeat motifs that are frequently found (Reiter, 2001). The length of a repeat sequence varies greatly, with different alleles varying in the number of units of the repeat motif. For example, di-nucleotide SSRs have alleles that either differs by two base pairs or multiples of two base pairs. The variability in the number of repeat units is typically the basis of the observed polymorphism.

The ability to detect polymorphisms is improved to such an extent that in species where the level of polymorphism detected by restriction fragment length polymorphism (RFLP) is low, acceptable levels of polymorphism are observed with SSRs (Akkaya *et al.*, 1992; Powell *et al.*, 1996). The high degree of observed polymorphisms appears to be the result of increased rates of sequence mutation affecting the number of repeat motifs present at an SSR locus (Edwards *et al.*, 1992), with the observed variation likely due to replication slippage or unequal crossing over.

SSR loci must be cloned and sequenced before a useful marker can be generated. Small fragment genomic libraries, enriched for SSR-repeat motifs, are screened for clones containing the SSR sequence using an oligonucleotide probe complementary to the repeat motif. In order to obtain single-copy DNA sequence flanking the SSR marker, each positive clone is sequenced. Oligonucleotide primers complementary to unique DNA sequence flanking both sides of the repeat are synthesised and used for PCR amplification of the SSR.

Microsatellites are attractive for population-level comparative analyses, since they are often highly polymorphic and very stable and reproducible. The high level of polymorphisms, relative to RFLPs and RAPDs, along with a high interspersion rate, makes them abundant sources of genetic markers. The main disadvantage of microsatellites is the cost of establishing polymorphic primer sites and the investment in synthesising the oligonucleotide. Gupta *et al*. (1999) reiterated that the high costs associated with the development of polymorphic primers sites. He stressed that the main limitation of SSRs in his study was the lack of polymorphisms across *Brassica* species for comparative mapping. However, once the system is developed, it is the most informative marker system available, because the number of repeat units at a locus is highly variable and can be visualised as AFLPs, even between closely related individuals (Mazur and Tingey, 1995).

Microsatellites (SSRs) have been extensively applied in genetic diversity studies for a number of crops such as maize (Pejic *et al*., 1998), sugarcane (Cordeiro *et al*., 2003), apricots (Zhebentyayeva *et al*., 2003), olives (Belaj *et al*., 2003) and in cultivar identification and pedigree studies in grapes (Sefc *et al*., 2000). SSRs have been used in creating a saturated map for the apple (*Malus* X *domestica* Borkh.) genome (Liebhard *et al*., 2003) and in generating a genetic map of white clover (Jones *et al*., 2003).

Selective amplification of microsatellite polymorphic loci (SAMPL)

SAMPL is a hybrid method which exploits features of both SSR sequence based methods and the AFLP sequence arbitrary method (Reiter, 2001). It utilises the same DNA template as that of AFLPs. Genomic DNA is first digested by restriction endonuclease as in the AFLP method. The selective amplification employs one of the AFLP primers in combination with an SAMPL primer. The SAMPL primer essentially comprises a compound microsatellite sequence, which is anchored. Such a SAMPL primer design ensures preferential amplification of microsatellite-like sequences (Singh *et al.*, 2002). The amplification products are radio-labelled and size-separated on high-resolution polyacrylamide gels (Reiter, 2001). This method is especially effective at revealing polymorphism when the SSR motif is a perfect compound SSR (Reiter, 2001). The method has not been used very extensively, but does constitute another potential tool with which to identify and exploit DNA polymorphisms. This technique is suitable for studies where low genetic variation is expected, since SAMPL primers target the hyper-variable microsatellite loci. It has been applied in

the assessment of intra-population genetic variation in neem (Singh *et al*., 2002), where it was found that SAMPL detect more polymorphic loci per assay than AFLP markers (Singh *et al.*, 2002).

Amplified fragment length polymorphisms (AFLPs)

The AFLP technique has been identified as a robust DNA fingerprinting technique that detects significant levels of polymorphism between accessions. Its replicability, resolution, ease of use and cost efficiency, make AFLPs superior to other molecular markers (Mueller and Wolfenbarger, 1999). AFLP markers offer the following advantages: (1) taxonomic scope i.e. AFLP markers can be generated for any organism with DNA, and no prior knowledge about the genomic makeup is needed, (2) AFLP amplifications are performed under conditions of high selectivity at high stringency, thus eliminating the artificial variation that is seen in RAPD-PCR, (3) AFLP analysis requires minimal amounts of DNA, (4) AFLP markers can be generated at great speed i.e. a high ratio of polymorphisms generated per PCR experiment and (5) due to the nearly unlimited number of markers that can be generated with AFLP-PCR, using a series of different primer combinations, at least some AFLP markers will be located in variable regions and will reveal even minor genetic differences within any given group of organisms (Mueller and Wolfenbarger, 1999). The method has one main disadvantage, namely the difficulty to identify homologous markers (alleles), rendering this method less useful for studies that require precise assignment of allelic states, such as heterozygosity analyses (Mueller and Wolfenbarger, 1999).

In a comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and their effectiveness in olives, it was found that SSRs presented a higher level of polymorphism and greater information content than AFLPs and RAPDs (Belaj *et al.*, 2003). However, AFLPs were the most efficient marker system due to their capacity to reveal the highest number of bands per reaction.

AFLPs, which was the DNA fingerprinting technique of choice for this study, is based on selective PCR amplification of restriction fragments generated by specific restriction enzymes. This technique involves specific double-stranded DNA biotinylated adapters to be ligated to the DNA restriction fragments (Vos *et al.*, 1995), so that the sequences of adapters and the adjacent restriction sites serve as primer-binding sites. The primers are designed to contain the sequences that are complementary to those of the adapters and the restriction sites, along with one to three

selective bases added at their 3' ends. The use of these selective bases allows amplification of only a subset of the restriction fragments, which still generate a large number of bands that facilitate the detection of polymorphisms. Polymorphisms are detected through the differences in the length of the amplified fragments via polyacrylamide gel electrophoresis (PAGE) (Matthes *et al.*, 1998). The AFLP methodology has been extensively applied in genetic diversity studies for many plants such as sunflower (Hongtrakul *et al*., 1997), Indian and Kenyan tea (Paul *et al.*, 1997), maize (Ajmone *et al*., 1998), jackfruit (Schnell *et al*., 2001), olives (Sensi *et al*., 2003) and *Aegilops* species (Sasanuma *et al*., 2004). The technique has also been applied as a useful tool in biodiversity conservation and management (Lucchini, 2003). Given both the broad spectrum that this technique covers and its versatility, we will apply AFLP molecular markers to detect genetic diversity for the three *Agathosma* taxa in question and to address the question of classification.

References

Aetna InteliHealth 2005. Pennyroyal (*Hedeoma pulegioides*), Complementary and Alternative Medicine. Natural Standard, http://www.intelihealth.com. [30/09/05].

Ajmone-Marsan, P., Castiglioni, P., Fusari, F., Kuiper, M. and Motto, M. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *Theor Appl Genet* 96:219-227.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132: 1131-1139.

Bean, A. 1993. *Agathosma crenulata*. The flora of Southern Africa. *Plate* 2075. Cape Town.

Belaj, A., Satovic, Z., Cipriani, G. and Baldoni, L. 2003. Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theor Appl Genet* 107: 736-744.

Besse, P., Da Silva, D., Bory, S., Grsopmo, M., Le Bellec, F. and Duval, M. 2004. RAPD genetic diversity in cultivated vanilla: *Vanilla planifolia*, and relationships with *V. tahitensis* and *V. pompona*. *Plant Sci* 167:379-385.

Blommaert, K.L.J. and Bartel, E. 1976. Chemotaxonomic aspects of the Buchu species *Agathosma betulina* (Pillans) and *Agathosma crenulata* (Pillans) from local plantings. *Journal of South African Botany* 42 (2): 121-126.

Chen, M., SanMiguel, P. and Bennetzen, J.L. 1998. Sequence organisation and conservation in *sh/2al* -homologous regions of sorghum and rice. *Genetics* 148: 435-443.

Cordeiro, G.M., Pan, Y.B. and Henry, R.J. 2003. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. *Plant Sci* 165:181-189.

Edwards, A., Hammond, H.A., Jin, L., Caskey, C.T. and Chakraborty, R. 1992. Genetic variation at five trimeric and tretrameric tandem repeat loci in four human population groups. *Genomics* 12: 241-253.

Endenburg, H.G. 1972. A study of the yield and composition of the essential oils from *Agathosma crenulata* and *A. betulina*. MSc-Thesis. University of Stellenbosch.

Fluck, A.A.J., Mitchell,W. and Perry, H.M. 1961. Composition of buchu leaf oil. *J. Sci Food Agric*.12: 290-292.

Fuchs, S., Sewenig, S. and Mosandl, A. 2001. Monoterpene biosynthesis in *Agathosma crenulata* (Buchu). *Flavour Fragr. J.* 16: 123-135.

Golan-Goldhirsh, A., Barazani, O., Wang, Z.S., Khadka, D.K., Saunders, J.A., Kostiukovsky, V. and Rowland, L.J. 2004. Genetic relationships among Mediterranean *Pistacia* species evaluated by RAPD and AFLP markers. *Plant Syst. Evol*. 246:9-18.

Goldblatt, P. and Manning, J. 2000. Cape plants. A conspectus of the Cape flora of South Africa *Strelitzia* 9. National Botanical Institute, Pretoria & Missouri Botanical Garden, Missouri, pp 68-73.

Gonzalez-Perez, M.A., Caujapé-Castellsm, J. and Sosam, P.A. 2004. Molecular evidence of hybridisation between the endemic *Phoenix canariensis* and the widespread *P. dactylifera* with Random Amplified Polymorphic DNA (RAPD) markers. *Plant Syst. Evol.* 247: 165-175.

Gupta, P.K., Varshney, R.K., Sharma, P.C. and Ramesh, B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding* 118: 369-390.

Hongtrakul, V., Gordan, M.H. and Knapp, S.J. 1997. Amplified fragment length polymorphism as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theor Appl Genet* 95:400-407.

Jones, E.S., Hughes, L.J., Drayton, M.C., Abberton, M.T., Michaelson-Yeates, P.T., Bowen, C. and Forster, J.W. 2003. An SSR and AFLP molecular marker-based genetic map of white clover (*Trifolium repens* L.). *Plant Sci* 165:531-539.

Karp, A. and Edwards, K.J. 1997. DNA markers: a global overview, in DNA Markers: Protocols, Applications and Overviews. (Ceatano-Anolles, G. and Gresshoff P. M. eds), Wiley, pp. 1-13.

Kaiser, R., Lamparsky, D. and Schudel, P. 1973. Analysis of buchu leaf oil. *J. Agric. Food Chem*. 23: 943-949.

Kumar, L.S. 1999. DNA markers in plant improvement: An overview. *Biotechnological Advances* 17:143-182.

Liebhard, R., Koller, B., Gianfranceschi, L. and Gessler, C. 2003. Creating a saturated map for the apple (Malus x domestica Borkh.) genome. *Theor Appl Genet* 106:1497-1508.

Lis-Balchin, M., Hart, S. and Simpson, E. 2001. Buchu (*Agathosma betulina* and *A. crenulata*, Rutaceae) essential oils: their pharmacological action on guinea-pig ileum and antimicrobial activity on microorganisms. *Journal of Pharmacy and Pharmacology* 53: 579-582.

Lucchini, V. 2003. AFLP: a useful tool for biodiversity conservation and management. C.R. Biologies 326:S43-S48.

Mazur, B.J and Tingey, S.V. 1995. Genetic mapping and introgression of genes of agronomic importance. *Current Opinion in Biotechnology* 6: 175-182.

Mueller, U.G. and Wolfenbarger, L.L. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* 14: 389-394.

Newall, C.A., Anderson, L.A. and Phillipson, J.D. 1996. Herbal Medicines: A Guide for Healthcare Professionals. London: *The Pharmaceutical Press* Vol. 51.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C. and Fuerst, P.A. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* March 1998.

Paul, S., Wachira, F.N., Powell, W. and Waugh, R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea *(Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theor Appl Genet* 94:255-263.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. and Motto, M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor Appl Genet* 97:1248-1255.

Pillans, N.S 1950. A revision of *Agathosma*. *Journal of South African Botany.* 16: 55-183.

Posthumus, M.A., van Beek, T.A., Collins, N.F. and Graven, E.H. 1996. Chemical composition of the essential oil of *Agathosma betulina*, *A. crenulata* and an *A. betulina* x *crenulata* Hybrid (Buchu). *Journal of Essential Oil Research* 8: 223-228.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2: 225-238.

Reiter, R. 2001. PCR-based markers systems, in DNA-Based Markers in Plants, Second Edition (Phillips R. L. and Vasil I. K. eds), Kluwer Academic Publishers, pp 9- 29.

Roberts, M. 1990. Indigenous Healing Plants. Southern Book Publishers, Halfway House, South Afica, pp. 42-43.

Rust, K., 2003. The beauty of buchu. *Food Review* 65: 55-56.

Sasanuma, T., Chabane, K., Endo, T.R. and Valkoun, J. 2004. Characterization of genetic variation in and phylogenetic relationships among diploid *Aegilops* species by AFLP: incongruity of chloroplast and nuclear data. *Theor Appl Genet* 108: 612-618.

Schnell, R.J., Olano, C.T., Campbell, R.J. and Brown, J.S. 2001. AFLP analysis of genetic diversity within a jackfruit germplasm collection. *Sci Hortic* 91:261-274.

Sefc, K.M, Glöβl, J. and Steinkellner, H. 2000. Broad range geotyping using microsatellite markers identified in V*itis riparia*. Proc. VII Int'l Symp. on Grapevine Genetics and Breeding. *Acta Hort*, 528.

Sensi, E., Vignani, R., Scali, M., Masi, E. and Cresti, M. 2003. DNA fingerprinting and genetic relatedness among cultivated varieties of *Olea europaea*. *Sci Hort*. 1867:1-10.

Singh, A., Chaudhury, A., Srivastava, P.S. and Lakshmikumaran, M. 2002. Comparison of AFLP and SAMPL markers for assessment of intra-population genetic variation in *Azadirachta indica* A.Juss. *Plant Science* 162: 17-25.

Simpson, D. 1998. Buchu – South Africa's amazing herbal remedy. *Scottish Medical Journal* 43: 189-191.

Spreeth, A.D. 1976. A revision of the commercially important *Agathosma* species. *Journal of South African Botany* 42(2): 109-119.

Turpie, J.K., Heydenrych, B.J. and Lamberth, S.J. 2003. Economic value of terrestrial and marine biodiversity in the Cape Floristic Region: implications for defining effective and socially optimal conservation strategies. *Biological Conservation* 112: 233-251.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M and Zabeau, M. 1995. *Nucleic Acids Research*. 23(21): 4407-4414.

Webber, L.N., Magwe, M.L. and van Staden, J. 1999. Alternative uses for some invader plants: turning liabilities into assets. *South African Journal of Science* 95: 329-331.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18:7213-7218.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531-6535.

Wu, K-s. and Tanksley, S.D. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. *Mol Gen Genet* 241:225-235.

Zhebentyayeva, T.N., Reighard, G.L., Gorina, V.M.and Abbott, A.G. 2003. Simple sequence repeat (SSR) analysis assessment of genetic variability in apricot germplasm. *Theor Appl Genet* 106:435-444.

Chapter Three

Assessing the genetic diversity of commercially important *Agathosma* species using amplified fragment length polymorphisms (AFLPs) as molecular markers

3.1. Abstract

AFLP markers were successfully employed to detect genetic differentiation among the buchu species *Agathosma betulina* (Berg.) Pillans, *A. crenulata* (L.) Pillans and a putative *A. betulina* X *A. crenulata* hybrid. Five primer combinations generated 6342 fragments of which 1186 (19%) were polymorphic for all genotypes. The data obtained were elaborated with the Dice genetic similarity coefficient, and analysed using different clustering methods and Principle Coordinate Analysis (PCoA). Cluster analysis of the genotypes revealed an overall genetic similarity between the populations of between 0.85 and 0.99. The AFLP-based dendrogram divided the populations into three major groups: (1) the *A. serratifolia* and *A. crenulata* populations, (2) the putative hybrid, *A. betulina* X *A crenulata* populations and (3) the *A. betulina* populations. Results from the AFLP-analyses using Unweighted Pair-Group Method Analysis (UPGMA) and PCoA plots were congruent, confirming the patterns of similarity displayed in the cluster analysis. The PCoA plot confirmed that the putative hybrid is not intermediately spread between the *A. crenulata* and *A. betulina* populations and that it is genetically more similar to *A. betulina.* The putative hybrid might therefore be viewed as a genetically distinct variant of *A. betulina*. To our knowledge this is the first report on the use of molecular evidence to investigate the affinities between *Agathosma* species and, more specifically, the question of hybridisation. This study sets the stage for the future accurate identification, protection, management and cultivation of this important indigenous resource.

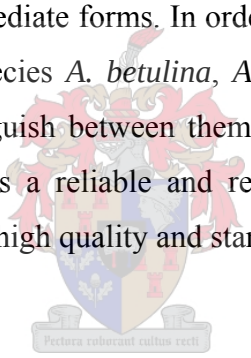
Key words
Buchu, *Agathosma betulina, Agathosma crenulata,* amplified fragment length polymorphisms, genetic diversity

3.2. Introduction

The Western Cape Province of South Africa is home to *Agathosma betulina* (Berg.) Pillans, *A. crenulata* (L.) Pillans and *A. serratifolia* (Curt.) Spreeth, a group of species collectively known as buchu (Von Willigh, 1913). The San and Khoi, two native tribes of the Western Cape of South Africa, used buchu as early as the 17th century. These native tribes also introduced buchu to the European settlers, who took it to Europe, where it found medicinal application as a diuretic (Lis-

Balchin *et al.*, 2001). Today buchu is one of three South African medicinal plants used internationally in medicine (Rust, 2003). In addition to its therapeutic qualities, the plant's oil has a distinct blackcurrant smell and flavour, which makes it an excellent ingredient in food flavourings and aromatic oils.

Due to buchu's huge national and international market potential increased exploitation of wild buchu species has been at the order of the day, jeopardising the future survival of the species in question. Cultivation of buchu was initiated to supply in the growing demand for oil. Consequently, the establishment of buchu plantings in various parts of the Western Cape led to major confusion around the true identity of many cultivated plants (Blommaert and Bartel, 1976). Alleged hybridisation has taken place over the years whenever the two species were interplanted. Seeds collected from these interspecific hybrids were used to establish new plantings. The result was that leaf shapes, which have always been the best characteristic by which to distinguish between the two species, started displaying many intermediate forms. In order to supply high quality oil, it is crucial to be able to correctly identify the species *A. betulina*, *A. crenulata* and the putative hybrid, *A. betulina* X *A. crenulata*, and to distinguish between them. The demand for essential oil, high in diosphenol and menthonethiol, requires a reliable and reproducible method of classification of buchu species to deliver essential oil of high quality and standard.



Since the discovery of the commercial value of these species, various studies have been directed at the correct classification of the two species. Linnaeus conducted phenotypic (morphology and anatomy based) studies on buchu as early as 1755 (Bean, 1993). Fluck *et al.* (1961) and Kaiser *et al.* (1975) studied buchu chemistry, and focussed on the composition of buchu oil. Blommaert and Bartel (1976) focussed on analytical assessments of buchu oil to distinguish between the taxa *A. betulina, A. crenulata* and *A. betulina* X *A. crenulata.* Posthumus *et al.* (1996) also studied the essential oil of buchu to distinguish between the three taxa *A. betulina*, *A. crenulata* and *A. betulina* X *A. crenulata*. Results of studies were, however, influenced by environmental factors (time of harvest) and experimental factors (method of distillation). Hence, the results elucidated based on such studies do not provide a true measure of classification.

DNA-based markers have been applied to, amongst others, genetic studies, variety characterisation and paternity analysis, because these markers are largely unaffected by environmental influences.

Of the various kinds of DNA-based markers characterised so far, restriction fragment length polymorphisms (RFLPs) were the first to provide means to directly detect variations present at the DNA level. RFLPs have been used to document genetic diversity in cultivated plants and their wild relatives (Tanksley *et al.*, 1989; Diers and Osborn, 1994). Although they are highly specific, performing RFLPs is rather tedious and expensive, since it requires large amounts of pure quality DNA and an expertise in handling radioactivity. Randomly amplified polymorphic DNAs (RAPDs) (Welsh and McCleland, 1990; Williams *et al*., 1990), a polymerase chain reaction (PCR)-based technique, resolved most of the technical obstacles and offered a cost-effective and easy-to-perform approach. This efficient technique obviates the need to work with radioisotopes and yields satisfactory results even with crude DNA preparations. RAPDs have therefore been extensively used in assessing genetic relationships amongst various accessions of different plant species (Chalmers *et al*., 1992; Adams *et al*., 1993; Belaj *et al*., 2003). One of the major drawbacks of RAPDs, however, is the lack of specificity and reproducibility. It has been observed that RAPD profiles are sensitive to variations in the concentrations of template DNA (Davin-Regli *et al*., 1995), Mg^{2+} ions, *Taq* polymerase and thermal cycler used. Thus the results obtained through RAPDs can be arbitrary.

Eukaryotic genomes are interspersed with tandem repeats of DNA, referred to as microsatellites or simple sequence repeats (SSRs). SSR polymorphisms have been extensively used as genetic markers in mammals (Tautz, 1989); they also occur frequently in plant genomes where they also show extensive variation in different individuals and accessions (Akkaya *et al*., 1992; Wu and Tanksley, 1993; Pejic *et al*., 1998; Cipriani *et al*., 2002). SSR loci are transferable, highly polymorphic, multiallelic PCR-based co-dominant markers that are more informative than RAPDs and RFLPs (Russel *et al*., 1997) and relatively simple to interpret (Rafalski *et al.*, 1996). These hallmarks justify the large initial effort necessary to obtain SSR markers, which entails the acquisition of sequence information (Morgante *et al*., 1998), which is expensive.

The introduction of amplified fragment length polymorphism (AFLP) as a technique for precision genotyping circumvents all the limitations of previous fingerprinting techniques mentioned above (Zabeau and Vos, 1993; Vos *et al*., 1995). The technique is highly specific, generates a high multiplex ratio and is repeatable. An added advantage is that it requires no prior knowledge of the genome being studied. AFLP methodology has thus been used to assess genetic diversity in *Lactuca* (Hill *et al*., 1996), soybean (Maughan *et al.*, 1996), *Lens* (Sharma *et al.*, 1996), sunflower

(Hongtrakul *et al.*, 1997), tea (Paul *et al.*, 1997), barley (Russel *et al.*, 1997), *Arabidopsis thaliana* ecotypes (Erschadi *et al*., 2000), apricots (Hagen *et al.*, 2001), durum wheat (Soleimani *et al.*, 2002) and olives (Belaj *et al.*, 2003). The analysis of biodiversity by means of AFLP has also occasionally been employed in rice (Zhu *et al*., 1998), hops (Hart and Seefelder, 1998), grapevine (Cervera *et al*., 1998) and maize (Peijic *et al*., 1998).

To the best of our knowledge, there has been no report on the extent of the genetic diversity prevalent in the commercially important *Agathosma* species i.e. *Agathosma betulina*, *A. crenulata* and the putative hybrid *A. betulina* X *A. crenulata*. The objective of the present study was to determine genetic variation between and within populations of commercially important *Agathosma* species using AFLP methodology.

3.3. Materials and methods

Plant material

The taxa *Agathosma betulina*, *A. crenulata* and *A. serratifolia*, as well as a putative hybrid *Agathosma betulina* X *Agathosma crenulata*, were included in this study (Table 1). A total of nine populations (representing two or more populations of each of the taxa) were sampled, including five replicate samples per population. *A. betulina* samples were collected from populations in Paarl (Floralea) (cultivated) and further north in Citrusdal (Haarwegskloof) (wild). *A. betulina* X *A. crenulata* samples were collected from two populations in Paarl (Floralea) and Paarl (Waterfalls) (cultivated) and also in Porterville (Berghof) (wild), while *A. crenulata* samples were collected from populations in Tulbagh (Bergplaas) (wild) and a southern population in Paarl (Floralea) (cultivated). The putative hybrid was also collected from the transitional zones between Paarl and Tulbagh and between Tulbagh and Porterville. *A. serratifolia* samples were collected from two populations in Swellendam, namely Merloth Nature Reserve and Grootvadersbosch Nature Reserve. These populations were selected to represent the main geographic distribution of each taxon. The identity of plant material collected from the field was confirmed through comparison with herbarium specimens from the Bolus Herbarium (BOL, University of Cape Town) and Compton Herbarium (NBG, Kirstenbosch). Leaf morphology of the samples is depicted in Fig. 2.1.

Table 3.1 Accessions of *Agathosma* taxa included in this study, with reference to the farms or Nature Reserves from which the material was sampled.

Species	Geographical location	Population Code	Sample Code
Agathosma serratifolia	Swellendam, Merloth Nature Reserve	Sm1	S1
	S 33° 59.927 E 20° 28.683	Sm2	S2
		Sm3	S3
		Sm4	S4
		Sm5	S5
	Swellendam, Grootvadersbosch Nature Reserve	Sg1	S6
		Sg2	S7
	S 33° 58.852 E 20° 49.622	Sg3	S8
		Sg4	S9
		Sg5	S10
Agathosma betulina X *Agathosma crenulata*	Porterville (Berghof)	Hport1	S11
	S 32° 57.795 E19° 09.954	Hport2	S12
		Hport3	S13
-		Hport4	S14
		Hport5	S15
	Paarl (Floralea)	Hp1	S16
	S33° 42.937 E19° 03.775	Hp2	S17
		Hp3	S18
		Hp4	S19
		Hp5	S20
	Paarl (Waterfalls)	Hwf1	S21
	S33° 42 845 E 19° 03.633	Hwf2	S22
		Hwf3	S23
		Hwf4	S24
		Hwf5	S25
Agathosma crenulata	Paarl (Floralea)	Cp1	S26
	S33° 42.917 E19° 03.764	Cp2	S27
		Cp3	S28
		Cp4	S29
		Cp5	S30
	Tulbagh (Bergplaas)	Ct1	S31
	S33° 10.388 E19° 09.954	Ct2	S32
		Ct3	S33
		Ct4	S34
		Ct5	S35
Agathosma betulina	Paarl (Floralea)	Bp1	S36
	S33° 42.958 E19° 03.836	Bp2	S37
		Bp3	S38
		Bp4	S39
		Bp5	S40
	Citrusdal (Haarwegskloof)	Bc1	S41
	S32° 22.589 E18° 52.333	Bc2	S42
		Bc3	S43
		Bc4	S44
		Bc5	S45

DNA extraction

Leaf tissue was ground in liquid nitrogen and stored at -80 °C. Genomic DNA was isolated from 1 g of freshly ground leaf material using the modified CTAB method (Murray and Thompson, 1980). Twelve milliliters of 5% CTAB extraction buffer (5% (m/v) CTAB, 100 mM Tris-HCl pH 8.0, 1.4 M NaCl, 20 mM EDTA) and 1% (v/v) β-mercaptoethanol were added to the tissue in 50 ml sterile test tubes and incubated at 65 °C for 2 hours with intermittent shaking. The samples were then incubated on ice for 10 minutes, extracted with 2 X equal volumes of chloroform: isoamyl alcohol (24:1 – v/v) to purify the DNA, and centrifuged (Sorvall RC5C PLUS) at 9000 rpm at 25 °C for 30 minutes between each extraction. The resultant supernatant was mixed with a double volume of 1% precipitation buffer (1% (m/v) CTAB, 50 mM Tris-HCl pH 8.0, 10 mM EDTA) and 1% (v/v) β-mercaptoethanol and incubated at room temperature overnight to precipitate. The DNA was spun down for 30 minutes at 9000 rpm at 25 °C. The DNA pellet was dissolved in 1 ml of 1 M CsCl. Twice the volume of ice-cold absolute ethanol was added to precipitate DNA for 30 minutes at -20 °C. DNA was spun down for 20 minutes at 9000 rpm. The pellet was then air dried and dissolved in 100 µl double-distilled water. Genomic DNA was stored at -20 °C. Approximately 100 ng of DNA per sample was used for the AFLP procedure.

AFLP analysis

The AFLP procedure was performed following the protocol of Matthes *et al.* (1998) with minor modifications. Genomic DNA (100 ng) was digested with 2.5 units of both *Mse*I and *Pst*I restriction enzymes at 37 °C for 2 hours and 1.25 µl of 10 x restriction-ligation buffer (100 mM Tris-acetate pH 7.5, 100 mM magnesium acetate and 500 mM potassium acetate) in a final volume of 11 µl. Non-phosphorylated adaptor sequences, one biotinylated and the other not, were ligated to the restriction fragments at 37 °C overnight. Ten µl of Dynabeads M280 (Dynal) (10 mg beads/ml) were used to select biotinylated DNA fragments, following the protocol of Matthes *et al.* (1998). TE wash buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 100 mM NaCl) was used to wash the beads. After the washing steps, the beads were resuspended in 25 µl of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA). This was used as template DNA. The *Mse*I primer was radioactively labelled using 0.5 µCi γ^{33}P-ATP. Selective amplification was performed with a [^{33}P]-labelled *Mse*I primer (5' -GATGAGTCCTGAGTAA+N-3') and biotinylated *Pst*I primer (5'-GACTGCGTACATGCAG+N-3') and extensions, where 'N' represents three selective nucleotides (Table 3.2).

Each 10 µl PCR reaction consisted of 1.0 µl template DNA, 50 ng labelled *Mse*I + 3 primer, 50 ng biotinylated *Pst*I + 3 primer, 10 mM dNTP, 0.05 µl Taq DNA polymerase (5 units/µl) (Bioline Taq) and 1.5 µl of 10x NH_4 Buffer (Bioline) and 0.5 µl $MgCl_2$ (50 mM) (Bioline). The following cycle profile ensured optimal primer selectivity: 1 cycle of 2 minutes at 94 °C; followed by 35 cycles of 30 seconds at 94 °C, 30 seconds at 56 °C, 30 seconds at 72 °C; followed by 1 cycle of 15 minutes at 72 °C.

Amplified products were heated at 95 °C for 6 minutes after addition of 5 µl formamide dye (98% (v/v) formamide, 10 mM EDTA pH 8.0 and 1 mg ml^{-1} each of bromophenol blue and xylene cyanol) and immediately chilled on ice. Five µl of amplified fragments were separated in 5% (m/v) denaturing polyacrylamide gels. All gels were run at 60 W for about 150 minutes. Gels were dried onto Whatman 3M paper on a slab gel dryer (Biorad Laboratories Inc., Hercules, CA, USA). Radioactively labelled DNA fragments were visualised on BioMAX MR film (Eastman Kodak Company, Rochester, New York) after exposure for 72 hours.

Data analysis

AFLP bands showing clear polymorphism were scored as present (1) or absent (0) on autorads. Thirty to forty bands per plant (only strong bands) were scored per autorad for each primer set, consisting of 45 samples, resolved on a single gel (Table 3.1). Genetic similarity (GS) between pairs was estimated according to Dice's coefficient (Sneath and Sokal, 1973). The formula is given as: $Gs_{ij} = 2a/2a + b + c$, where Gs_{ij} is the measure of GS between individuals *i* and *j*, a is the number of polymorphic fragments that are shared by *i* and *j*, b is the number of fragments present in *i* and absent in *j*, and c is the number of fragments present in *j* and absent in *i*. For phenetic tree construction, the binary character matrix was converted into a distance matrix based on the Nei and Li (1979) index of similarity (also known as the Dice index). The resulting similarity matrix thus generated was used to construct a phenetic dendrogram using the unweighted pair group method of averages (UPGMA) and the SAHN clustering analysis (Sneath and Sokal, 1973). The goodness-of-fit (Mantel test) of the clustering (Mantel, 1967) was tested using the MXCOMP program, which directly compares the original similarity matrix and the cophentic value matrix, as suggested by Rohlf (1993). The programs DCENTER and EIGEN were used to perform a principal coordinates analysis. All the statistical analyses were performed using the NTSYS-pc-software (Version 2.02k, Rohlf, 1999).

3.4. Results

The summary of all AFLP markers produced by five primer-pairs across all the genotypes is given in Table 3.2. The five primer-combinations generated 6342 fragments (Fig. 3.1 as an example) ranging from 50 to 1100 bp, of which 1186 (19%) were polymorphic over all the genotypes. The capability of different primer-pairs to generate polymorphic AFLP markers varied significantly, ranging from 190 to 284 polymorphic bands per primer-pair (Table 3.2) over all the genotypes. It thus confirms the high multiplex ratio expected from the AFLP technique.

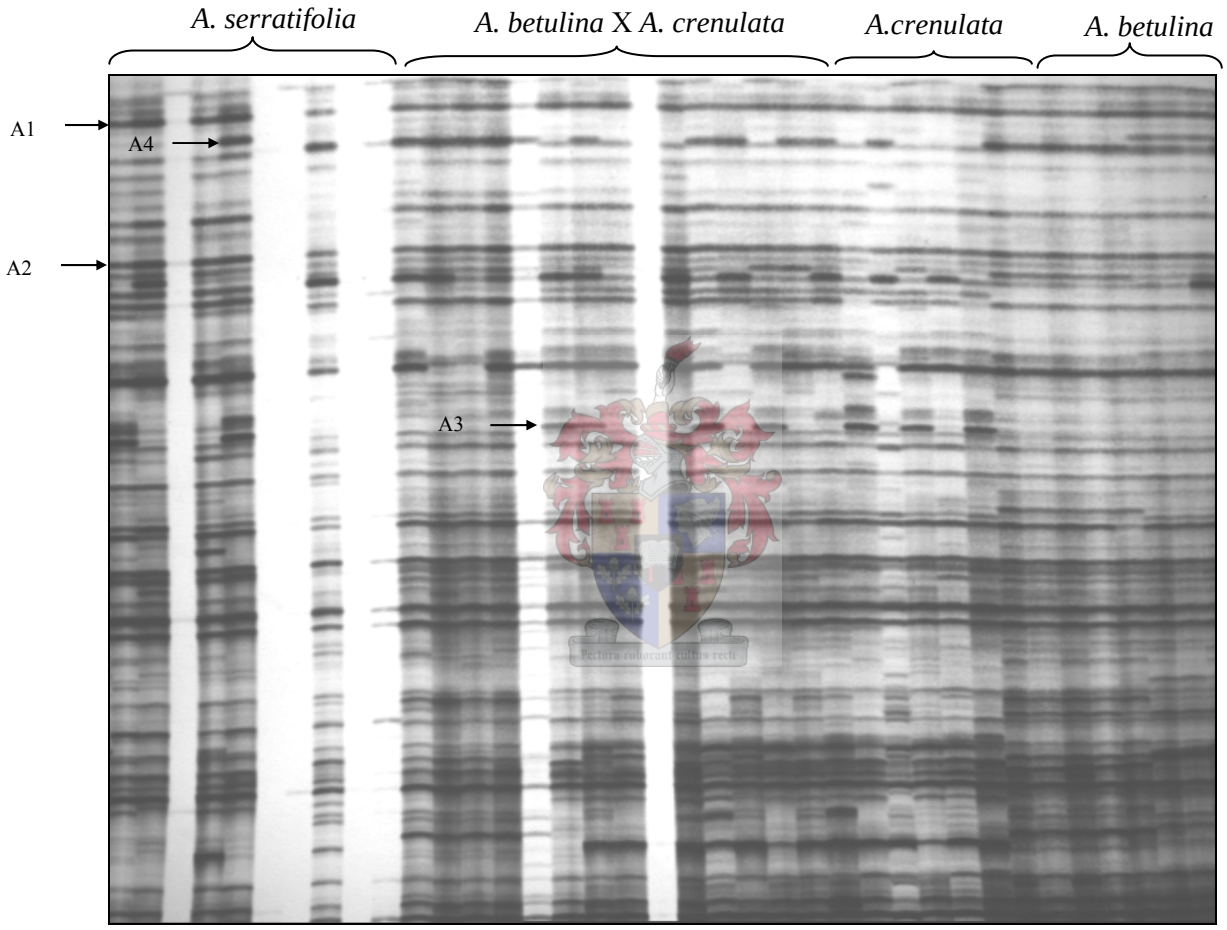


Figure 3.1. AFPL profile of *A. betulina*, *A. crenulata*, *A. serratifolia* and the putative hybrid showing monomorphic- (A1, A2) and polymorphic bands (A3, A4) with primer combination *Pst*I (GTA) + *Mse*I (CAG).

Agathosma genotypes were grouped by cluster analysis (Fig. 3.2) based on Dice's genetic similarity coefficient, with genetic similarity between populations ranging from 0.85 to 0.99. The dendrogram had a cophenetic correlation r = 0.82, which is interpreted as a good fit (Mantel, 1967). The AFLP-based dendrogram divides the populations into three major groups (Fig. 3.2): (1) the *A. serratifolia* and *A. crenulata* populations (2) the putative hybrid, *A. betulina* X *A crenulata* populations and (3) the *A. betulina* populations. A further evaluation of the relationship among the populations was

obtained through the principle coordinate analysis based on the same similarity matrix (Fig. 3.3) and these groupings were found to be congruent with the main groups identified in the dendrogram.

Table 3.2 Summary of the AFLP data generated for 45 Agathosma genotypes using five different primer combinations.

Selective amplification primer pairs	Total number of bands	Number of polymorphic bands	Average polymorphism (%)
*Pst*I – GTA + *Mse*I – CAA	1264	210	16.6
*Pst*I – GTA + *Mse*I – CAG	1261	247	19.6
*Pst*I – GTA + *Mse*I – GAG	1254	190	15.1
*Pst*I - GTA + *Mse*I – CTG	1262	284	22.5
*Pst*I – TTT + *Mse*I – CAA	1301	255	19.6
Total	6342	1186	Average 18.7 ± 2.9

Table 3.3 The average percentages of polymorphic bands generated within different populations of the different *Agathosma* species.

A. serratifolia		Putative *A. crenulata* X *A. betulina*		*A. crenulata*		*A. betulina*	
Population	% polymorphism	Population	% polymorphism	Population	% polymorphism	Population	% polymorphism
Merloth	19.7	Porterville	17.1	Paarl	21.4	Paarl	22.8
Grootvadersbosch	13.5	Paarl	18.1	Tulbagh	20.4	Citrusdal	23.6
		Waterfalls	15.1				
Average	16.6 ± 4.3		16.8 ± 1.5		20.9 ± 0.7		23.2 ± 0.6

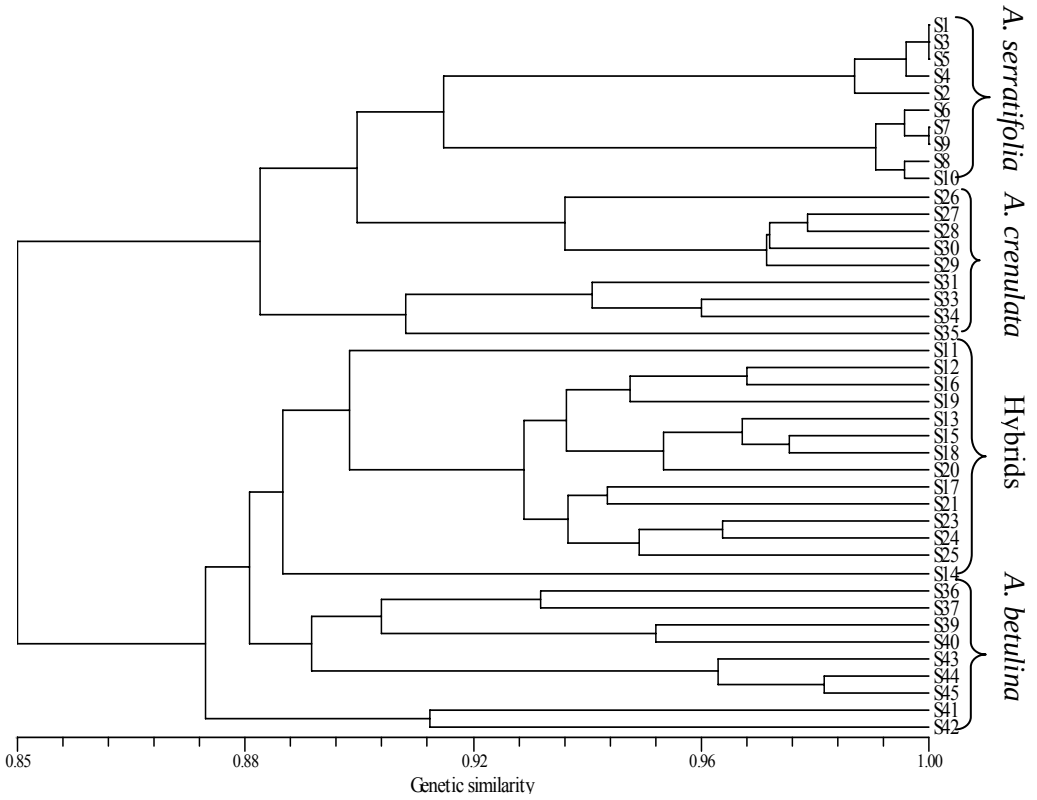


Figure 3.2. Dendrogram of the three commercially important *Agathosma* species resulting from UPGMA cluster analysis based on Dice's similarity coefficients obtained from 6342 fragments. (Cultivars included in the analysis are terminally indicated by the codes as defined in Table 3.1.).

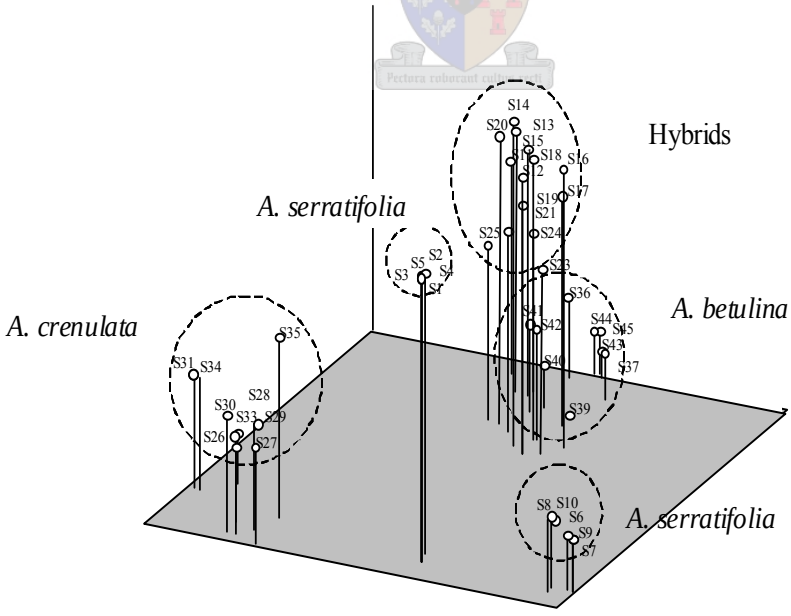


Figure 3.3. Principle coordinate analysis plot of commercially important *Agathosma* species based on 6342 AFLP fragments derived from five primer sets using Dice's similarity matrix (Cultivars included in the analysis are indicated by the codes as defined in Table 3.1.).

3.5. Discussion

In this study, a group of genotypes including both cultivated and wild varieties of the commercially important *Agathosma* species were genetically analysed. The ability of the AFLP technique to distinguish different levels of variability has been established in other taxa, such as lentil (Sharma *et al*., 1996), sunflower (Hongtrakul *et al*., 1997), bananas (Ude *et al*., 2002) and grapevines (Fanizza *et al*., 2003) and proved to be a powerful tool for *Agathosma* species as well. Among the nine populations sampled, the percentages of polymorphic bands detected by AFLPs ranged from 15.1 - 22.5% (Table 3.2). This showed that there is a high level of genetic similarity across all populations sampled (Table 3.3). Within the *A. serratifolia* populations, the genetic variation is between 13 - 19%, with *A. serratifolia* (Merloth) demonstrating the highest level of genetic variation (approximately 20% polymorphic). The *A. serratifolia* and hybrid populations display the highest genetic similarity between taxa, with approximately 17% variation (Table 3.3). Between the *A. crenulata* and *A. betulina* populations, genetic variation ranges between 20 and 23%. The putative hybrid exhibits lower levels of genetic diversity than either of the two putative parents, as measured by estimates of percentage polymorphic loci (Table 3.3).

The genetic variation found within all the populations included in this study is low (20% polymorphic) although *Agathosma* species are cross-pollinated by bee pollinators (Gold, 1992). The uniqueness of this finding is in contrast to what Sasanuma *et al.* (2004) found for cross-pollinating species of *Aegilops speltoides* (Tausch) and *Aegilops mutica* (Boiss.), where the levels of genetic variation were higher (61- and 68% polymorphic respectively) than among self-pollinating species. The high genetic similarity could be ascribed to the limited number of accessions used in this study. A larger number of accessions should be investigated in order to assess the overall genetic diversity of these species more accurately. Our aims were, however, not to assess the complete within-species variation of these taxa, but rather to compare the relative levels of genetic similarity and/or dissimilarity between different populations of different commercially important *Agathosma* species.

The first important cluster to recognise in Fig. 3.2 is the one that groups the two populations of *A. serratifolia* together as 0.915 similar. Within this cluster, *A. serratifolia* specimens form two sub-clusters that group together as being 0.985 and 0.99 similar respectively. These subclusters do not correspond with specific localities or populations, as genotypes S1, S3, S5, S7 and S9 are very similar at the molecular level despite originating from two geographically isolated localities, namely the Merloth (S 33° 59.927; E 20° 28.683) and Grootvadersbosch (S 33° 58.852; E 20° 49.622) nature reserves. The observed similarity (lack of difference) between these genotypes at the

molecular level might be ascribed to the small fraction of the genome explored, even though a large number of AFLP markers were used in this analysis.

The two *A. crenulata* populations do not resolve in a single cluster, but effectively split along population boundaries (Paarl and Tulbagh). The *A. crenulata* (Tulbagh) population forms a sister cluster to S26, an *A. crenulata* (Tulbagh) specimen from the same population. In turn, the *A. crenulata* (Tulbagh) population and S26 cluster is sister to the two *A. serratifolia* populations, at a genetic similarity (GS) level of 0.904. The other *A. crenulata* population from Paarl, along with S35 from the same population, resolves as a sister to the cluster containing the two *A. serratifolia* populations, S26 and the *A. crenulata* (Tulbagh) populations at a 0.884 genetic similarity level. This implies that the *A. crenulata* (Tulbagh) population is genetically more similar to both the *A. serratifolia* populations (0.90) than it is to the *A. crenulata* (Paarl) population. Members of the *A. serratifolia* (Grootvadersbosch) population clustered together as a subgroup with the highest similarity of GS = 0.99. The *A. serratifolia* (Merloth) population showed the second highest similarity of GS = 0.986, and also clustered together as a subgroup. The two *A. serratifolia* populations are linked at a high level of similarity of GS = 0.915 to form a well-defined cluster, as would be expected for two populations of the same species. There is, however, some genetic variation both within and among these populations. These *A. serratifolia* and *A. crenulata* populations are therefore genetically very similar to each other. Pillans (1950) lumped *A. serratifolia* and *A. crenulata* as the single species *A. crenulata*. Based on biochemical evidence, however, Spreeth (1976) proposed that the two different species *A. serratifolia* and *A. crenulata* are different enough to validate their recognition as separate species. He also substantiated this view with anatomical and morphological evidence. Our results support the close affinity between the two species, as proposed by Pillans (1950).

Fig. 3.2 also shows that the *A. betulina* and the putative hybrid populations are quite different (only 0.85 similar) from the *A. serratifolia* and *A. crenulata* populations. This is supported by the morphological distinctions between *A. crenulata* and *A. betulina* proposed by Pillans (1950). Spreeth (1976) also confirmed anatomical and morphological differences between these two species. According to Spreeth (1976), leaf morphology constitutes the main morphological difference between the two species. Essential oil analysis by Posthumus *et al.* (1996) showed that the oil constituents of the two species vary in yield. Within the *A. betulina* and the putative hybrid cluster (*A. betulina* X *A. crenulata*), the *A. betulina* genotypes S41 and S42 are distinct, and resolve as less similar (and thus basal) to both a well-defined *A. betulina* (Citrusdal) and a hybrid cluster. However, in the principle coordinate analysis plot (Fig. 3.3), the genotypes S41 and S42 group with the other *A. betulina* population. The genotypes of *A. betulina* species from the Citrusdal

populations cluster together at the 0.886 similarity level, and form a sister cluster to the hybrid populations HPort., HPaarl and HFalls.

The legitimacy of the putative hybrid (*A. betulina* X *A. crenulata*) was also investigated. It is difficult to determine from morphological characters whether the intermediacy in chemical characters (oil profile) and leaf shape (Fig. 2.1) resulted from hybridisation. Intermediate leaf shape (Fig. 2.1) and oil profile data (Blommaert and Bartel, 1976; Posthumus *et al.*, 1996) of the putative hybrid suggests hybridisation between the two parental species. However, there is neither experimental proof of such inter-specific breeding capabilities, nor any molecular evidence to support these claims. Therefore, it is difficult at this stage to know whether the intermediate individuals are real hybrids, morphological variants or ecotypes. Traditionally, putative hybrids demarcated through biosystematic data such as morphology or chemotaxonomy (e.g. leaf shape or oil profiles) were not often verified by molecular evidence. Molecular markers represent a more powerful tool for identifying hybrid taxa, but even this approach can generate ambiguous results. Molecular data would support the hypothesis of a hybrid origin of morphological intermediate individuals if any unique markers found in the putative parents were additive in the putative hybrids. Hybrid origin could be refuted if there were no combinations of the parental markers in the putative hybrids and/or many unique markers were observed in the parents and the putative hybrid.

DNA-based assessment of the putative hybrid in the present study produced ambiguous results. The putative hybrid is not intermediate to the putative parents, but is very similar to *A. betulina.* The hybrid population is genetically very well-demarcated with 0.887 similarity (Fig. 3.2). Instead of being intermediate between the two putative parental species, the putative hybrid clustered more closely with *A. betulina* populations, with a similarity level of 0.872. The reason for this could be that gene flow has occurred through repeated back-crossing between an original hybrid (*A. betulina* X *A. crenulata*) and just one parental species, *A. betulina,* through a process of introgression. The resulting segregating generations will then mostly be *A. betulina* variants and will have multi-locus associations with this species (Gonzalez-Perez *et al.*, 2004). An alternative possibility could be that the hybrid is merely an ecological variant of *A. betulina*.

A further evaluation of the relationship among the populations was obtained through the principle coordinate analysis based on the same similarity matrix. The result from the analysis of AFLP fragments using UPGMA and PCoA plot are congruent. In the PCoA plot, a different algorithm is used that allows a three-dimensional view of the data. The PCoA plot (Fig. 3.3) shows three distinct groups with the putative hybrid and *A. betulina* grouped separately, thus confirming the pattern of similarity displayed in the cluster analysis (Fig. 3.2). The PCoA plot also confirmed that the

putative hybrid is not intermediately spread between the *A. crenulata* and *A. betulina* populations. The putative hybrid is clearly closer to the *A. betulina* group, supporting the conclusion that the putative hybrid could be an ecological variant of *A. betulina*. The *A. serratifolia* populations form two separate groups, indicating that they originate from different locations. In the *A. betulina*, *A. crenulata* and the putative hybrid populations, the respective populations cluster together to form distinct groups. This could be a confirmation that genetic similarity is very high within the respective populations even though they are from different locations. Geographic distribution of the *A. serratifolia* populations is clearly reflected in the dendrogram (Fig. 3.2) and PCoA plot (Fig. 3.3), whereas that of the other populations i.e. *A. betulina*, *A. crenulata* and the putative hybrid *A. betulina* X *A. crenulata* is not so clear in the respective figures.

These results clearly show that there is genetic variation within and between populations of all the different species. By using AFLP markers, this study showed that there is a high level of genetic similarity among populations (Table 3.3). Individuals among populations have different morphological characteristics, e.g. leaf shape and oil profiles (Spreeth, 1976; Posthumus *et al.*, 1996). Anatomical studies also confirmed that there is very little variation between species (Bissiengou, 2003). This high level of genetic similarity found in this study correlates with the findings of Spreeth (1976), who identified leaf morphology as the only diagnostic difference between the different species. The high level of genetic similarity could be a result of the seed dispersal strategies of buchu species. Seeds are dispersed from the capsule upon ripening via ballistic dispersal (Gold, 1992) that does not project the seeds very far from their parental plant. Genetically related offspring thus form relatively dense stands in close proximity to their parent plant. Wu *et al.* (2004) stated that in wild rice populations, where seeds are dispersed by gravity, genetic variation is low. Hamrick and Godt (1989) also reported that the genetic variation of a population that transmit seeds by gravity is lower than that of other seed-dispersed types.

This is the first molecular study of the genetic relationships of *Agathosma* in which the three commercially important cultivars and/or populations were grouped into three distinct groups by AFLP analysis. AFLP markers implemented in this study confirm genetic diversity within and among commercially important *Agathosma* species. Furthermore, the AFLP profiles and the clusters established can be used as a basis for comparison with other *Agathosma* species or cultivars. The genetic similarity among cultivars established could further help *Agathosma* germplasm identification and preservation. It also sheds further light on the origin of the putative hybrids populations, suggesting that the putative hybrid is genetically more closely related to *A. betulina* than to any other species. Although these molecular data still do not fully explain the origin of the putative hybrid populations, the results could suggest a hybridisation event followed by

repeated back-crossing with *A. betulina*. The results from this study open new avenues for better management and cultivation programmes, and sets the stage for future research on finding molecular tools to screen large numbers of commercially important buchu cultivars.

References

Adams, R.P., Demeke, T. and Abulfaith, H.A. 1993. RAPD DNA fingerprints and terpenoids: clues to past migrations of *Juniperus* in Arabia and east Africa. *Theor Appl Genet* 87:22-26.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132: 1131-1139.

Bean, A. 1993. *Agathosma crenulata*. The flora of Southern Africa. *Plate* 2075. Cape Town.

Bissiengou, P. 2003. A morphological and anatomical comparison of selected *Agathosma betulina* and *A. crenulata* populations in the Western Cape, South Africa. Honours mini-thesis. University of Stellenbosch.

Blommaert, K.L.J. and Bartel, E. 1976. Chemotaxonomic aspects of the Buchu species *Agathosma betulina* (Pillans) and *Agathosma crenulata* (Pillans) from local plantings. *Journal of South African Botany* 42 (2): 121-126.

Chalmers, K.j., Sprent, J.I., Simons, A.J., Waugh, R. and Powell, W. 1992. Patterns of genetic diversity in tropical tree legume *(Gliricidia)* revealed by RAPD markers. *Heredity* 69:465-472.

Cervera, M.T., Cabezas, J.A., Sancha, J.C., Martinez de Toda, F. and Matinez-Zapater, J.M. 1998. Applications of AFLPs to the characterization of grapevine *Vitis vinifera* L. genetic resources. A case with accessions from Rioja (Spain). *Theor Appl Genet* 97:51-59.

Cipriani G., Marrazzo M.T., Marconi , R., Cimato, A., Testolin, R. 2002. Microsatellite markers isolated in olive are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars (*Olea europaea* L.). *Theor Appl Genet* 104:223-228.

Davin-Regli, A., Charrel, R.N., Bollet, C. and de Mico, P. 1995. Variations in DNA concentrations significantly affect the reproducibility of RAPD fingerprint patterns. *Res Microbiol* 146:561-568.

Diers, B.W. and Osborn, T.C. 1994. Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymophisms. *Theor Appl Genet* 88:662-668.

Erschadi, S., Haberer, G., Schöniger, M. and Torres-Ruiz, R.A. 2000. Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP). *Theor Appl Genet* 100:633-640.

Fanizza, G., Chaabane, R., Lamaj, F. and Ricciardi, L. 2003. AFLP analysis of genetic relationships among aromatic grapevines (*Vitis vinifera*). *Theor Appl Genet* 107:1043-1047.

Fluck, A.A.J, Mitchell,W. and Perry, H.M. 1961. Composition of buchu leaf oil. *J. Sci Food Agric*.12: 290-292.

Gold, M. 1992. The buchus. Cultivation and propagation. National Botanical Institute, Kirstenbosch, Cape Town, pp 48-50.

Gonzalez-Perez, M.A., Caujapé-Castellsm, J. and Sosam, P.A. 2004. Molecular evidence of hybridisation between the endemic *Phoenix canariensis* and the widespread *P. dactylifera* with Random Amplified Polymorphic DNA (RAPD) markers. *Plant Syst. Evol.* 247: 165-175.

Hagen, L.S., Lambert, P., Audergon, J.M. and Khadari, B. 2001. Genetic relationships between apricot (*Prunus armeniaca* L.) and related species using AFLP markers. Proc. Int. Symp. on Molecular Markers. *Acta Hort*. 546.

Hamrick, J.L. and Godt, M.J. 1989. Allozyme diversity in plant species, in: Brown A.H.D., Clegg M.T., Kahler A.l., Weir B.S. (Eds.), Plant Population Genetics, Breeding, and Genetic Resources, Sineaur, Associates Publishers, MA, USA, 1989, pp. 43-63.

Hill, M., Witsenboer, H., Zabeau, M., Vos, P., Kesseli, R. and Michelmore, R. 1996. PCR fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. *Theor Appl Genet* 93:1202-1210.

Hongtrakul, V., Gordan, M.H. and Knapp, S.J. 1997. Amplified fragment length polymorphism as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theor Appl Genet* 95:400-407.

Kaiser, R., Lamparsky, D. and Schudel, P. 1973. Analysis of buchu leaf oil. *J. Agric. Food Chem*. 23: 943-949.

Lis-Balchin, M., Hart, S. and Simpson, E. 2001. Buchu (*Agathosma betulina* and *A. crenulata*, Rutaceae) essential oils: their pharmacological action on guinea-pig ileum and antimicrobial activity on microorganisms. *Journal of Pharmacy and Pharmacology* 53: 579-582.

Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.

Matthes, M.C., Daly, A. and Edwards, K. 1998. Amplified Fragment Polymorphism (AFLP). In: Karp, A., Isaac, P. G. and Ingram D. S. (eds.) Molecular Tools for Screening Biodiversity. Chapman & Hall, London, pp 183-190.

Maughan, P.J., Saghai Maroof, M.A., Buss, G.R. and Huestis, G.M. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance and near-isogenic lines. *Theor Appl Genet* 93:392-401.

Morgante, M., Pfeiffer, A., Jurman, I., Paglia, G., and Olivieri, A.M. 1998. Isolation of microsatellite markers in plants. In: Karp, A., Isaac, P.G., Ingram, D.S. (eds) Molecular tools for screening biodiversity. Plants and animals. Chapman and Hall, London, pp 75-134.

Murray, M.G and Thompson, W.F. 1980. Rapid isolation of high-molecular weight plant DNA. *Nucleic Acids Research* 8: 4321-4325.

Nei, M. and Li, W. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76: 5269-5273.

Paul, S., Wachira, F.N., Powell, W. and Waugh, R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea *(Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theor Appl Genet* 94:255-263.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. and Motto, M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor Appl Genet* 97:1248-1255.

Pillans, N.S 1950. A revision of *Agathosma*. *Journal of South African Botany.* 16: 55-183.

Posthumus, M.A., van Beek, T.A., Collins, N.F. and Graven, E.H. 1996. Chemical composition of the essential oil of *Agathosma betulina*, *A. crenulata* and an *A. betulina* x *crenulata* Hybrid (Buchu). *Journal of Essential Oil Research* 8: 223-228.

Rafalski, J.A., Vogel, J.M., Morgante, M., Powell, W., Andre, C., and Tingey S.V. 1996. Generating and using DNA markers in plants. In: Birren, B, Lai, E. (eds) Nonmammalian genomic analysis. A pratical guide. Academic Press, San Diego, pp 75-134.

Rohlf F.J. 1993. NTSYS-pc Numerical Taxonomy and Multivariate Analysis System, Version 1.8. Exeter Publications, Setauket, NY.

Rohlf, F.J. 1999. NTSYS-pc-numerical taxonomy and multivariate analysis system, Version 2.02k. Exeter Publishing, Ltd., Setauket, New York.

Rust, K., 2003. The beauty of buchu. *Food Review* 65: 55-56.

Sasanuma, T., Chabane, K., Endo, T.R. and Valkoun, J. 2004. Characterization of genetic variation in and phylogenetic relationships among diploid *Aegilops* species by AFLP: incongruity of chloroplast and nuclear data. *Theor Appl Genet* 108: 612-618.

Sharma, S.K., Knox, M.R. and Ellis, T.H.N. 1996. AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theor Appl Genet* 93:751-758.

Sneath, P.H.A. and Sokal, R.R. 1973. Numerical Taxonomy. Freeman. San Francisco, p. 73.

Soleimani, V.D., Baum, B.R. and Johnson, D.A. 2002. AFLP and pedigree-based genetic diversity estimates in mordern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. *Theo Appl Genet* 104:350-357.

Spreeth, A.D. 1976. A revision of the commercially important *Agathosma* species. *Journal of South African Botany* 42(2): 109-119.

Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. 1989. RFLP mapping in plant breeding: new tools for an old science. *BioTechnology* 7:257-264.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17:6463-6471.

Ude, G., Pillay, M., Nwakanma, D. and Tenkouano, A. 2002. Analysis of genetic diversity and sectional relationships in *Musa* using AFLP markers. *Theor Appl Genet* 104:1239-1245.

Von Willigh, G.R. 1913. The culture of the buchu plant. *Agric. J. Un. S. Afr*. 6: 80-87.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M and Zabeau, M. 1995. *Nucleic Acids Research*. 23(21): 4407-4414.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18:7213-7218.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalsky, J.A. and Tingey, S.V. 1990. DNA polymorphisms useful as genetic markers. *Nucleic Acids Res* 18:6531-6535.

Wu, C-J., Cheng, Z-Q., Huang, X-Q. and Yin, S-H. 2004. Genetic diversity among and within populations of *Oryza granulata* from Yunnan of China revealed by RAPD and ISSR markers: implications for conservation of the endangered species. *Plant Science* 167:35-42.

Zabeau, M., and Vos, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application of 92402629 (Publ No. 0534858A1).

Zhu, J., Gale, M.D., Jackson M.T. and Bryan, G.J. 1998. AFLP markers for study of rice biodiversity. *Theor Appl Genet* 96:602-611.

Chapter Four

Identification of commercially important *Agathosma* species using high throughput molecular tools

4.1. Abstract

Agathosma betulina, *A. crenulata* and the putative *A. betulina* X *A. crenulata* hybrid are indigenous crops of the Western Cape Province, South Africa. They are widely used for their reported medicinal properties and more recently also as a food flavorant. Widespread harvesting of wild-grown shrubs has led to rapid reduction in natural populations, causing a potential threat to genetic diversity and the future survival of these taxa. Suitable genetic markers could aid both the commercialisation (improve propagation and breeding programs) and conservation of these taxa, but no such markers currently exist. Due to the superior quality of *A. betulina* oil, the development of markers is especially critical for this species. To meet this need, a reliable method was developed by using AFLP-based SNPs for the rapid screening of large numbers of commercially important *Agathosma* species. Several species-specific AFLP markers were identified during the first phase of this project, but a quick, robust marker specific for *A. betulina* could not be developed successfully. Such a marker, or set of markers, would eventually filter out all unwanted *A. crenulata* and "hybrid" specimens from the buchu oil industry. In this paper we introduce a PCR-based SNP marker that is clone specific. We propose that this SNP marker be further optimised in future studies to be species-specific.

Key words: Sequence characterised amplified regions; single nucleotide polymorphism; Buchu, *Agathosma betulina; Agathosma crenulata; Agathosma betulina* X *Agathosma crenulata*

4.2. Introduction

Buchu (*Agathosma betulina* (Berg.) Pillans, *A. crenulata* (L.) Pillans and *A. serratifolia* (Curt.) (Spreeth)), which belongs to the family Rutaceae, is distributed throughout the Western Cape Province of South Africa. These three shrubby species have considerable commercial value; they have medicinal use and have recently also been introduced in the food and flavour industries. As a medicine, buchu was introduced to the European settlers in the 17th century. They took it to Europe where it was used as a diuretic (Lis-Balchin *et al.*, 2001). It was introduced to the pharmaceutical industry as early as 1821, when Reece and Company imported round-leaf buchu (*Agathosma betulina*) from South Africa to be used as a medicine. Today buchu is one of three South African

medicinal plants used in international medicine (Rust, 2003). Buchu oil has a distinct blackcurrant smell and flavor, which makes it an excellent ingredient in food flavorings and aromatic oils (Posthumus *et al.*, 1996; Fuchs *et al.*, 2001).

Increased harvesting of wild shrubs has led to rapid thinning of natural populations and the danger of these species becoming extinct. Cultivation of buchu was initiated, since wild-growing buchu could no longer satisfy the demand. Initially three buchu species were commercially used, namely *A. betulina*, *A. crenulata* and *A. serratifolia*, but the latter gradually became less important and marketable. Only *A. betulina* and *A. crenulata* were recognized as "true" buchu, with *A. betulina* as the most sought-after species (Blommaert and Bartel, 1976). Intermediate forms of these two species, with oil of a lesser quality, also appeared in the industry. To supply the demand for "true" buchu species it became increasingly important to distinguish between the commercially important species *Agathosma betulina* (round-leaf buchu), *A. crenulata* (oval-leaf buchu) and their putative hybrid, *A. betulina* X *A. crenulata*. Morphological, anatomical, taxonomical and chemo-taxonomical studies have been conducted to distinguish between these three taxa, but to date all of these methods have failed to be reliable or reproducible.

DNA-based molecular markers have provided an opportunity for genetic characterisation, allowing direct comparison of different genetic material without the confusion imposed by environmentally induced variation (Potter *et al.*, 2002). At present, various molecular techniques are available for assessing genetic diversity in plants, i.e. amplified fragment length polymorphisms (AFLPs) (Vos *et al.*, 1995), random amplified polymorphic DNA (RAPDs), restriction fragment length polymorphisms (RFLPs) and microsatellites. There are, however, limitations to these techniques such as the presence of low numbers of polymorphic loci, their requirements for large amounts of DNA, and their poor reproducibility and labour intensity. Despite these problems, AFLPs have been used in many genetic studies of different plant species (Maughan *et al.*, 1996, Hongtrakul *et al.*, 1997, Hagen *et al.*, 2001, Belaj *et al.*, 2003). Most of the genetic studies conducted thus far have improved on the reliability and reproducibility of these methods.

The identification of molecular markers specific to *A. betulina* would be an essential step towards marker-assisted selection. Although the AFLP technique is powerful in terms of the identification of markers closely linked to genes of interest, it has some disadvantages for use in marker-assisted

selection. Limitations to the large scale, locus specific application of AFLPs includes their dominant type of inheritance, labour intensity and the high costs of this method. The conversion of AFLP markers into sequence-specific PCR products is required if large breeding populations are to be screened at low cost. The conversion of AFLP markers into PCR-based markers has been accomplished for several crop plants, including carrots (Bradeen and Simon, 1998), barley and wheat (Shan *et al.*, 1999), brassicas (Negi *et al.*, 2000), soybean (Meksem *et al.*, 2001) and apple (Xu *et al*. 2001). However, Shan *et al.* (1999) found the conversion of AFLP markers to be more difficult than the conversion of other markers types such as RFLPs, RAPDs and SSRs due to loss of sequence specificity after amplification of the AFLP-derived internal primer.

Marker conversion requires the characterisation of the linked marker sequences and the design of locus-specific primers (Dussle *et al.*, 2002). The success rate of converting AFLP fragments to SCAR markers in other crop plants has been very low (Dussle *et al.*, 2002). To circumvent this shortcoming, AFLPs that are derived through the AFLP technique (Vos *et al*., 1995) include SNPs, insertions and deletions (indels) and microsatellites (Meksem, 2001). Single nucleotide polymorphisms (SNPs) have been found to be the most abundant resource of genetic variation among individuals of a species (Nicod and Largiader, 2003). SNPs can be defined as a substitution, insertion or deletion at a single base position on a DNA strand (Pacey-Miller and Henry, 2003). SNPs represent the most abundant class of genetic variation found in eukaryotic genomes (Garg *et al.*, 1999; Batley *et al.*, 2003) and are usually identified by sequencing and comparing homologous gDNA fragments (Nicod and Largiader, 2003). SNP markers derived from genome sequencing have been used for saturated map development (Froelich *et al*., 1999), for candidate gene identification (Cargill *et al*., 1999) and for high-throughput-screening (Germer *et al*., 2000). They have mostly been used in human genome studies, especially in human disease gene mapping (Zhou *et al.*, 2001; Nicod and Largiader, 2003). SNP markers also have large potential in evolutionary studies. The number of SNP studies in plants have to date been limited (Pacey-Miller and Henry, 2003), despite the great potential of these techniques to enable screening of large numbers of specimens (Meksem *et al.*, 2001).

The objective of the present study was therefore to sequence selected polymorphic AFLP bands in order to convert these markers into sequence characterised amplified regions (SCARs) that can be used to identify the commercially important *Agathosma* species. We report here on the conversion

of AFLPs to SCARs, and failing to achieve this, resort to characterising the SNPs present in the AFLPs and to develop SNP markers.

4.3. Materials and Method

Cloning and sequencing of AFLP fragments

Polymorphic AFLP fragments amongst the three *Agathosma* species were selected, cut from the polyacrylamide gel, eluted and re-amplified (Chapter 3). DNA fragments were recovered from each band after heat treatment at 95°C in 30 µl distilled water for 10 minutes. The fragments were re-amplified in a 30 µl standard PCR reaction with the appropriate selective AFLP primer combination and PCR conditions. The resulting PCR products were analysed on ethidium bromide-stained 2% (m/v) agarose gels, and purified from the agarose gels using QIAquick Gel Extraction Kit (250) (Qiagen) following the manufacturer's instructions. Quantification of samples was done using the DyNA Quant 200 (Hoefer) fluorometer and cloned into pGEM® T-Easy vector, according the manufacturer's recommendations (Promega Corporation, Madison, USA), and then transformed into the DHα5 competent cells. Selected DNA clones were sequenced (ABI PRISM™ dye terminator cycle sequencing) using a reaction kit with AmpliTaq® DNA polymerase (The Perkin Elmer Corporation, Norwalk, USA) and submitted to the University of Stellenbosch's Central Analytical Facility for sequencing.

Sequence characterised amplified regions (SCARs) design and analysis

Eight polymorphic and three monomorphic AFLP fragments were cloned and sequenced. The DNA sequences were edited to discard the vector/linker and primer sequences. Primers designed were A1 (forward (frw) and reverse (rev)) for *A. betulina*, F2 (frw and rev) for *A. crenulata*, I1 (frw and rev) for the putative hybrid and J2 (frw and rev) for the internal standard and/or positive control. Three monomorphic fragments (one for *A. betulina*, one for *A. crenulata* and one for the putative hybrid) were pooled to serve as an internal standard. The designed primers were synthesised by Integrated DNA Technologies, Inc. and used in the final PCR to identify the species. Primer sequences are listed as follows:

For *A. betulina*:

A1 (frw), 5'- CTA GAG AGA GAA GTC CAG AGC TTG -3';

A1 (rev), 5'- GAA AGA TTA GGT TAT GTT CAG TCA TTG -3';

For *A. crenulata*:

F2 (frw), 5'- ACA ATT CTC AGA CTG TAC CGT -3';

F2 (rev), 5'- TGA CAA AAT AAT CTA TAA ACT ATG AAA AAA AAT TAC -3';

For the putative hybrid:

I1 (frw), 5'- TGC GGA AAG GAG ACT CAA CCC -3';

I1 (rev), 5'- GAA GTG GGC AGC TTT TAA GGT -3';

As internal standard:

J2 (frw), 5'- CGA GTT TGA ACA AAT TAG TGC TTC -3';

J2 (rev), 5'- AAC AAC CTC AAC CAT GTG AAC -3'

Target sequences were amplified in a 25 µl reaction volume containing 100 ng gDNA, dNTPs (10 mM each of dATP, dCTP, dGTP and dTTP), 0.5 µl Taq DNA polymerase (5 units/µl) (Bioline Taq) and 2.5 µl of 10x NH_4 Buffer (Bioline) and 0.75 µl $MgCl_2$ (50 mM) (Bioline), 0.5 µl of forward and reverse primer (10 µM) and distilled water. The following cycle profile ensured optimal amplification: 1 cycle of 3 minutes at 94 °C; followed by 35 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C, 30 seconds at 72 °C; followed by 1 cycle of 5 minutes at 72 °C. Bands were visualised on ethidium bromide-stained 2% (m/v) agarose gel.

Characterisation of SNPs

To clone the genomic sequences immediately upstream of the *Pst*I and *Mse*I restriction site, long-range inverse PCR amplifications (LR-iPCR) were conducted. Based on the sequence of the AFLP fragments, inverse primer pairs were designed at the 5' and 3' ends of the sequence, extending in an "outward" direction to allow LR-iPCR. The oligonucleotide primers used to amplify the upstream and downstream regions are listed as follows:

A1 - (Upstream (Upstr.)), 5'- TCA AGC TCT GGA CTT CTC TCT C –3';

A1 - (Downstream (Downstr.)), 5'- AAC AAA GAG AAC AAT GAC TGA ACA TA –3';

F2- (Upstr), 5'- TAC AGT CTG AGA ATT GTT GAA CGA AT –3',

F2 - (Downstr), 5'- GCG AGT GCT ATG TGT AAT TTT TTT TCA TAG –3';

I1 - (Upstr), 5'- GCT CTG GGT TGA GTC TCC TT –3';

I1 - (Downstr), 5'- CCT TAA AAG CTG CCC ACT TCT C –3'

Inqaba Biotechnologies, South Africa, synthesised all the primers used in this study.

Genomic DNA was restriction digested at 37 °C overnight with a 6 bp cutter restriction enzyme (*BamH*I or *Hind*III) that does not cut in the selected DNA fragment from the AFLP gel. A restriction enzyme map analysis was done to select for these specific enzymes (DNAsis software). DNA fragments were ligated at 16 °C overnight, and the ligated DNA was precipitated at –20 °C and spun down at 13000 rpm for 20 minutes at 4 °C. Target sequences were amplified in a 25 µl reaction volume containing 150 ng ligated DNA, 0.5 µl dNTPs (10 mM), 0.5 µl Taq DNA polymerase (5 units/µl) (Bioline Taq), 2.5 µl of 10x NH_4 Buffer (Bioline), 0.75 µl $MgCl_2$ (50 mM) (Bioline), 0.5 µl of upstream and downstream primer (10 µM) and distilled water. The following cycle profile ensured optimal amplification: 1 cycle of 3 minutes at 94 °C, followed by 35 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C, 1 minute at 72 °C, followed by 1 cycle of 5 minutes at 72 °C. Inverse PCR bands (4kb) were cut from ethidium bromide-stained 0.8% (m/v) agarose gel and cleaned using QIAquick Gel Extraction Kit (250) (Qiagen) following the manufacturer's instructions. Quantification of samples was done using a DyNA Quant 200 (Hoefer) fluorometer.

Selective cloning

The single-band inverse PCR products were converted to blunt-ends in a 50 µl volume containing 5 µl of 10x DNA polymerase buffer, T4 DNA polymerase (9 units/µl) (Promega), 1 µl dNTPs (10mM), 30 µl template DNA (ranging between 20-30 ng/µl) and dH_2O. The reaction mixture was incubated for 10 minutes at 37 °C and then heat inactivated for 10 minutes at 75 °C. Blunt-end fragments were restriction enzyme digested with *Hind*III and *Sac*I respectively to create cohesive ends, in order to perform directed cloning. Restriction digestion was confirmed on an ethidium bromide-stained 0.8% (m/v) agarose gel and the large and small bands (insert) were cut from the gel

and cleaned using QIAquick Gel Extraction Kit (250) (Qiagen) following the manufacturer's instructions. Quantification of samples was done using a DyNA Quant 200 (Hoefer) fluorometer.

A vector, pBluescript SK (pBS), was restriction enzyme digested in a reaction volume of 20 µl containing 2 µl Multicore Buffer (Promega), 3 µg pBS (700 ng/µl), restriction enzymes (*Sma*I, *Hind*III and *Sac*I) (Promega) and distilled water. First, the vector was restriction enzyme digested with *Sma*I for 6 hours at 25 °C to create blunt ends and then with *Hind*III and *Sac*I overnight at 37 °C respectively to create cohesive ends. The prepared vector was loaded on an ethidium bromide-stained 0.8% (m/v) agarose gel; bands were cut from the gel and cleaned using QIAquick Gel Extraction Kit (250) (Qiagen) following the manufacturer's instructions. Quantification of samples was done using the DyNA Quant 200 (Hoefer) fluorometer. The respective large and small inserts were ligated at 16 °C overnight into the prepared vector in a reaction volume of 25 µl containing 2.5 µl of 10x T4 DNA ligase buffer, 1 µl of T4 DNA ligase (3 units/µl), vector (pBS) (50 ng/µl) and the inserts (1:3 ratio). The ligated product was transformed into the DHα5 competent cells. Colony PCR with T7 and T3 primers confirmed the presence of the large and small inserts in the pBluescript vector. Selected DNA clones were sequenced (ABI PRISM^{TM} dye terminator cycle sequencing) using a reaction kit with AmpliTaq® DNA polymerase (The Perkin Elmer Corporation, Norwalk, USA). The DNA sequences were edited to discard the vector and primer sequences.

Sequences were aligned using DNAsis software package. Primers were designed at the single nucleotide polymorphism (SNP): (1) an artificial mismatched base at the SNP at the third position from the 3'- end of the specific primer (SNPrev1) was introduced to destabilize the primer and (2) a primer set (SNPrev2 and I1frw) that would allow restriction enzyme digestion at the SNP specific for *Hinc*II. The amplified product of (SNPrev2 and I1frw) primers would be 510 bp and would cut only with *Hinc*II that is specific only for the *A. crenulata* and the putative hybrid. Target sequences were amplified in a 25 µl reaction volume containing 100 ng gDNA, 1 µl of dNTPs (10 mM), 0.5 µl Super Therm Taq DNA polymerase (5 units/µl) (Super Therm Taq) 2.5 µl 10x NH_4 Buffer (Super Therm), 1.5 µl $MgCl_2$ (25 mM) (Super Therm), 0.5 µl I1frw (10 µM) and 0.5 µl SNPrev1 (10 µM) and 0.5 µl SNPrev2 (10 µM) primer and distilled water. The following cycle profile ensured optimal amplification: 1 cycle of 2 minutes at 94 °C, followed by 35 cycles of 30 seconds at 94 °C, 15 seconds at 65 °C, 30 seconds plus 1 second after every cycle at 72 °C; followed by 1 cycle of 5 minutes at 72°C.

4.4. Results and Discussion

In the previous chapter (Chapter 3), AFLP analysis revealed genetic variation amongst the commercially important *Agathosma* species. However, the AFLP technique cannot be directly applied for large-scale screening or locus-specific research efforts, such as map-based gene cloning and marker-assisted breeding, due to its high cost and complicated technology. Therefore, the conversion of AFLP markers into other, more applicable types of markers, such as PCR-based markers, is indispensable to expand the usefulness of the AFLP technique. A high throughput screening system that involves a rapid and non-radioactive PCR screening system is needed to routinely screen large numbers of plants. Sequence-characterised amplified region markers, based on polymorphic AFLP fragments, were therefore designed to establish a robust, PCR-based method for distinguishing between the commercially important species.

Eight polymorphic AFLP fragments (Fig. 4.1 as an example), i.e. two from *A. betulina*, two from *A. crenulata*, three from the putative hybrid and three monomorphic fragments were excised, re-amplified and re-checked in 2% agarose gel (Fig. 4.2). The re-amplified AFLP fragments that were the same size as on the polyacrylamide gels were cut from the agarose gel and than sequenced for further analysis and primer design.

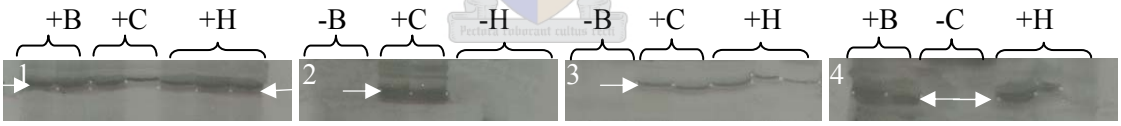


Figure 4.1 Original AFLP fragments selected from polyacrylamide gels. Fragment 1 was used as internal standard and polymorphic fragments 2-4 were used for the development of putative species-specific markers. *A. betulina* (B), *A. crenulata* (C) and the putative Hybrid (H).

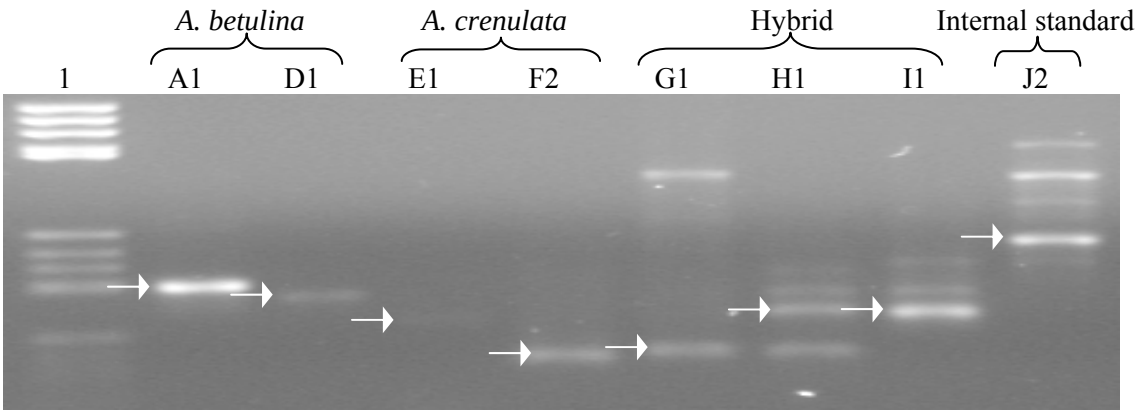


Figure 4.2 Agarose gel electrophoresis of re-amplified bands of selected AFLP fragments as identified in Fig. 1. Lanes*:* (1) DNA molecular marker V. (A1-J2) Re-amplified AFLP fragments. Fragments corresponding to the size of the original AFLP fragments are indicated by arrows and were cut from the gel, cloned and sequenced. (J2) Internal standard.

During the cloning and sequencing of AFLP fragments, eight AFLP fragments were successfully cloned and sequenced. From the eight sequences only four primer sets were designed from sequences that were long enough. The primers were used to amplify the specific polymorphic sequences from gDNA. The putative species-specific primers based on the sequences of polymorphic fragments A1, F2, I1 and J2 were not able to distinguish between the three species (Fig. 4.3). A possible reason for this could be insufficient sequence variability in the primer target sequences between the different species. Although a point mutation in the restriction sites of the *Mse*I or *Pst*I restriction enzymes could lead to polymorphisms, it will not be sufficient to influence primer binding. Moreover, during the preparation of the genomic DNA for the generation of AFLP profiles, gDNA is restriction enzyme digested with *Mse*I and *Pst*I restriction enzymes, which cut the gDNA at different sites in the DNA (Vos *et al.*, 1995). Consequently, point mutations in the DNA may alter the restriction sites for the specified restriction enzymes, resulting in polymorphisms. In other words, the polymorphic band could be the result of a small mutation not queried by the primers, i.e. the designed primer sets are targeting homologous sequences. This prompted the characterisation of the enzyme restriction sites at the ends of the polymorphic fragments.

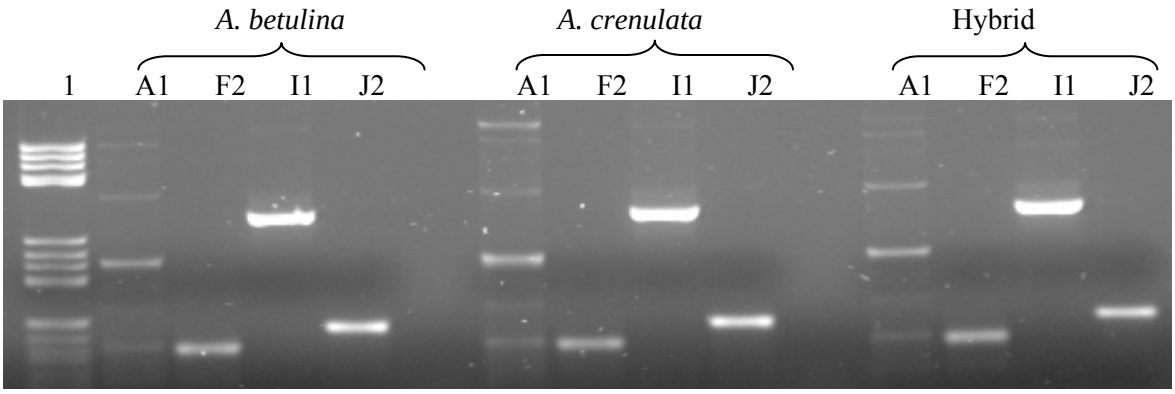


Figure 4.3 Agarose gel electrophoresis of the amplification products of various primer sets designed to distinguish between the three *Agathosma* species i.e. *A. betulina*, *A. crenulata* and the putative Hybrid. Lanes: (1) DNA molecular marker V. Lanes (A1, F2, I1, J2) PCR products of primer sets A1 (frw and rev), F2 (frw and rev), I1 (frw and rev) and J2 (frw and rev) for *A. betulina*, *A. crenulata* and the putative hybrid gDNA respectively.

The long-range inverse PCR (LR-iPCR) method was used to characterise the restriction sites flanking the polymorphic bands. Based on the sequences of the selected AFLP fragments, inverse primer pairs were designed at the 5' and 3' ends of the sequence. Ligated gDNA prepared from pooled gDNA of *A. betulina*, *A. crenulata* and the putative hybrid were used as template in the inverse PCR. All four primer sets (A1, F2, I1 and J2) were used to amplify the target sequence from *A. betulina*, *A. crenulata* and the hybrid respectively. Only the I1 (upstream and downstream) primer set was able to amplify the targeted fragment from the ligated gDNA of the three species. Digestion of the amplified PCR product with *BamH*I or *Hind*III restriction enzymes confirmed the product to be a true inverse PCR product. The resulting fragments amplified from the three different species' ligated gDNA were cloned and sequenced. The sequence data obtained revealed no point mutations in the restriction site for *Mse*I and *Pst*I restriction enzymes. A possible reason for the absence of point mutations could be that the point mutations are clone specific. In addition, the conversion of AFLP fragments to SCARs is problematic, since the extracts of the fragment from the polyacrylamide gel often contain multiple fragments, which are the result of co-isolation of background amplification products of the AFLP fragment of interest. The resulting DNA sequence is often too short to optimally design PCR primers, and too short to expect internal polymorphisms, which can be used to differentiate between alleles (Brugmans *et al.*, 2003).

Since no point mutations were found in the enzyme restriction sites, sequence data obtained with the I1 (upstream and downstream) primer set were aligned to determine if there were any polymorphisms which could be used to distinguish between the species. In the alignment of the inverse PCR sequences obtained from the I1 upstream primer, no single nucleotide polymorphisms

were identified. However, in the inverse PCR sequences obtained from the I1 downstream primer, a single nucleotide polymorphism was identified in *A. betulina* (Fig. 4.4). This data revealed no difference in the sequences between the *A. crenulata* and the hybrid for this PCR product.

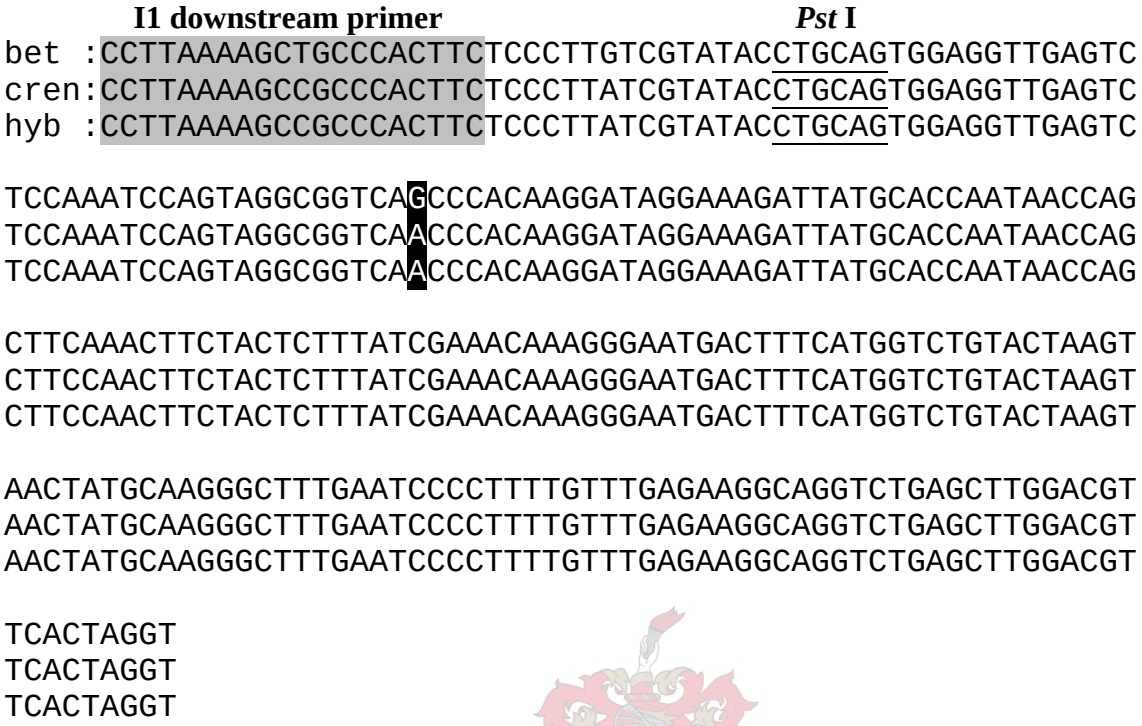


Figure 4.4 Alignment of downstream inverse PCR sequences of the three *Agathosma* species showing the I1 downstream primer (shaded), *Pst*I restriction site and single nucleotide polymorphism. *A. betulina* (bet), *A. crenulata* (cren) and the putative Hybrid (hyb).

Based on this polymorphism a locus-specific primer (SNPrev1 and I1frw) was designed (Table 4.1) to distinguish between species. Furthermore, to improve the specificity of this primer, a deliberate artificial mismatched base was introduced at the third base near the 3'- terminal region of the SNPrev1 primer (Zhou *et al.*, 2001). Amplification with this primer set also amplified fragments for *A. crenulata* and the hybrid (Fig. 4.5).

Table 4.1 Sequence of the putative *A. betulina*-specific primer set used in distinguishing between the *Agathosma* species. (The SNP is underlined and the artificial mismatched base is indicated in bold.).

Primer	Sequence
SNPrev1	3′- CG**C** GTG TTC CTA TCC TTT CTA ATA CGT – 5′
I1frw	5′-TGC GGA AAG GAG ACT CAA CCC – 3′

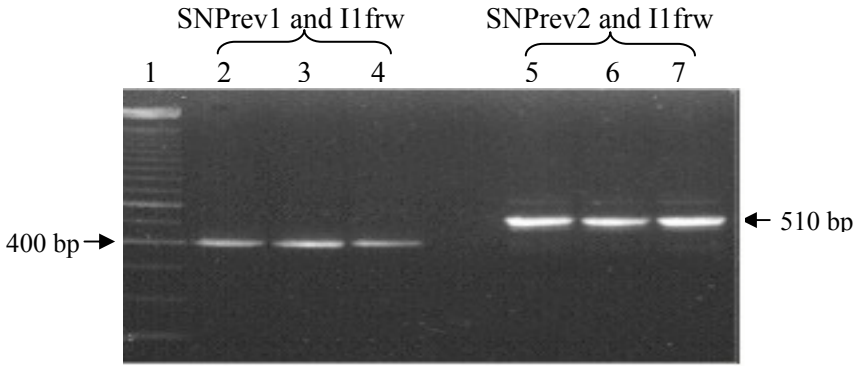


Figure 4.5 Agarose gel electrophoresis of the amplification products of the two primer sets (SNPrev1 and I1frw) and (SNPrev2 and I1frw) amplified from gDNA of the three *Agathosma* species *A. betulina*, *A. crenulata* and the putative hybrid. Lanes: (1) 100 bp DNA Molecular Marker, (2) *A. betulina;* (3) *A. crenulata;* (4) *A. betulina* X *A. crenulata* amplified with (SNPrev1 and Ifrw) (400 bp product); (5) *A. betulina;* (6) *A. crenulata;* (7) *A. betulina* X *A. crenulata* amplified with (SNPrev2 and Ifrw) yielding a 510 bp product.

In an attempt to increase their specificity, the hybridisation conditions (different annealing temperatures) were optimised for the SNPrev1 and Ifrw primer sets. Generally, high temperatures are preferred, as they prevent mismatch primer extension (Zhou *et al.*, 2001). The *A. betulina* band and the bands for the other two species, *A. crenulata* and the hybrid, was also amplified. Amplification at annealing temperatures of 62 and 65 °C produced two fragments of 400 and 360 bp for the hybrid (Fig. 6-A). When the annealing temperature was increased to 67 °C for 15 seconds, the 400 bp fragment disappeared (Fig. 4.6-B).

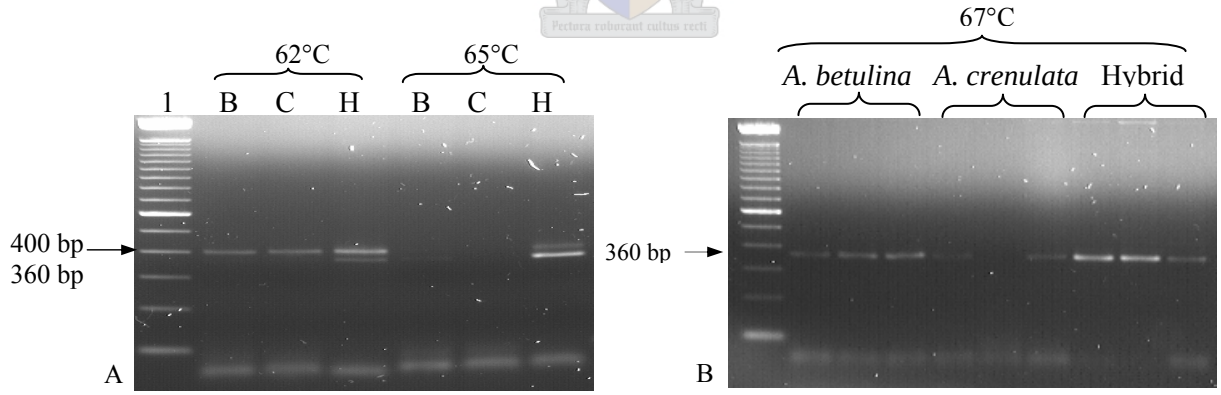


Figure 4.6 Agarose gel electrophoresis of PCR products for optimising the hybridisation conditions of the primer set (SNPrev1 and Ifrw) at different temperatures. Lane: (1) 100 bp DNA Molecular Marker. gDNA of the three species (*A. betulina* - B, *A. crenulata* - C and the putative Hybrid – H) amplified with the primer set showing two bands 400 bp and 360 bp respectively (Fig. 6 A). gDNA of the three species amplified with primers showing that the 400 bp band had disappeared (Fig. 6 B).

Since optimising of the PCR conditions at different temperatures did not result in the expected band being amplified, optimising the hybridisation of primers by varying the Mg^{+2} concentrations was investigated. The stability and therefore specificity of the primer:target hybrid is known to be very sensitive to magnesium ion concentrations (Simsek and Adnan, 2000). With the SNPrev1 and Ifrw

primer set only one band (360 bp) was detected. The PCR products were amplified and with increasing Mg^{+2} ion concentration gradient the specificity of the primer between species was increased. However, the amplification at different annealing temperatures and with different $MgCl_2$ concentrations did not produce the desired specific band for *A. betulina* as the corresponding alleles in the other species were also amplified. Only the hybrid alleles were amplified more effectively (Fig. 4.7-C). A possible reason why this SNP could not be used for the specific amplification of an *A. betulina* fragment could be that the original SNP was erroneously identified due to an amplification or sequencing error. Alternatively, the designed primers might just not be able to discriminate between the two different alleles.

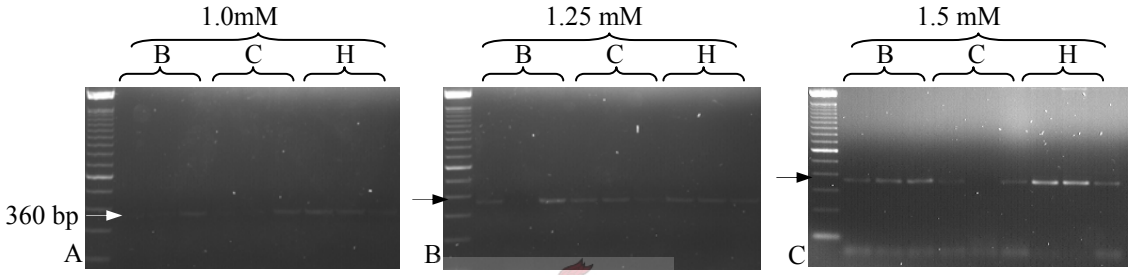


Figure 4.7 Agarose gel electrophoresis of PCR product for the optimisation of primer set (SNPrev1 and Ifrw) with different $MgCl_2$ concentrations. Only one band is detectable (360 bp, indicated by the arrow) throughout the concentration gradient in the three *Agathosma* species, i.e. (*A. betulina* - B, *A. crenulata* - C and the putative Hybrid – H).

Since amplification with SNPrev1 and Ifrw primer set under various conditions could not amplify the *A. betulina* band specifically, a primer set was designed so that its amplification product would include an enzyme restriction site. Digestion of this amplification product would reveal differences in the species. This was achieved when the amplification products of the primer set (SNPrev2 and Ifrw) of the different species were sequenced to characterise any other sequence variations that could be utilised to differentiate between the species, e.g. using restriction enzymes. This primer set, SNPrev2 and I1frw, was designed in such a way that the resulting amplified product would include a polymorphic *Hinc*II restriction site. The *Hinc*II site (Fig. 4.8) was present in *A. crenulata* and the "hybrid", but was absent in *A. betulina*, and should therefore be able to distinguish *A. betulina* from the other taxa. Amplification of gDNA from three varieties of the three *Agathosma* taxa with SNPrev2 and I1frw produced the expected PCR fragments (510 bp) (Fig. 4.5). Unfortunately, although the restriction enzyme did cut the fragments as expected, yielding two fragments of 370 and 140 bp respectively, it also cut the *A. betulina* fragment (Fig. 4.9). The reason for this could be that the enzyme restriction site GTTAAC in the *A. betulina* sequence (Fig. 4.8) would allow for *Hinc*II to recognise this restriction site. The enzyme restriction site for *Hinc*II is GTY/RAC in which the Y and R bases could be any combination of C/A, C/G or T/A, T/G.

Ifrw →

*TGCGGAAAGGAGACTCAACCC*AGAGCTTAGTTGCCGAGCACATCGTATTCTAAAACCTACAAAAACCTAGCCT
GGTTACGAACTCTCTCTTGAAAAAAATCCCTTTACGAGCTTGAATTGGTATCCCTTCTGAAACGGATGAAATT
GATTTCAACCTCTACAGTTATGTCTCCTGCTGAAAGGTTCTCCCAAGTACCCTGGAAAAGTGATAGATGGTTT
TGCTAAATAAGCATCGGAATTGACAGATGCAAGCCACCTGAACTAAGATGGATTTATAGATGTAGCCCATTTG
AACCTTAAAAGCTGCCCACTTCTCCCTTATCGTATACCTGCAGTGGAGGTTGAGTCTCCAAATCCAGTAGGCG
GTCAACCCACAAGGATAGGAAAGATTATGCACCAATAACCAGCTTCCAACTTCTACTCTTTATCGAAACAAAG
GGAATGACTTTCATGGTCTGTACTAAGTAACTATGCAAGGGCTTTGAAT***CCCCTTTTGTTTGAGAAGGCAGG***

← SNPrev2

Figure 4.8 Sequence of PCR product for *A. betulina* (510 bp) of SNPrev2 and Ifrw primer set with the restriction site for *Hinc*II (indicated in bold) used to differentiate between *Agathosma* species.

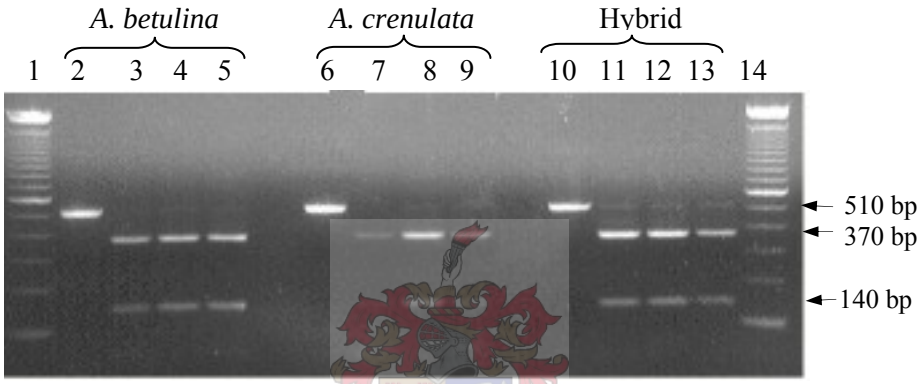


Figure 4.9 Agarose gel electrophoresis of the undigested and digested amplification products of the SNPrev2 and Ifrw primer set from gDNA of three varieties of the three *Agathosma* species. Lanes: (1 & 14) 100 bp DNA Molecular Marker, (2, 6, 10) undigested PCR product. Amplification products cut with restriction enzyme *Hinc*II that should have been specific for *A. crenulata* (7-9) and the putative Hybrid (11-13) also cut *A. betulina* (3-5).

In this study, AFLP markers specific to *A. betulina*, *A. crenulata* and the putative hybrid were identified. These species-specific markers were converted to SCARs, which proved to be non-specific and SNPs were therefore ultimately characterised. The reason for this lack of specificity found with the SCARs could be that in the conversion of AFLP fragments to SCARs, extracts of the fragment from the polyacrylamide gel often contain multiple fragments, which are the result of co-isolation of background amplification products of the AFLP fragment of interest. Also, with the conversion of AFLPs to SCARs, non-specific co-migrating DNA fragments may be cloned and sequenced. It is also found that when the number of sequenced clones is low or when the cloning efficiency of a particular AFLP fragment is low, co-isolated fragments may outnumber the sequence of the correct fragment and ultimately hamper the determination of the right sequence (Brugmans *et al.*, 2003). Negi *et al.* (2000) also experienced difficulties in their study of *Brassica* to convert short AFLP fragments in to SCAR markers. Dussle *et al.* (2002) also found a lack of sequence specificity

in their study of the sugarcane mosaic virus (SCMV) in maize, and attributed this to the shortness of the AFLP fragment (150-300 bp). It is easier to convert longer, randomly amplified polymorphic DNA (RAPD) fragments (500-1500 bp) to a SCAR (Barret *et al.*, 1998) than AFLP markers, since the DNA sequence of the AFLP markers is often too short to optimally design PCR primers, and too short to expect internal polymorphisms that can be used to differentiate between alleles (Brugmans *et al.*, 2003).

SNPs were thus characterised and it was found that locus-specific designed primers could not amplify species-specific bands. Theoretically, the locus-specific primer (SNPrev1 and I1frw) should have amplified a band specific only for *A. betulina*. The introduction of the artificially mismatched base at the third base from the 3'- terminus of the primer should have prevented the mismatch primer extension reaction of the *A. crenulata* and the putative hybrid samples, ultimately increasing the specificity of the primer. The sensitivity of this technique also depends on the DNA polymerase as well as the annealing temperatures used in the extension step (Zhou *et al.*, 2001). Finally, with the optimisation of the primers under different PCR conditions, i.e. different temperatures and Mg^{+2} ion concentrations, no differences were observed in the fragments amplified, other than that of the putative hybrid varieties amplified more effectively. This phenomenon could also be ascribed to the quality of the gDNA, such as PCR inhibitors present that might inhibit the amplification of the desired product.

The digestion of the PCR product of the second set of primers, SNPrev2 and I1frw, with *HincII,* should have been specific only for *A. crenulata* and the hybrid. However, the digested amplification products are rather an indication of the presence of a *Hinc*II enzyme restriction site. The reason for this might be that the SNP that was characterised was only specific to the particular clone and not species specific. Alternatively, the single nucleotide polymorphism identified between *A. betulina* and the hybrid could simply be a false single nucleotide polymorphism. This might explain why none of the SNPs that were characterised were able to differentiate between the three different species using either locus-specific amplification or restriction enzyme digestion. In addition, the *Hinc*II restriction enzyme that cuts the PCR products of the amplified samples have a GTY/RAC recognition sequence, where the Y and R bases could be any combination of C/A, C/G or T/A, T/G. Thus, although the SNP might be genuine, the restriction enzyme was not able to discriminate between the two alternative sequences.

In conclusion, although several species-specific AFLP markers were identified during the first phase of the project, a quick, robust marker specific for *A. betulina* has not yet been generated. Future studies might involve the exploitation of RAPDs, which would be ideal for the conversion of these RAPDs to SCARs. Since RAPDs are notorious to form artifactual amplification products (Parker *et al*., 1998), great care should be taken in order to ensure that a reproducible fragment is being sequenced for ultimate development of SCAR markers. In the development of RAPDs, careful control of DNA quality and amplification conditions is necessary to ensure reproducible banding patterns (Parker *et al*., 1998). Through this technique a robust marker specific for *A. betulina* is feasible. A marker or set of markers specific for *A. betulina* would be extremely useful, since the elevated oil quality of this species makes it the most sought after buchu species for cultivation on a large scale. Such a marker would be advantages to the farmer as well as the distillers knowing that the seedlings received and/or the essential oil processed is pure *A. betulina*. Application of such a marker or set of markers will eventually filter out all the unwanted putative hybrid and *A. crenulata* species from the buchu oil industry.

References

Barret, P., Delourme, R., Foisset, N. and Renard, M. 1998. Development of a SCAR (sequence characterized amplified region) marker for molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus* L. *Theor Appl Genet* 97: 828-833.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K.J. and Edwards, D. 2003. Mining for single nucleotide polymorphisms and insertions–deletions in maize expressed sequence tag data. *Plant Physiol*. 132: 84–91.

Belaj, A., Satovic, Z., Cipriani, G. and Baldoni, L. 2003. Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theor Appl Genet* 107: 736-744.

Blommaert, K.L.J. and Bartel, E. 1976. Chemotaxonomic aspects of the Buchu species *Agathosma betulina* (Pillans) and *Agathosma crenulata* (Pillans) from local plantings. *Journal of South African Botany* 42 (2): 121-126.

Bradeen, J.M. and Simon, P.W. 1998. Conversion of an AFLP fragment linked to carrot Y_2 locus to a simple, codominant PCR-based marker form. *Theor Appl Genet* 97: 960-967.

Brugmans, B., Van der Hulst, R.G.M., Visser R.G.F., Lindhout P and van Eck, H. J. 2003. A new and versatile method for the successful conversion of AFLP markers into simple single locus markers Nucleic Acids Research 31,10 e55.

Cargill, M., Altshuler, D., Ireland, J. Sklar, P. Ardlie, K., Patil, N. Shaw, N., Lane, C.R., Lim, E.P., Kalyanaraman, N., Nemesh, J. Ziaugra, L., Friedland, L., Rolfe, A. Warrington, J., Lipshutz, R., Daley, D.Q., and Lander, E.S. 1999. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet* 22:231-238.

Dussle, C.M., Quint, M., Xu, M.L. and Melchinger, A.E. 2002. Conversion of AFLP fragments tightly linked to SCMV resistant genes Scmv1 and Scmv2 into simple PCR-based markers. *Theor Appl Genet* 105: 1190-1195.

Froelich, S., Houlden, H., Rizzu, P., Chakraverty, S., Baker, M. Kwon, J., Nowotny, P., Isaacs, A., Nowotny, V., Wauters, E., van Baren, M.J. Oostra, B.A., Hardy, J. Lannfelt, L., Goate, A., Hutton, M., Lendon, C.L. and Heutink, P. 1999. Construction of a detailed physical and transcript map of the FTDP-17 candidate region on chromosone 17q21. *Genomics* 60:129-136.

Fuchs, S., Sewenig, S. and Mosandl, A. 2001. Monoterpene biosynthesis in *Agathosma crenulata* (Buchu). *Flavour Fragr. J.* 16: 123-135.

Garg, K., Green, P. and Nickerson, D.A. 1999. Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags, *Gen. Res.* 9: 1087–1092.

Germer, S., Holland, M.J. and Higuchi, R. 2000. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome Res* 10:258-66.

Hagen, L.S., Lambert, P., Audergon, J.M. and Khadari, B. 2001. Genetic relationships between apricot (*Prunus armeniaca* L.) and related species using AFLP markers. Proc. Int. Symp. on Molecular Markers. *Acta Hort*. 546.

Hongtrakul, V., Gordan, M.H. and Knapp, S.J. 1997. Amplified fragment length polymorphism as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theor Appl Genet* 95:400-407.

Lis-Balchin, M., Hart, S. and Simpson, E. 2001. Buchu (*Agathosma betulina* and *A. crenulata*, Rutaceae) essential oils: their pharmacological action on guinea-pig ileum and antimicrobial activity on microorganisms. *Journal of Pharmacy and Pharmacology* 53: 579-582.

Maughan, P.J., Saghai Maroof, M.A., Buss, G.R. and Huestis, G.M. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance and near-isogenic lines. *Theor Appl Genet* 93:392-401.

Meksem, K., Ruben, E., Hyten, D., Triwitayakorn, K. and Lightfoot, D.A. 2001. Conversion of AFLP bands into high-thoughput DNA markers. *Mol. Gen Genom* 265: 207-214.

Negi, M.S., Devic, M., Delseny, M. and Lakshmikumaran, M. 2000. Identification of AFLP fragments linked to seed coat color in *Brassica juncea* and conversion to SCAR marker for rapid selection. *Theor Appl Genet* 101: 146-152.

Nicod, J-C. and Largiader, C.R. 2003. SNPs by AFLP (SBA): a rapid SNP isolation strategy for non-model organisms. *Nucleic Acids Research* Vol. 31 No.5.

Pacey-Miller, T. and Henry, R., 2003. Single-nucleotide polymorphism detection in plants using a single-stranded pyrosequencing protocol with a universal biotinylated primer. *Analytical Biochemistry* 317: 165-170.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C. and Fuerst, P.A. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* March 1998.

Posthumus, M.A., van Beek, T.A., Collins, N.F. and Graven, E.H. 1996. Chemical composition of the essential oil of *Agathosma betulina*, *A. crenulata* and an *A. betulina* x *crenulata* Hybrid (Buchu). *Journal of Essential Oil Research* 8: 223-228.

Potter, D., Gao, F., Aiello, G., Leslie, C. and McGranahan, G. 2002. Intersimple sequence repeat markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia*) cultivars. *Journal of American Society for Horticultural Science* 127: 75-81.

Rust, K. 2003. The beauty of buchu. *Food Review* 65: 55-56.

Shan, X., Blake, T.K. and Talbert, L.E. 1999. Conversion of AFLP markers to sequence-specificPCR markers in barley and wheat. *Theor Appl Genet* 98: 1072-1078.

Simsek, M. and Adnan, H. 2000. Effect of single mismatches at 3'- end of primers on polymerase chain reaction. *SQU Journal for Scientific Research*, Medical Sciences.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M and Zabeau, M. 1995. *Nucleic Acids Research*. 23(21): 4407-4414.

Xu, M.L., Huaracha, E. and Korban, S.S. 2001. Development of sequence-characterized amplified regions (SCARs) from amplified fragment length polymorphisms (AFLP) markers tightly linked to the Vf gene in apple. Genome 44:63-70.

Zhou, G-h., Masao, K., Kazunori, O., Gao, C., Kunio, H. and Hideki, K. 2001.Quantitaive detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER), *Nucleic Acids Research* 29, e19.

Chapter Five

General Discussion and Conclusion

The major goal of this study was to deal with the question of differentiating between the commercially important *Agathosma* species, i.e. *A. betulina, A. crenulata* and *A. betulina* X *A. crenulata* in the buchu industry. The need for a reliable and reproducible method of classification cannot be overemphasised, given that historic and even present-day methods of classification are often problematic and inaccurate. Morphological, phenotypic and chemo-taxonomic studies performed by Spreeth (1976), Endenburg (1972), Blommaert and Bartel (1976) and Posthumus *et al.* (1996) aimed at improving the methods of classification of this commercially important species, had shortcomings of being subject to both environmental and experimental factors and, these methods remain inefficient. If we could accurately distinguish between these species, industry would benefit tremendously in their selection of pure *A. betulina* plant material, which has a superior oil quality. The market driven demand and the subsequent pressure on farmers to match this with an adequate supply forced the buchu industry to become regulated in terms of product quality. The product that goes out has to meet international standards, hence the need for a more reliable and reproducible analytical method.

DNA-based molecular marker techniques such as AFLPs, RFLPs, RAPDs and SSRs proved to be reliable and reproducible in studies of other plants. In this study, AFLPs were established successfully for application in the genetic characterisation of the commercially important *Agathosma* species. More than 6300 polymorphic and monomorphic fragments were generated using five primer pairs, of which 19% were polymorphic over all the genotypes (Table 3.2). Subsequent statistical analyses using the NTSYS software package enabled the quantification of the genetic variation between the different *Agathosma* species and between the individuals within a specific species. The level of genetic polymorphisms found between the *Agathosma* species was very low for cross-pollinating species, ranging between 16% and 23% with generalist bee pollinators (Fig. 2.2) (Gold, 1992). This finding was unique and in direct contrast to what Sasanuma *et al.* (2004) observed for cross-pollinating species of *Aegilops speltoides* and *Aegilops mutica*. In their studies, they reported higher levels of genetic variation among self-pollinating species than their cross-pollinating counterparts. This low genetic variation could be a result of the seed dispersal strategies in buchu species. In *Agathosma* species, seeds are dispersed from the capsule upon ripening via ballistic dispersal (Gold, 1992). This dispersal method does not project the seeds very far from their parental plant, thus causing genetically related offspring to form relatively dense

stands in close proximity to their parents. In this way then, cross-pollination favours very high levels of inbreeding, hence the low genetic variability. The aim of this study was, however, not to assess the complete within-species variation of these taxa, but rather to compare the relative levels of genetic similarity and/or dissimilarity by using AFLPs between different populations of different commercially important *Agathosma* species.

From the results of AFLP analysis it was also apparent that the development of a DNA fingerprint for each individual buchu plant is achievable and that the data generated clearly groups these three species into three distinct clusters, i.e. *A. betulina, A. crenulata* and the putative *A. betulina* X *A. crenulata* hybrid. The AFLP technique can thus be successfully utilised in the differentiation of the *Agathosma* species, which is a major contribution to the global competitiveness of the buchu industry. In addition, the status and genetic composition of the putative hybrid between *A. betulina* and *A. crenulata* was clarified. This putative hybrid's position on the dendogram generated from results of the AFLP analysis (Figs. 3.2 and 3.3) suggests that it is not genetically intermediate between the two supposed parental species as should be expected of hybrids. This therefore suggests that the putative hybrid is most likely only a genetic or ecological variant of *A. betulina*. This is quite a remarkable finding, since it has never before been reported in literature. Thus the AFLP profiles generated for each species can be used to distinguish between the commercially important *Agathosma* species.

However, the success of this technique is limited and it cannot be directly applied on a large-scale. Its applicability in large-scale screens or locus-specific research efforts on a commercial level is highly compromised by the inherent high costs and complications associated with the technology. This necessitates the urgent need to convert these AFLP markers into more robust and easy to run PCR-based markers. This will not only expand the usefulness of the AFLPs based on monetary costs, but will also reduce the processing time to a full and accurate result.

To this end a high throughput screening system involving a rapid and non-radioactive PCR screening system is needed to routinely screen large numbers of plants. Sequence-characterised amplified region markers based on polymorphic AFLP fragments were designed. This facilitated the establishment of a robust PCR-based method that can distinguish between the commercially important species. AFLP markers specific to *A. betulina*, *A. crenulata* and the putative hybrid were

identified and converted to SCARs. These, however, lacked the level of specificity (Fig. 4.3) required and ultimately led to the characterisation of SNPs. The conversion of polymorphic AFLP fragments to SCARs proved to be problematic, since the DNA fragments extracted from polyacrylamide gels often contained multiple fragments. This problem is caused largely by the co-isolation of background amplification products of the AFLP fragment of interest. Secondly extracted DNA sequences were often too short either for the optimal design of PCR primers or for harbouring internal polymorphisms that could be used to differentiate between alleles (Brugmans *et al.*, 2003). These problems were also experienced in other crop varieties in which this conversion was attempted (Negi *et al.*, 2000; Dussle *et al.*, 2002).

The use of SNP-based markers was also compromised by the occurrence of their own set of problems. They showed a low level of sensitivity that was inadequate to distinguish between the *A. betulina* and *A. crenulata* and/or the putative hybrid populations (Figs. 4.6 and 4.7). However, it is suspected that the original SNP that was characterised could either have been an artefact or was only specific to the particular clone from which it was obtained. Finally, the non-species specificity of this SNP could also be attributed to possible errors brought about by *Taq* in the PCR amplification or possibly sequencing errors (Fig. 4.9).

Although several species-specific AFLP markers were identified during the first phase of the project, a quick and robust marker that is specific for *A. betulina* has not yet been generated. This remains a top priority in future studies, since this species is the most sought after because of its superior oil composition. There are a number of ways in which this could be achieved; one of which would be to test other larger fragments for their potential to develop a robust marker specific for *A. betulina*. Alternatively, RAPDs should be exploited further in order to develop SCARs from the polymorphic fragments given that they produce larger fragments that can be quite easily converted (Dussle *et al.*, 2002).

References

Blommaert, K.L.J. and Bartel, E. 1976. Chemotaxonomic aspects of the Buchu species *Agathosma betulina* (Pillans) and *Agathosma crenulata* (Pillans) from local plantings. *Journal of South African Botany* 42 (2): 121-126.

Brugmans, B., Van der Hulst, R.G.M., Visser, R.G.F., Lindhout, P. and van Eck, H.J. 2003. A new and versatile method for the successful conversion of AFLP markers into simple single locus markers Nucleic Acids Research 31,10 e55.

Dussle, C.M., Quint, M., Xu, M.L. and Melchinger, A.E. 2002. Conversion of AFLP fragments tightly linked to SCMV resistant genes Scmv1 and Scmv2 into simple PCR-based markers. *Theor Appl Genet* 105: 1190-1195.

Endenburg, H.G. 1972. A study of the yield and composition of the essential oils from *Agathosma crenulata* and *A. betulina*. MSc-Thesis. University of Stellenbosch.

Gold, M. 1992. The buchus. Cultivation and propagation. National Botanical Institute, Kirstenbosch, Cape Town, pp 48-50.

Negi, M.S., Devic, M., Delseny, M. and Lakshmikumaran, M. 2000. Identification of AFLP fragments linked to seed coat color in *Brassica juncea* and conversion to SCAR marker for rapid selection. *Theor Appl Genet* 101: 146-152.

Posthumus, M.A., van Beek, T.A., Collins, N.F. and Graven, E.H. 1996. Chemical composition of the essential oil of *Agathosma betulina*, *A. crenulata* and an *A. betulina* x *crenulata* Hybrid (Buchu). *Journal of Essential Oil Research* 8: 223-228.

Sasanuma, T., Chabane, K., Endo, T.R. and Valkoun, J. 2004. Characterization of genetic variation in and phylogenetic relationships among diploid *Aegilops* species by AFLP: incongruity of chloroplast and nuclear data. *Theor Appl Genet* 108: 612-618.

Spreeth, A.D. 1976. A revision of the commercially important *Agathosma* species. *Journal of South African Botany* 42(2): 109-119.