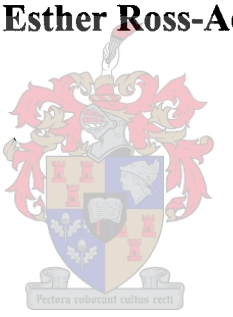
THE CHARACTERISATION OF SELECTED GRAPEVINE CULTIVARS USING MICROSATELLITES

by

Helen Esther Ross-Adams



Thesis presented in partial fulfilment of the requirements for the degree of Master of Science at the University of Stellenbosch

December 2002

Supervisor: Dr JT Burger

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 university for a degree.

Signature: **Date:**

Summary

Grapevine supports one of the oldest industries in South Africa today, and is also of significant international importance. With increasing international trade and the transport of fruit and other grapevine-derived products between borders, it has become increasingly important for South African farmers and viticulturalists to ensure their products conform to strict international market requirements if they are to remain competitive. Such requirements include the correct and accurate identification of berries and wines according to cultivar. In light of this, 26 different wine, table grape and rootstock cultivars, as well as a number of clones from KWV's core germplasm collection were characterised at 16 microsatellite marker loci. Microsatellite markers are known for their high level of informativeness, reliability and reproducibility, and are widely used in the identification and characterisation of plant varieties, population analyses and forensic applications. Unique allelic profiles were obtained for all but two plants, which proved to be identical at all loci considered, and thus 'clones'. These profiles were collated to form a database, containing the DNA fingerprints of each sample at each locus. The relative levels of informativeness of each marker used were also determined, and compared with those found in the literature. Six markers proved to be highly informative, and are promising in the potential application of this technology to other cultivars. The applicability of microsatellite markers to such studies is confirmed; this approach could easily be extended to include any number of cultivars of national and international interest. The results of such an investigation would have important implications for both the farming and commercial industries alike.

Opsomming

Wingerd ondersteun een van die oudste industrieë in Suid-Afrika vandag, en is ook van groot internasionale belang. Met die toenemende internasionale ruilhandel en die vervoer van vrugte en ander wingerd produkte tussen grense, het dit toenemend belangrik geword vir Suid-Afrikaanse wingerdboere om te verseker dat hulle produkte voldoen aan die streng vereistes van die internasional mark, indien hulle kompeterend wil bly. Hierdie vereistes sluit in die korrekte en akkurate identifisering van druiwe en wyn volgens kultivar. Met hierdie vereistes in ag geneem, is 26 verskillende wyn, tafeldruif en wortelstok kultivars, asook 'n aantal klone van die KWV se kern kiemplasma versameling, gekarakteriseer by 16 mikrosatelliet merker loki. Mikrosatelliet merkers word gekenmerk deur 'n hoë vlak van informatiwiteit, betroubaarheid en herhaalbaarheid en word wydverspreid gebruik in die identifisering en karakterisering van plant variëteite, populasie analises en forensiese toepassings. Unieke alleliese profiele is vir al die plante verkry, behalwe vir twee plante wat identiese resultate by alle loki opgelewer het en dus as "klone" beskou kan word. Hierdie profiele is bymekaar gevoeg om 'n databasis te vorm wat die DNA vingerafdrukke van elke monster by elke lokus bevat. Die relatiewe vlak van informatiwiteit van al die merkers is ook bepaal en vergelyk met merkers in die literatuur. Ses van die merkers blyk om hoogs informatief te wees en lyk belowend in die potensiële toepassing van hierdie tegnologie op ander kultivars. Die toepaslikheid van mikrosatelliet merkers op sulke studies is bevestig; hierdie benadering kan maklik aangepas word om enige aantal kultivars van nasionale en internasionale belang in te sluit. Die resultate van só 'n ondersoek sal belangrike implikasies inhou vir beide die boerdery en kommersiële industrieë.

Acknowledgements

My sincerest thanks to the following:

- **Dr Johan Burger, for his patience, guidance and most importantly, his infectious enthusiasm for his work.**
- **THRIP, Winetech and Stellenbosch 2000, for their financial support.**
- **KWV (Pty) Ltd, for providing plant material.**
- **My nearest and dearest, for their love and support.**
- **My friends, for making all the good times great.**

I may not have gone where I intended to go, but I think I have ended up where I intended to be

Douglas Adams, 1952 - 2001

Abbreviations

%	percent
°C	degrees centigrade
µCi	microcurie
µg	microgram
µl	microlitre
µM	micromolar
^{33}P	γ-^{33}P-ATP
A	adenosine
ADP	adenosine diphosphate
AFLP	amplified fragment length polymorphism
$AgNO_3$	silver nitrate
α	alpha
ATP	adenosine triphosphate
bp	base pair(s)
C	cytosine
cm	centimetre
CTAB	N-cetyl-N, N, N-trimethyl ammonium bromide
cv.	cultivar
ddH_2O	double distilled water
dH_2O	distilled water
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate(s)
dsDNA	double-stranded DNA
e.g.	exempli gratia (for example)
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
EtBr	ethidium bromide
F-	forward
g	gram

G	guanine
g's	gravitational force
GAA	glacial acetic acid
γ	gamma
gDNA	genomic DNA
Ha	hectares
HCl	hydrochloric acid
hr	hour(s)
i.e.	id est (that is)
IAM	Infinite Allele Model
Isopropanol	isopropyl alcohol
ISSR	inter-simple sequence repeat
IWBT	Institute of Wine Biotechnology, University of Stellenbosch
K	potassium
kb	kilobase
KCl	potassium chloride
KWV	Kooperatiewe Wynbouers Vereeniging (Wine makers Cooperative)
M	molar
Mg	magnesium
mg	milligram
$MgCl_2$	magnesium chloride
min	minute(s)
ml	millilitre
mm	millimetre
mM	millimolar
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
ng	nanogram
NH4 buffer	ammonium buffer
NH_4OAc	ammonium acetate
nM	nanomolar
nt	nucleotide(s)

OIV	Office International de la Vigne et u Vin (International Office of the Vine and Wine)
PCR	polymerase chain reaction
PEG	polyethylene glycol
pg	picogrammes
pH	hydrogen ion potential
PI	Probability of Identity
PNK	polynucleotide kinase
PVP	polyvinylpyrrolidone
QTL	Quantitative Trait Loci
R-	reverse
RAPD	random amplified polymorphic DNA
RE	restriction enzyme
RFLP	restriction fragment length polymorphism
RNase A	ribonuclease A
SCAR	sequence characterised amplified region
SDS	sodium dodecyl sulphate
sec	second(s)
SMM	Stepwise Mutation Model
SPARs	single primer amplification reactions
SSCP	single strand conformation polymorphism
ssDNA	single-strand DNA
SSM	Slipped strand mispairing
SSR	simple sequence repeat
STMS	sequence tagged microsatellite site
STS	sequence tagged site
T	thymine
T_A	Annealing temperature
Taq	*Thermus aquaticus*
TBE	89mM Tris-borate and 2.5mM EDTA, pH 8.0
TEMED	N,N,N',N'-tetramethylethylenediamine
T_M	Melting temperature
TPM	Two Phase Model
Tris	2-amino-2- (hydroxymethyl)-1,3-propanediol

U	unit(s)
UCO	Unequal crossing over
USA	United States of America
UV	ultra violet
V	volts
V. riparia	*Vitis riparia*
V. vinifera	*Vitis vinifera*
v/v	volume:volume ratio
VMC	Vitis Microsatellite Consortium
W	watts
w/v	weight:volume ratio

Contents

Chapter 2 – Materials and Methods

Chapter 3 – Results and Discussion

Chapter 1

Introduction

1.1 Biology of the Grapevine

The genus *Vitis* falls into the family Vitaceae, and is itself split into two sub-genera, namely *Muscadinia* and *Euvitis*. The *Euvitis* sub-genus is the only one cultivated in Europe (Classification of vine varieties..., 1993), and contains the small number of agronomically important cultivars in use today (Bourquin *et al.*, 1995).

The Vitaceae comprises 12 genera containing ±700 species of tropical or sub-tropical woody vines (Olmo, 1976). One, *Vitis,* has been estimated to contain 60 species, but knowledge of this genus is acknowledged as being incomplete (Olmo, 1976). One species native to the Far East and Northwest India, *Vitis vinifera* (the common grapevine), has proved highly valuable as a crop plant (Perold, 1927; Cangelosi, 2001).

The genus *Vitis* is unique in the Vitaceae family in that it has 38 very small chromosomes (n=19) – among the smallest eukaryotic chromosomes (Böhm and Zyprian, 2000), which made the karyotyping of *Vitis* very difficult in the past (Haas, 2000). Most related genera, including *Muscadinia* (which has only 3 known species, and is only found naturally in the southern parts of North America), have 2n=40 (Olmo, 1976). The basic chromosome number of *Vitis* is x = 13 (Patil and Patil, 1992), and it is estimated to have a genome size of approximately 475Mb (the same as rice), of which 96% is non-coding (Lodhi and Reisch, 1995).

Vitis species are dioecious out-crossers, and therefore display a high degree of heterozygosity, with the result that deleterious recessive traits have accumulated in the genome. A certain high level of heterozygosity is now a prerequisite for the successful breeding of hardy types – inbreeding depression can be so severe as to result in sterility within 3 generations (Guerra and Meredith, 1995; Passos *et al.*, 1999, Olmo, 1976). Heterozygosity is known to confer hybrid vigour (Lamboy and Alpha, 1998), which results in higher fruit production, larger plant size and faster root growth. Such highly

heterozygous plants (hardier types with better agricultural characteristics) have consistently been selected for in the process of domesticating the wild vine, and this type of selection pressure (selecting *against* homozygosity) is still exerted today in traditional breeding practices. This, then, explains the continued use of vegetatively propagated material in the form of grafted stocks (which is expensive and laborious) – essentially, the 'winning formula' of a good genome is retained.

1.2 History of the Grapevine

Grapevine is one of the oldest perennial crops, and has adapted to different climates around the world, although cultivation is largely in areas with a Mediterranean climate of hot, dry summers and cool, wet winters (Hinrichsen *et al.*, 2000).

Domestication of the vine started with nomadic groups preserving forest trees that supported especially fruitful vines near the watering holes their herds used. As these roaming tribes settled into agricultural communities and the forests were cleared, vines on the spared boundary-line trees survived, and were eventually incorporated into the village settlement (Olmo, 1976).

The wine grape had been domesticated in the Near East by 4000 BC, and had spread west to Greece and beyond by 1000 BC. Christianity, in the hands of the Romans, played a great part in the spread of the vine. Vineyards were established and maintained by monks, and wine features prominently in the consecration of the Catholic mass (Olmo, 1976; Vidal *et al.*, 1999). The *V. vinifera* grape eventually accompanied the Spanish and Portuguese on their voyages of discovery, supported again by the Roman Catholic church, and in this way had spread to the New World by 1570.

1.2.1 Rootstocks

Vines were initially vegetatively propagated using cuttings, and so remained relatively free of pests and diseases until the advent of *Phylloxera.* A mere root louse, *Phylloxera vastatrix* represented the greatest crisis in the history of European viticulture during the last decades of the 19th century, destroying nearly two thirds of the continent's vineyards (Pongrácz, 1983). The insect is indigenous to the eastern USA, and was exported to the continent between 1854 and 1860 on resistant vines.

The name is derived from the Greek *phyllon* (leaf) and *xeros* (dry). Symptoms of infestation include the stunted summer growth of leaves, which become a dull green and drop early. The root systems are also affected by nodosities and tuberosities, and are eventually unable to form new roots (Pongrácz, 1983). The vine dies when the roots are so affected as to be unable to absorb nutrients from the soil (Perold, 1927).

It was this assault on the industry that prompted the investigation into crossing apparently resistant American stocks, which had not been adversely affected by this insect, with susceptible European varieties. These resistant stocks were found to include *Vitis riparia, V. rupestris* and *V. berlandieri,* (Olmo, 1976) which are still commonly used as rootstocks for commercial *V. vinifera* cultivars, resulting in hybrids. These hybrid plants are created by grafting a susceptible European variety onto the root system of a resistant American vine (Pongrácz 1983; Perold, 1927).

Several aspects of rootstocks must be considered when developing grafts for a vineyard. For instance, their tolerance of saline soils, uptake of nutrients and minerals, resistance to bacterial and fungal pests and ability to flourish in compact or sodden soils (Pongrácz, 1983). An example is the *V. berlandieri x V. riparia* hybrid, which is one of the most important rootstocks today and combines resistance to *Phylloxera* with lime tolerance (Guerra and Meredith, 1995).

1.2.2 Grapevine Diversity

This gradual spread of grapevine cultivation together with the spread of civilisation across Europe from the East was probably a major contributing factor in the diversity of *Vitis vinifera* genotypes. There were many opportunities for crossings between domesticated varieties and wild varieties, and crossings between different types of domesticated varieties (Thomas and Scott, 1993). Crosses between cultivated varieties would most likely have been spontaneous until the last two centuries, when controlled crosses were made (Bowers *et al.*, 2000).

There are no documented records of grapevine breeding or crosses performed before the 19th century (Bowers *et al.*, 2000; Regner *et al.*, 2000a). With the exchange of grapevine cultivars between different vine growing regions, the true origin of most of

today's cultivars is unknown, and can only be traced back to the vines, which are clones of the original plant (Reisch, 2000; Silvestroni *et al.*, 1997; Sefc *et al.*, 2000b).

Silvestroni *et al.* (1997) hypothesised that these source plants have one of two origins. Either a single seedling produced different biotypes during the process of vegetative propagation as a result of somatic mutations; or more than one morphologically similar but genetically distinct seedling resulted in a variety of possible progenitors. Interestingly, the polyclonal origin of cultivars has until recently had little evidence to support it.

Several authors have explored the concept of geographic grouping of grapevine cultivars. Bowers *et al.* (2000) now offer the first evidence of this – that geographic groups represent different genetic groups. Biological characteristics of different species have been investigated with the aid of both molecular markers and isozymes. It is generally expected that self-pollinating crops have a larger proportion of their total variation distributed between populations rather than within populations, whereas in cross-pollinating crops (like grapevine) the variation within populations is expected to be more significant (Hodgkin *et al.*, 2000).

1.2.3 Grapevine in South Africa

In 1652, the Dutch East India Company sent Jan van Riebeeck to the Cape of Good Hope with a mandate to establish a victualling station. Its sole intended purpose was to re-supply ships travelling around Africa to the Indian sub-continent with fresh produce.

Van Riebeeck, the first governor of the Cape, had brought Hanepoot, Muscadel and Stein vines with him from Holland (Perold, 1927), which he planted in 1655 and from which the first Cape wine was produced in 1659. Unfortunately, the Dutch didn't have a strong winemaking tradition and floundered in foreign conditions. It was only when Simon van der Stel succeeded van Riebeeck as governor of the Cape in 1679 that things improved. Van der Stel was an avid wine enthusiast, and knew a thing or two when it came to winemaking. He planted the first vineyard at his Groot Constantia homestead, which is still now renowned for its wine. The historic town of Stellenbosch is named after Van der Stel, and lies at the centre of the Cape winegrowing region about 50km from Cape Town (Wines of South Africa, 2000).

The arrival of the French Huguenots at the Cape during 1688 brought an upturn in the wine industry - they brought with them a culture of winemaking and knowledge of viticulture since many came from the south of France (Perold, 1927). Just over 200 years later, however, the area was thrown into turmoil with the discovery of *Phylloxera* (which severely affected the wine industry) and the outbreak of the Anglo-Boer War (1899 – 1902) (Pongrácz, 1983).

In 1918 Charles Kohler established the Ko-operatiewe Wijnbouwers Vereniging van Zuid Afrika Beperkt (Co-operative Winegrowers Association of South Africa Limited), which was to act as an umbrella organisation for its 4 300 wine farmer members, and brought order and direction to the industry (Wines of South Africa, [2000]; KWV International, [N.d.]).

1.3 The Viticulture Industry

Viticulture is one of man's oldest agricultural activities – Egyptians were harvesting grapes around 4700BC and Mesopotamians were producing wine by 3000BC (Sefc *et al.*, 1998c). Greece got the vine 2 000 years later (1000BC), from where it spread to Europe and on to the New World by the 1550s AD.

European palates of the Middle Ages were not quite as refined as they are today – the only distinction made being between 'huntsch' (a poor quality wine) and 'frentsch', a high quality wine. Only from the 1400s onwards did grapevine cultivar names appear (Sefc *et al.*, 1998c).

Today, the majority of cultivars grown worldwide for wine and fruit production are *V. vinifera* (Thomas and Scott, 1993). The primary product of the grapevine is wine, but it is by no means the only product. Brandies, fortified wines and sweet wines are also produced, as well as non-alcoholic beverages, vinegar and other distillates.

The grape also supports a large fresh and dried fruit industry – table grapes, raisins, preserves, grape juice, grape syrup and grape-seed oil (Perold, 1927). Products of the vine also feature prominently in Mediterranean cuisine. Herbal medicines are derived from grape berries, grape leaves (tea) and grape seed extracts. These can be used to

treat a number of ailments, including inflammation of the joints and poor circulation (Cangelosi, 2001).

1.4 The SA Wine industry

There is a total of 106 331 Ha under vine in South Africa, which includes both wine and table grapes, currants and rootstocks. Although South Africa vineyards account for only 1.5% of the world's total, we produce 3% of the world's wine, and rank eighth internationally with respect to volume output (Wines of South Africa, 2000).

In 2000, just over 540 million litres of 'good wine' (for drinking purposes; this for example excludes distilling wine) was produced, of which 79% was white wine, the remaining 21%, red. (South African Wine Industry Information and Systems (SAWIS), 2000). Exports increased by nearly 18% in 2001, with sales of red wines growing the most rapidly. According to Mr Dennis Dykes, chief economist for Nedcor, the wine industry "yielded significant indirect benefits for the economy as a major employer, exporter and foreign tourist draw card" (ProAgri, 2002).

A study commissioned by SAWIS concluded that the Cape wine industry contributed some R14.6 billion to the regional economy in 1999, and made up almost 10% of the gross geographic product in the Western Cape. Furthermore, and possibly most importantly, the South African wine industry employs 348 500 people, 98.5% of which are unskilled or seasonal labourers.

1.5 Ampelography

Ampelography is derived from the Greek *ampelos* (vine) and *graphein* (to describe, to draw), and describes the science of classifying *Vitis vinifera* vines based on morphological features (Pongrácz, 1983; Perold, 1927). This technique is not helpful in identifying accessions of other *Vitis* species, which make up a significant proportion of the core germplasm at USDA-ARS Plant Genetic Resources Unit (PGRU) at Cornell University, Geneva, New York, and USDA-ARS National Germplasm Repository (NGR) at the University of California, Davis, two premier repositories of American and European germplasm (Lamboy and Alpha, 1998).

Ampelography is the description of organs and organ traits in a specific code, while ampelometry is the study of vine leaves with the aid of quantitative linear and angular

methods in order to identify or characterise each genotype (Boselli *et al.*, 2000). Until relatively recently, vines were distinguished solely on the basis of Ampelography (Hinrichsen *et al.*, 2000) and Ampelometry (Boselli *et al.*, 2000). Generally however, the physical features used in identifying a vine are 'not sufficiently constant to prevent doubt and errors from occurring' (Perold, 1927). Features can, and do change in appearance under different environmental conditions and disease states (Lopes *et al.*, 2000), and some (e.g. berries) may not always even be present. It is therefore difficult to accept that a method based on the, albeit detailed, phenotypic inspection of a grapevine is totally accurate and objective.

Texts such as that of Galet (1979), cited in Thomas *et al.* (1993), exist to describe quantitative ampelographic methods; i.e. to describe how to identify a vine on the basis of its appearance. There are currently 151 descriptors (criteria) used in grapevine identification, which explains why even expert ampelographers can disagree on the identity of a cultivar (Weihl and Dettweiler, 2000). There is still as yet no infallible classification system for vine – and as the difference between many cultivars is small, it is easy to conceive of mis-identification and misnaming of such cultivars.

In total, the world's collection of grapevine plant material has been estimated to contain between 10 000 and 24 000 different cultivars, but the true figure is probably much lower (Regner *et al.*, 2000a) - a realistic estimate is around 5000 cultivars (Thomas *et al.*, 1994). However, this discrepancy can explain the high number of synonymies (multiple names for the same cultivar) and homonymies (different cultivars described by the same name) in the vine world today (Thomas *et al.*, 1994; Bowers *et al.*, 1996; Borrego *et al.*, 2000).

The global spread of vegetatively propagated cuttings of a small original selection of vines over hundreds of years and countless cultural and political boundaries into new environments, has resulted in many problems for this phenotype-based identification system (Sánchez-Escribano, 1999). Most varieties are known under different names in different countries, depending on what is most easily recognised (Perold, 1927).

Many wines have also acquired new names as they have moved across the world. In many countries, wine is identified by the cultivar from which it is made. The accurate

identification and labelling of cultivars is a legal prerequisite in Europe, where wines are identified primarily by their *geographical* origin, but where the *varietal composition* of each is dictated by law (Bowers *et al.*, 1996; Sánchez-Escribano, 1999).

Many examples exist of misnaming or misidentification of cultivars and rootstocks, which has resulted in duplicate accessions in germplasm banks (Walker and Boursiquot, 1992), mixed commercial plantings and the planting of completely incorrect cultivars (Thomas *et al.*, 1994; pers. comm. JT Burger, 2002).

Strict guidelines and rules for most aspects of the wine and table grape industries in Europe are laid out by the European Commission. These include which varieties may be grown on which rootstock species, the composition of any wines or alcohol products, as well as the procedure for accurate labelling of the product, grape or wine. Guidelines are also given for the conclusive identification of varieties, and recognition is given to the fact that it is still very difficult to find two different ampelography experts to agree on grapevine identities. To this end, the commission reports the streamlining of the definitions used by the 'Office International de la Vigne et du Vin' (OIV), the International Union for the Protection of new varieties of Plants (UPOV) and the International Board for Plant Genetic Resources (IBPGR) in the identification of varieties (Classification of vine varieties..., 1993). These institutions have previously been criticised for their lack of communication, which has resulted in distinct codes for the same characters and different methods of notation for varying levels of expression (Boursiquot, 2000). It would however, obviously be more desirable to have a quick, easy and reliable method to conclusively identify cultivars; molecular markers are one way for even non-experts to accurately determine the identity of a sample (Boselli *et al.*, 2000).

A similar commission – the Wine Certification Board – was established in South Africa in the 1970s, to protect both the producer and consumer. Legislation was also put in place to protect wines of origin (WO), and wines made from certain cultivars or vintages. This legislation was drawn up to comply with European Union guidelines, as the EU is a major export destination for South African wines (Wines of South Africa, 2000).

1.6 The Need for Molecular Markers

The accurate identification of grapevine is essential for the viticulture industry. Errors in the identification of a cultivar at the nursery or during the establishment of a vineyard can make a significant financial impact on a farming enterprise. The distribution of incorrectly labelled fruit or wine can have adverse effects on the professional reputation of a winery or winemaker and consequently affect market share (Thomas *et al.*, 1994).

In 1985 grafted vines were illicitly smuggled into the Western Cape from abroad. These were purportedly clones of a high quality Chardonnay mother vine, and buyers had high hopes. However, the vastly inferior crop that was subsequently harvested by numerous farmers prompted an investigation into the true identity of this "Chardonnay". It was eventually found that the grafted vine that was smuggled in was 'Auxerrois'. Morphologically, Auxerrois and Chardonnay are virtually identical – Auxerrois scions (buds) are tinged red; Chardonnay scions are green – but Auxerrois produces a wine greatly inferior to Chardonnay. This error in identification resulted in large financial losses for all parties concerned, as well as a number of lawsuits (pers. comm. P. Goussard, 2002).

Global industry depends on the production of established cultivars and clonal selections (Thomas *et al.*, 2000), as a result of both industry and consumer preferences. The increase in the international trade of grapevine and rootstock plant material necessitates a reliable method of genotype identification. (Sefc *et al.*, 1998b)

The problems experienced by the international wine community in this regard have led to proposals for a vine identification system satisfying three conditions – international co-operation in the systematic evaluation of all major germplasm collections, a method of identification and a standardised testing and evaluation system (Thomas, 1994).

Although cultivar identification is based primarily on morphological characters described by the OIV, markers based on variation at the protein or DNA level have also proved useful. In fact, most cultivars can be distinguished from each other on the basis of almost any marker due to the vegetatively propagated grapevine's high level of heterozygosity. None of these methods though, has the resolution to identify an unknown sample as belonging to a certain cultivar, or to identify clones of one cultivar,

which have arisen because of somatic mutation or clonal selection (Martínez-Zapater *et al.*, 2000).

According to the OIV definition C137, clones are "a group of individual plants propagated asexually from a single ancestor" (Classification of vine varieties..., 1993). As such, clones should be genetically identical to the parent plant. However, in the case of grafting, the scion (the shoot grafted onto the rootstock species) may sometimes undergo somaclonal variation (somatic mutation). These are genetic changes in a 'parent' genotype that were first noted to occur spontaneously in tissue culture conditions (Larkin and Scowcroft, 1981), and which have since been observed in various stages of a plants development. The exact cause of this variation is unknown, but may be the result of changes in chromosome number and/or structure, point mutations and epigenetic changes. If a cultivar's scion develops somatic mutations as a result of stresses associated with the grafting process, the adult plants derived from these cells may express the mutation and are then classified as somaclones.

Clones of a variety (cultivar) are essentially slightly different versions of the same thing, which developed in specific locations over the centuries, and whose subtle differences suit them to their local environment. Often they differ only with respect to berry colour. Pinot noir is perhaps the best-known variety for having many different clones, at least 46 at the last count (The Wine Room, 2000). To date, it had not been possible to distinguish between the clones of 'Pinot noir' using any form of molecular marker analysis (Ye *et al.*, 1998).

Clones of the same cultivar are very similar in appearance. Since ampelography is based on the phenotype of the plant (and as such is influenced by external environmental factors), it is almost impossible to refine a classification beyond the level of the cultivar. These same external factors however, have no effect on molecular DNA markers (Merdinoglu *et al.*, 2000), which therefore provide an objective means of identifying cultivars as direct indicators of genotype, avoiding problems associated with environmental influences, physiological factors and developmental and tissue-specific expression (Botta *et al.*, 1995).

Molecular markers have significant advantages over other types of markers including the fact that many are detectable, and heterozygotes and homozygotes can often be distinguished – a benefit in a crop for which a high level of heterozygosity is advantageous (Hodgkin *et al.*, 2000).

DNA profiles have been recommended as supporting evidence for determining cultivar identity and purity in many horticultural and crop species, thus preventing the infringement of inventor's or breeder's rights (Jain *et al.*, 1999). Local genetic resources should also be protected, their genetic diversity evaluated and used to establish possible relationships among cultivars grown today (Malossini *et al.*, 2000).

Many instances exist where incorrect vine identification (based on ampelography) has resulted in errors in naming of cultivars and subsequent mixed plantings. In Australia, Cabernet Franc and Merlot were planted together (they were subsequently found to differ significantly using molecular techniques), and both Chenin Blanc and Crouchen have been incorrectly labelled as Semillon. Incorrectly labelled rootstocks have also recently been correctly designated at the UC Davis germplasm repository where 'Teleki 5C' had been labelled as 'SO4' (Thomas *et al.*, 1994).

Molecular markers therefore offer a system whereby grapevine cultivar and rootstock species can be conclusively identified and distinguished, with no negative influences on the results from external factors. This approach would invariably be more efficient, robust and reliable than traditional methods, as well as providing molecular markers which could potentially find further application in a number of different research areas like crop improvement, conservation, germplasm management and mapping, to name but a few.

1.7 The Applicability of Molecular Markers

Research into grapevine genetics has been previously hindered by lack of stocks, inbreeding depression, large space requirements and a long juvenile period. The recent availability of comparatively inexpensive and easy to use molecular markers has facilitated research into *Vitis* genetics (Reisch, 2000). A number of markers are now available to the researcher for investigations into nearly every plant type known, the

choice of which is governed by the nature and objective of the investigation, as well as properties of the species (Hodgkin *et al.*, 2000).

Another use of molecular markers is in the management of germ banks. Markers may be used to identify duplicate accessions as well as in the identification of misnamed accessions. They may also be used in the analysis of genetic diversity and redundancy in a core collection (Hodgkin *et al.*, 2000).

Molecular markers have contributed to a better understanding of heterosis (hybrid vigour), and have helped improve ways of identifying potentially vigorous cross combinations (Hodgkin *et al.*, 2000). These markers also facilitate the early detection of both desirable and undesirable traits. Marker assisted selection (MAS) is ideally suited to long-cycle vegetatively propagated crops like grapevine; if important genes are associated with a given marker, early selection of seedlings can take place. Pyramiding of multiple genes for a single trait is also possible using molecular markers (Reisch, 2000).

Various molecular markers have been used in genetic diversity estimates, evolution studies and the analysis of migration. They also open up the opportunity of using MAS in young plants for characteristics like disease and pest resistance, drought tolerance, sugar content and seedlessness (Hodgkin *et al.*, 2000; Reisch, 2000).

Genome synteny (where some genes are found in similar map positions in different organisms) is a useful tool that provides the opportunity for using information gained on one crop on another, related species and has important implications for horticultural and food crops, as well as for those generally under-utilised or used mainly by subsistence communities (Hodgkin *et al.*, 2000). These types of crops tend to be neglected as far as research efforts are concerned, and there often simply is not the funding available for the development of novel markers, unique to a marginal crop.

DNA markers can be used for forensic analyses involving grapevine identification and parentage determination. Field identification of unknown cultivars and the identification of duplications or synonymies in germplasm collections is another potential application. They can also be used to determine the extent of diversity within

a collection and so ensure that every representative of a species is present (Reisch, 2000).

Finally, and particularly relevant to this study, the identification and characterisation of varieties is important for any selection programme and the subsequent agricultural/ commercial use of such varieties (Monteiro *et al.*, 2000). Molecular markers represent an objective tool with this can be accomplished (Crespan and Milani, 2000).

Ideally, markers for research or production need to display a high degree of polymorphism, be easy to use, inexpensive, abundant and reliable. Each type of marker represents a trade-off between these characteristics - markers offer different things in terms of level of polymorphism, ease of analysis and stability. Only some are sufficiently reliable to allow for the exchange of data between laboratories (Reisch, 2000; Crespan and Milani, 2000).

The following section concentrates on the various markers that have fallen in and out of favour in grapevine research in chronological order, with their respective advantages and disadvantages. It also briefly explains each markers mechanism of working, application and limitations, focussing in detail on the marker that formed the basis of this work, the simple sequence repeat.

1.8 Molecular Markers in Grapevine – An Overview

1.8.1 Isozymes

Isozymes are differently charged functional forms of the same enzyme that catalyse reactions with the same mechanism but which have different kinetic parameters and which differ in electrophoretic ability (Reisch, 2000). These isomers are separated from each other using starch gel electrophoresis, resulting in patterns similar to those generated when microsatellite data is separated on agarose gels (De Woody and Avise, 2000).

Isozymes were one of the first molecular techniques in the identification of vines, and two isozyme systems are currently accepted by the OIV as useful for the initial screening process of cultivars. Over 20 isozyme polymorphisms have

been identified in grapevine (Reisch, 2000), and Walker and Boursiquot (1992) successfully proved that two differently named rootstocks were in fact identical using isozyme analyses.

Although a number of groups have reported successes in identifying cultivars using this system (Bowers *et al.*, 1993), an equal number of researchers have found their discriminant power not sufficient to consistently and reliably distinguish varieties (Crespan and Milani, 2000). The enzymes used in this type of analysis must be stable in different environments, present and active in the tissue used as well as polymorphic between different samples (cultivars) (Bowers *et al.*, 1993). Unfortunately, tissue-specific expression patterns, variable stability and relatively low levels of polymorphism have limited the widespread adoption of isozyme analysis (Bowers *et al.*, 1993; Thomas *et al.*, 1993; Monteiro, 2000).

It was these shortcomings of the isozyme technique that spurred on the search for new, more reliable molecular markers (Crespan and Milani, 2000).

1.8.2 Restriction Fragment Length Polymorphisms (RFLPs)

Restriction Fragment Length Polymorphisms (Botstein *et al.*, 1980) are variable lengths of DNA generated by the digestion of gDNA with restriction endonucleases. Polymorphism results from mutations that either create or destroy enzyme recognition sequences, giving longer or shorter fragments. These fragments are separated using gel electrophoresis, and visualised using a labelled probe and Southern blotting. Many different types of DNA sequences can be used as probes, the choice depending on the level of polymorphism required and the objective of the study (Thomas and Scott, 1993).

This technique requires large amounts of high quality DNA and is time-consuming and relatively expensive (Reisch, 2000). The fact that the extraction of high molecular weight DNA from grapevine is difficult due to the high levels of secondary compounds present, such as polysaccharides and polyphenols, (Monteiro *et al.*, 2000) could pose a problem to groups who choose RFLPs as their marker of choice. Grapevine DNA is also extensively methylated, which

means RFLP analysis requires the use of methylation-insensitive restriction enzymes (Thomas *et al.*, 1993).

Co-dominantly inherited RFLP markers have been successfully used for cultivar identification in a variety of species, including rice, apple, roses and avocado. They were also the first widely used and generally accepted method of DNA fingerprinting in humans where it found application in the courts and was considered conclusive legal evidence of identity in criminal cases, cases of paternity and immigration (Bowers *et al.*, 1993).

Specifically concerning grapevine, RFLPs were used to show that a Californian cultivar 'Zinfandel' was actually the same as the Italian cultivar 'Primitivo' (Bowers *et al.*, 1993). The same study also confirmed the correction of a homonymy between Pinot noir 1 and 19 – they were previously thought to be distinct but are in fact identical. Two cultivars were also confirmed as being different ('Petit Syrah' and 'Durif'), even though they had previously been considered identical. Bourquin *et al.* (1995) had used RFLPs to distinguish different rootstock species, and Bowers *et al.* (1993) and (1996) also used RFLPs, applying them to the identification of wine grape cultivars.

RFLPs had been the dominant marker type since the early 1980s (Thomas and Scott, 1993), and was one of the first DNA-based marker systems used in grapevine studies. It was largely abandoned with the advent of PCR-based techniques (Mullis and Faloona, 1987), which were less laborious to use, safer because radioactive exposure was avoided and required much less DNA as a starting material. They were generally also much more informative with regards to genetic characterisation than RFLPs (Regner *et al.*, 2001).

1.8.3 Random Amplified Polymorphic DNA (RAPD)

In RAPD (Williams *et al.*, 1990) reactions, single decamer primers of arbitrary sequence are used to amplify DNA fragments of variable length. Amplification depends on the single primer annealing in opposite orientations (i.e. facing each other) on a strand of DNA at a distance close enough to allow for extension of a complete fragment by *Taq* polymerase – from 200bp to 2000bp apart.

Polymorphisms result when primers anneal at different positions, in different orientations and at varying distances from each other. Amplification products are easily separated by agarose gel electrophoresis and visualised via ethidium bromide staining and UV illumination.

One of the first PCR-based methods employed, RAPDs followed soon after the advent of the polymerase chain reaction in 1987 (Mullis and Faloona, 1987). This technique is simple, quick and easy to perform and much cheaper than RFLPs, as it does not require a labelled probe, many reaction reagents or specific apparatus in order to view amplification products. Furthermore, no genomic information is required for the design of specific primers (Ibañez, 2000; Ryan *et al.*, 2000). RAPDs are therefore the method of choice for the identification of species and cultivars when there is no previous sequence knowledge of the genome.

Although a popular technique, it soon became evident that results from RAPD analyses were influenced by many variables, which compromised the repeatability of results even within the same laboratory and therefore seriously hindered the exchange of results between laboratories (Reisch, 2000; Büscher *et al.*, 1993; Cipriani *et al.*, 1994). The RAPD technique is also a dominant marker system, and this, together with the numerous bands amplified in a single reaction, make gel scoring and statistical analysis difficult (Ibañez, 2000).

Despite these drawbacks, Büscher *et al.* (1993) described the potential use of RAPD markers in the identification of clones of a grapevine cultivar, and preliminary results of Lodhi *et al.* (1995) have suggested that the RAPD markers used would be useful in distinguishing between clones, but this has yet to be proven.

RAPDs have also previously been used to conclusively identify synonymies and homonymies in grapevine collections (Moreno *et al.*, 1995), as well as to distinguish between different cultivars of grapevines. However, these results were not confirmed as being constant between different locations (Ulanowsky *et al.*, 2000).

RAPD analyses done on the Pinot group [Pinot blanc, Pinot noir, Pinot gris and Meunier – typically difficult to distinguish (Ye *et al.*, 1998)] did not yield any clonal differences (Bellin *et al.*, 2000), and also had problems with reproducibility (Regner *et al.*, 2001) - RAPDs have been shown to produce false positives and negatives, due to the competitive nature of the reaction (Vidal *et al.*, 2000).

While RAPDs are still commonly used today, the problems many have encountered while using them, most significantly their lack of reproducibility and difficulty in sharing data between groups, prompted a move towards more robust, reliable marker systems whose results could be replicated and shared between laboratories.

1.8.4 Simple Sequence Repeats (SSRs)

Since a thorough description of this marker system follows at a later stage - its evolution, development, advantages and disadvantages and applications, only a basic outline of this technique is described here.

This marker has alternately been referred to as Simple Sequence Length Polymorphism (SSLP) (Tautz *et al.*, 1989), Simple Sequence Repeat (SSR) (Morgante and Olívieri, 1993), Sequence Tagged Microsatellite Site (STMS) (Beckman and Soller, 1990) and Sequence Tagged Repeats (STR) (Hamada *et al.*, 1982). In future in this script, the marker *system* as well as the individual marker *locus*, will be interchangeably referred to as either 'microsatellite(s)' or 'SSR(s)'.

After RAPDs, SSRs became popular even though their development is initially expensive and laborious (Crespan and Milani, 2000), involving the generation of a genomic library, screening of clones and sequencing. This large initial financial outlay has prohibited the development of microsatellite markers on many species, but the fact that many microsatellite loci show sequence conservation between species and even some taxa (Cipriani *et al.*, 2000) has allowed for the transfer of useful SSRs with a concomitant reduction in research

costs. Until recently, for example, only 3 grapevine research groups worldwide had sufficient resources to develop microsatellite markers independently – Bowers and Meredith (USA), Steinkellner (Austria) and Thomas and Scott (Australia).

Microsatellites are tandemly repeated arrays of short (1-6bp) motifs of DNA, repeated end-to-end between 50 and 200 times (Cipriani *et al.*, 2000; Scott *et al.*, 2000c). They are found scattered throughout eukaryotic genomes and reside in predominantly non-coding regions of DNA (Tautz and Renz, 1984). The number of times the core repeat unit (motif) is present varies, with the result that these loci give DNA fragments of varying lengths when amplified using PCR. These length polymorphisms are resolvable using polyacrylamide gel electrophoresis.

Sequences adjacent to the repeat motif show a high degree of conservation, enabling the design of site (microsatellite) specific primers. This information, which essentially describes a single simple sequence repeat locus, can easily be shared between research groups, as can the generated data (allele lengths expressed in base pairs), in a digital format (Borrego *et al.*, 2000). It is this ability to express allele lengths easily in the form of quantitative data that has allowed databases of allelic profiles at SSR markers to be developed for a number of agriculturally important crop species like wheat, barley, sorghum, maize and cotton (Sánchez-Escribano *et al.*, 1999; Monteiro *et al.*, 2000). Just such a database is being developed for the exclusive use of all viticulturalists and grapevine geneticists around the world. This effort falls under the aegis of Agrogene SA, which co-ordinates and directs the efforts of laboratories participating in the Vitis Microsatellite Consortium (VMC).

The SSR technique is renowned for its high level of informativeness (polymorphism), reliability and reproducibility as well as its locus-specificity, co-dominant inheritance and simple PCR-based method of detection – it is therefore an ideal marker for this type of international effort. For example, the amenability of SSRs to automation makes them quicker and easier to work with than RFLPs (Grando and Frisinghelli, 1998), they are more polymorphic than

isozymes, as well as being independent of the tissue analysed and stage of development. Allelic profiles generated are also reliable and repeatable between laboratories, – an advantage over RAPDs.

1.8.5 Single Primer Amplification Reactions (SPARs) and Inter-simple Sequence Repeats (ISSRs)

Gupta *et al.* (1994) developed SPARs as a new way of discovering microsatellite DNA markers quickly and simply. As the name suggests, only a single primer is used in each PCR consisting of a microsatellite core motif. PCR amplifications using such primers provide a shortcut to determining sequences flanking variable microsatellite loci. Results obtained have proved to be polymorphic, reproducible and heritable. In addition, there are usually more polymorphic loci amplified per SPAR than per RAPD. It is also possible to generate fragments in a number of evolutionary diverse species, like pine, lettuce, tomato and grapevine (Gupta *et al.*, 1994), using similar primers.

A related technique, using *two* primers based on microsatellite core motifs, is ISSRs (Zietkiewicz *et al.*, 1994). This is also a PCR-based multi-locus marker system and works according to the same principles as RAPDs. This method generates many more bands than SSRs, since the 'gaps' *between* numerous microsatellite loci of a certain type are amplified. In this regard, ISSRs are approximately as informative as AFLPs, but with the added advantage of being more reproducible on the same target DNA (Arnau *et al.*, 2000; Regner *et al.*, 2001).

The complexity of patterns generated is influenced by the repeat used – di-, tri- or tetranucleotide repeat give patterns of decreasing complexity respectively, corresponding to the frequency of the type of repeat in the genome. In this way, ISSRs are therefore helpful in predicting the frequency and level of polymorphism of simple sequence repeat motifs (Zietkiewicz *et al.*, 1994).

ISSRs provide a convenient way of detecting and measuring common genetic events underlying plant genomic instability and leading to somaclonal variation (Leroy and Leon, 2000). They have also proven useful in genetic diversity

studies and have been extensively used in the characterisation of grains (Blair *et al.*, 1999).

Moreno *et al.* (1998) used ISSRs to evaluate the level of inter-varietal polymorphisms between grapevine cultivars, and although no variation was detected (supporting previous findings), the high reproducibility of the results was confirmed. Furthermore, although Regner *et al.* (2001) disagrees that ISSRs would be useful in inter-varietal identifications, he anticipates great potential for their use in clonal discrimination. As yet, however, very few reports exist in the literature of the use of ISSRs in grapevine research.

1.8.6 Amplified Fragment Length Polymorphism (AFLPs)

AFLPs, developed by Zabeau and Vos (1993), involve the restriction enzyme digestion of gDNA (as for RFLPs), followed by the selective amplification of a proportion of the fragments using primers complementary to adaptors ligated to the fragment ends. In this way, the entire genome is scanned for many polymorphisms (compared to site-specific marker systems like SSRs) (Markert *et al.*, 2001). Multiple fragments (50-100), separated by polyacrylamide gel electrophoresis, are generated in each reaction and yield DNA markers of a genome-wide origin.

Although no prior sequence information is required and large numbers of polymorphic markers can be generated with ease (Bellin *et al.*, 2000; Ryan *et al.*, 2000), AFLPs are more complex to work with than most other (PCR-based) markers. Bearing this in mind, Hinrichsen *et al.* (2000) recommends using AFLPs only for specific tasks like the identification of clones, which is difficult to achieve with other methodologies. Furthermore, according to Cervera *et al.*, (1998) the technique's comparative complexity may 'preclude its general use in laboratories' – clone-specific SCAR markers would probably find wider applicability as they are PCR-based and therefore easier to use.

Although Bellin *et al.* (2000) questioned the reproducibility of AFLPs, this non-reproducibility could have been due to an incomplete restriction digestion of gDNA. On the other hand, Cervera *et al.* (1998) has found AFLPs to be reliable

and reproducible; and has criticised 'excessive selection' in the selective amplification step of the reaction for any inconsistent and unreliable results.

High levels of sugars and polyphenolic compounds in certain organs and at certain stages of development of the strawberry plant affect the restriction enzyme digestion of gDNA (Arnau *et al.*, 2000). This could then also be expected to be a possible source of inconsistency in AFLP work on grapevine, which is known to have high levels of sugars and polyphenols, the presence of which makes extraction of clean gDNA problematic (Wang *et al.*, 1996).

A high level of genetic variability in grapevine has allowed for the differentiation of cultivars using RAPDs, RFLPs and STRs (Bellin *et al.*, 2000), yet none of these systems is able to distinguish between clones of the same cultivar. AFLPs however, have the ability to screen the highest number of anonymous loci of all the molecular markers in use today, and are therefore more efficient and more likely to detect differences in closely related cultivars or clones of the same cultivar (Martínez-Zapater *et al.*, 2000; Lamboy and Alpha, 1998).

This dominant marker system is expected to be particularly useful for the classification of grapevine, as amplification with only a single primer combination has been shown to be sufficient to identify cultivars. In fact, there are conflicting reports regarding the efficacy of AFLPs in grapevine cultivar and clonal analyses:

Cervera *et al.* (1998) successfully used AFLPs to distinguish between clones in numerous cultivars in a Spanish collection using only two primer combinations. Cases of synonymy and homonymy were also cleared up. In contrast, Bellin *et al.* (2000) found AFLPs could only distinguish between different varieties, with no higher resolution possible. Additionally, Sefc *et al.* (2000a) used AFLPs in an attempt to distinguish between the Pinot group of cultivars (Pinot blanc, Pinot noir and Pinot gris). However, apart from differences in the intensities of some bands, these 3 varieties could not be separated.

Besides work involving varietal identification, AFLPs have also successfully been used to estimate genetic relationships and in the fine mapping of the genomes of a variety of organisms (Bellin *et al.*, 2000). They have also been used to develop markers for the purpose of MAS of traits in grapevine at the USDA-ARS Horticultural Crops Research Laboratory, and a linkage map is being constructed. Traits marked for selection include berry and seed size, disease resistance and seasonality (Ryan *et al.*, 2000).

AFLPs have obvious advantages as a marker system – they are highly polymorphic (informative), reasonably reproducible and relatively easy to use. However, problems can arise if data is to be shared between research groups working with different mapping populations. In such cases, a highly reliable co-dominant system that allows for the easy sizing of specific amplification products (alleles), and the sharing of both locus information and data in a digital format is preferable.

1.8.7 Sequence Characterised Amplified Regions (SCARs)

For characterisation purposes, RAPD and AFLP techniques generate relatively complex patterns, making the analysis of data tedious and time consuming (Vidal *et al.*, 2000). SCARs (Michelmore and Paran, 1993) offer a way of finding single DNA markers, possibly linked to a specific trait, which could be amplified in a simple PCR to determine the presence/absence of a particular trait.

SCARs fall under a broader group of Sequence Tagged Sites and are usually developed from fragments generated from RFLP, RAPD or AFLP reactions as follows: A fragment that is unique to a specific trait is cloned and sequenced in order for primers (16-24mers) to be designed for that specific marker. In a simple PCR, these locus-specific primers can then be used to test a single genetically defined locus.

In this way, SCAR markers for seedlessness were subsequently developed on grapevine from RAPD markers (Lahogue *et al.*, 1998), and show potential to be developed for a number of commercially important traits for which a genetic basis is being sought. For instance, a RAPD marker associated with phylloxera

resistance in rootstocks was converted into a SCAR marker (Eimert and Schroder, 2001), allowing for the early detection of resistant and susceptible lines.

RAPD markers associated with disease resistance have also been converted into SCAR markers for marker assisted selection and to anchor linkage groups on a map and also to use in determining homology with respect to published maps of unrelated grapevine populations (Reisch *et al.*, 1995).

1.8.8 SSRs derived from Expressed Sequence Tags (ESTs)

ESTs fall into the general class of sequence tagged sites (STSs), and are derived from cDNA (Paterson, 1996, cited in Reisch, 2000). Messenger RNA is extracted and converted into cDNA by way of reverse transcription PCR. These DNA fragments are then cloned into a library of vectors and the sequences of these clones are maintained in a database.

Scott *et al.* (2000b) developed 16 SSR markers using an Expressed Sequence Tag database of *Vitis*. Using a publicly available database is a more convenient and cheaper alternative to developing SSR markers from scratch. However, this approach can only obviously be used for species for which a database already exists.

There was speculation that these EST derived SSRs would be less polymorphic (and therefore less informative than other SSRs). Scott *et al.* (2000b) found that the functional primer pairs were all highly transferable across grapevine cultivars, species and even genera. EST derived SSRs were also found to be more representative of all repeat motifs than enriched SSRs (microsatellites derived from a database obtained by screening a gDNA library with microsatellite core repeat probes).

1.8.9 Organellar genome analysis

Three characteristics of extra-nuclear DNA make it useful for systematic studies:

i. It is inherited cytoplasmically and essentially asexually in plants, as there is no recombination during sexual reproduction.

ii. In nearly all eukaryotes, it is usually inherited from only one parent - mitochondria are usually paternally inherited and the chloroplasts are usually maternally inherited (Birky, 1995), although the opposite situation has been found to apply in some species, like *Actinidia* (Testolin and Cipriani, 1997). Some plants also tend to display maternal, paternal and biparental inheritance simultaneously (Birky 1995).

iii. Extra-nuclear DNA generally displays very little within-species variation, because of the much lower rate of base pair substitutions in this DNA compared to that of nuclear DNA (Testolin and Cipriani, 1997).

Information regarding the inheritance of the chloroplast and mitochondrial genomes is a pre-requisite if they are to be used to trace the evolution of a species. The development of molecular markers has lead to a wealth of information being gleaned from these organelles with respect to inheritance, mutation rates and levels of polymorphism (Testolin and Cipriani, 1997).

1.8.9.1 Mitochondrial inheritance

To use mitochondrial markers, gDNA is extracted from the sample and specific mitochondrial primers are used to amplify specific genic regions. Polymorphisms may occur as length differences or as restriction endonuclease recognition site alterations.

mtDNA markers are widely used in population studies and genetic diversity analyses in aquaculture (Ferguson *et al.*, 1995), and are also popular markers in tracing human movements across continents and gaining insight into specific mitochondrial pathologies (Wallace *et al.*, 1999). As yet, very little work has been done using mtDNA markers on *Vitis* species, but this is an area that holds much promise for the future especially considering the interest in origins and development of cultivars commonly used today.

1.8.9.2 Chloroplast DNA analysis

Chloroplast genomes appear to be conserved between species in as far as gene function and order is concerned. Chloroplast DNA (cpDNA) is inspected using

PCR-RFLP and is a powerful tool for phylogenetic reconstruction at both inter- and intra-specific levels (Palmer, 1987, cited in Cipriani and Testolin, 1997), as well as being highly reproducible (Lefort *et al.*, 2000). Analysis of chloroplast DNA polymorphisms can shed light on the direction of a cross to reveal the male and female parents, as the chloroplast genome is usually maternally inherited in angiosperms (Regner *et al.*, 2000a).

The possibility of screening polymorphic microsatellite loci in the chloroplast genome, as well as the much lower rate of mutation in cpDNA compared to nuclear DNA makes this a useful technique to study the inheritance of plastids, cytoplasmic diversity and also to monitor gene flow. Owing to the high level of conservation between chloroplasts, it is possible that 'universal' primers will be able to be developed, to allow for cpDNA analysis of a number of diverse species (Arroyo-García and Martínez-Zapater, 2000).

Conserved cpDNA sequences for the 16S and 23S rDNA were used by Primikirios *et al.* (2000) to confirm the placing of Vitis (family Vitaceae) in the phylogenetic tree of flowering plants. Lefort *et al.* (2000) also used universal primers designed for dicotyledonous angiosperms (Weising and Gardner, 1999) to characterise 77 grapevines at cpDNA microsatellite loci. As a result, the first large-scale genomic deletion in chloroplast microsatellites was detected – a rare occurrence in the chloroplast genome as there is no recombination and a high degree of conservation. In addition, the Cabernet Franc x Sauvignon Blanc cross, which yielded Cabernet Sauvignon, was confirmed (Bowers and Meredith, 1997); the female parent (chloroplast donor) was found to be Sauvignon Blanc.

Finally, an effort to sequence the grapevine chloroplast genome is currently underway at the University of Stellenbosch. Once completed, this promises to provide valuable information regarding the inheritance of particular chloroplast genes, as well as facilitate the possible transformation of the grapevine chloroplast (another project currently underway at Stellenbosch University, Department of Genetics).

1.9 Repetitive DNA

Eukaryotic genomes comprise a combination of coding (in the minority) and non-coding (the vast majority) DNA sequences. Non-coding DNA can roughly be divided into highly repetitive and moderately repetitive sequences, of which moderately repetitive DNA can further be split into interspersed or tandemly repeated sequences. It is these tandemly repeated sequences that have proved to be the most useful as far as DNA marker development is concerned.

Satellite DNA consists of short sequences repeated many times in the same orientation, and without intervening spacer sequences (*tandemly*), and is usually found in the heterochromatic regions of the genome (Klug and Cummings, 2000). There are a number of different types of these short repeats, which are present in the genomes of many species (Tautz and Renz, 1984).

Hypervariable 'minisatellite' regions, discovered in the human genome by Jeffreys *et al.* in 1985, are characterised by tandemly repeated core sequences of 10bp-200bp, stretching over a few kilobase lengths of DNA. Their high level of variability is due to variation in the number of copies of the core repeats present. As a result of this variability, they were able to provide individual-specific fingerprints for most organisms, and indeed launched what is known today as 'DNA fingerprinting'. This requires the extraction of high molecular weight DNA, restriction enzyme digestion, Southern blotting (Edwards *et al.*, 1991) and a suitable method of hybridisation detection.

Microsatellites – short DNA motifs of 1-6bp – have the same characteristics as minisatellites: They are repeated end-to-end, each locus shows a high degree of polymorphism as a result of varying numbers of core repeats present, and they are inherited in a co-dominant, Mendelian fashion. Scattered throughout eukaryotic genomes, usually in non-coding regions, Tautz *et al.* (1986) showed that these microsatellite sequences are nearly ten times more frequent in eukaryotic genomes than non-repetitive sequences of the same length. Collectively, these middle-repetitive sequences are generally known as VNTRs (Variable Number of Tandem Repeats) (Nakamura *et al.*, 1987).

Although researchers had been aware of microsatellites for several years, it was only in 1989 that they were finally isolated from a genome, by three separate groups simultaneously (Litt and Luty, 1989; Weber and May, 1989; Tautz, 1989). Unique sequences flanking these loci were revealed to be conserved between organisms of the same species and were therefore used to design locus-specific primers, which could easily be used to amplify numerous polymorphic bands (length variants), resolvable on agarose or polyacrylamide gels (Cipriani, *et al.*, 2000). This technique quickly became another handy method of developing a genomic fingerprint – each conserved locus has multiple alleles, and there were hundreds of possible loci available for testing (Reisch, 2000). Furthermore, it soon became evident that there were several types of microsatellites (Thomas and Scott, 1993; Hancock, 1999; Scott *et al.*, 2000c):

- Perfect repeats with no interruptions (CTCTCTCTCTCT)
- Imperfect repeats, where one repeat is changed (CTCTCT**G**TCT)
- Interrupted repeats, where the motif is interrupted by several foreign bases (CTCT**GGG**CTCTCT)
- Compound repeats, where one type of repeat directly follows another (CTCTCTCTCT**GAGGAGGAGGAG**)

With the advent of PCR in 1987, the automated amplification of microsatellites became relatively simple (once the daunting task of developing SSR primers had been achieved) in comparison to the laborious method of fingerprinting using minisatellites, which are also generally too large to amplify (up to 50kb), due to limitations of the DNA polymerases used. Since microsatellites are detected using a simple PCR, this technique can be applied to samples containing as little as 100pg DNA (Balding, 1999), in comparison to the several microgrammes of DNA required for minisatellite detection. Consequently, SSRs have superseded minisatellites as the standard DNA markers for genome fingerprinting.

1.10 Mechanisms of Microsatellite Polymorphism

Microsatellites are particularly informative markers due to their high level of polymorphism, which manifests itself in the many possible alleles at any given locus.

These polymorphisms are length variants, resulting from variable numbers of core repeats amplified by the site-specific primers, and can vary by as little as one base pair.

There are two mechanisms of mutation thought to give rise to variations in allele length:

1. Unequal crossing over
2. Slipped strand mispairing

Simple sequence repeat motifs are found arranged in long, uninterrupted chains. It is this physical arrangement that is thought to predispose these areas of DNA to mispairing events and unequal crossing over of homologous strands (Levinson and Gutman, 1987a).

1.10.1 Unequal crossing over (UCO)

Unequal crossing over during recombination between homologous chromosomes is generally the mechanism by which new alleles are generated in longer repeats, like minisatellites (Jeffreys *et al.*, 1985). This faulty physical exchange of DNA occurs over longer stretches of repetitive sequences when the 'recombination machinery' of the replicating cell fails to determine accurately the alignment or register of the repeat sequence between homologues. As a result, gene 'conversion' can occur, creating new allele lengths. Interestingly, Jeffreys (2002) observed that the mutation processes at these loci are controlled by DNA factors near the locus, and are not as a result of the repeat's inherent instability.

Also, although microsatellites are not thought to mutate as a result of unequal crossing over, as the number of core units in an SSR increases, so does the possibility that the locus will mutate as a result of this process. This is more likely because homologous chromosomes with longer repetitive motif areas have an increased chance of mis-aligning at meiosis (Smith, 1998)

1.10.2 Slipped strand mispairing (SSM)

Slipped strand mispairing during replication ('replication slippage') is the primary mechanism by which new microsatellite alleles are generated (Regner *et*

al., 2000b; Scott *et al.*, 2000c). Further, it is only once a critical number of repeat units is reached that the locus has the ability to become hypervariable (Messier, 1996).

Slipped strands are thought to occur mainly on the lagging strand during DNA synthesis when the DNA polymerase molecule leaves the transcription complex. A protruding bulge is then created, which is either removed by polymerase proofreading activity (resulting in a shorter strand), or the cellular repair mechanism 'fills in' the gap (creating a longer strand), (Eisen, 1999). This is a short-range effect, dependant on the rate of slippage of the specific core repeat sequence. For instance, dinucleotides mutate faster than trinucleotides, and repeats with a high AT content mutate faster than those containing many GC repeats (Schlötterer and Tautz, 1992). Further, the more core repeats present, the greater the rate of mutation. Long repeats (like tetra- and pentanucleotides) are more likely to suffer large, multi-repeat deletions, while shorter repeats generally tend towards getting longer. These trends may explain why, over long periods of time, the number of repeats at a particular locus is stable around a mid-value (Eisen, 1999).

Slipped strand mispairing is highly unlikely during DNA replication of unique sequences, as the lagging strand has a precise copy to which it can bind in a specific manner. However, when repetitive sequences are amplified there is a much higher chance for the new strand to align 'out of synch' with the template strand, even though the base pairs are still complementary (Hancock, 1999).

This process is considered the primary mode of microsatellite mutation over unequal crossover for several reasons:

> Firstly, mutation rates in mitotic and meiotic yeast cells are similar, even though the rate of recombination during meiosis is much higher than during mitosis. Furthermore, the mutation rates at microsatellite repeats are much higher than the rate of recombination (Hancock, 1999).

Secondly, microsatellite stability is affected when genes for correction of replication errors are defective, which supports the SSM hypothesis, as it would occur during the replication of DNA.

Thirdly, repeat length stability is unaffected when genes with a role in recombination are defective (unequal crossover would require these genes before this model resulted in repeat instability).

Fourthly, the direction of the repeat relative to the leading and lagging strand influences its stability, which is consistent with the SSM model, as these two strands have slightly different mechanisms of replication.

Finally, changes in copy number without a corresponding re-arrangement of flanking markers have been detected at human microsatellite loci (Eisen, 1999).

1.11 Microsatellite mutation rates

The mutation rate of microsatellites is significantly higher than the rate of point mutations (approximately 10^{-9} to 10^{-10}) (Hancock, 1999), and the two models above - SSM and UCO - serve as explanation for this phenomenon (Levinson and Gutman, 1987a; Jeffreys *et al.*, 1985).

Evolutionary studies have shown microsatellite DNA to be less stable than coding DNA, probably due to its tolerance of mutations (Cipriani *et al.*, 2000), and is further explanation for the very high levels of polymorphism observed in repetitive sequences. Generally, SSR loci have been estimated to have mutation rates of approximately 10^{-4} per generation (Ciofi *et al.*, 1998).

In humans, the mutation rate of microsatellite loci is inversely proportional to the size of the motif – non-disease-causing dinucleotides have mutation rates nearly twice as high as that of tetranucleotides. Interestingly, disease-causing trinucleotides have mutation rate nearly seven times greater than that of tetranucleotides (Chakraborty *et al.*, 1997). Prokaryotic evidence also shows that longer microsatellites mutate faster than shorter repeats (Levinson and Gutman, 1987b).

There are several complications involved in the mutations of microsatellites. These include the apparent dependence of mutation rates on repeat motif structure, a preference for allele expansion rather than contraction, heterozygote instability and allele size constraints (Markert *et al.*, 2001).

To explain the observed range of allele sizes over a wide variety of organisms, as well as the tendency of allele sizes to increase over time, a number of theoretical models have been proposed.

1.12 Mutation models

Mutation models have been proposed to provide a description of the likelihood of change between allelic states. For example, two conflicting models exist to answer the question of the total number of possible alleles – is there an upper limit, or is there an infinite number of alleles possible?

Four basic models are briefly described below, as well as their ability to describe adequately the observed allele frequencies and mutation rates.

1.12.1 *K*-allele model

This model assumes that only K different alleles are possible, and any mutations of a parent allele give rise to already existing alleles states (present in the population) – any parent allele has the same constant probability of mutating towards any already existing allele (Crow and Kimura, 1970). There are evidently size constraints on allele lengths as very long alleles are rare, so this model seems more realistic than the Infinite Alleles model, even though it is rarely cited in microsatellite evolution literature.

1.12.2 Infinite Alleles Model (IAM)

The Infinite Alleles Model developed by Kimura and Crow (1964) assumes that each mutation event creates a new allele – either a gene inherits an allele type of its ancestor (no mutation), or if a mutation does occur, the gene inherits an allele type not previously present in the population. There is assumed to be no constraint on allele size, hence an 'infinite' number of alleles is theoretically possible.

A mutation process that depends on the number of core repeats present as well as the tendency for microsatellite arrays to get longer (Markert *et al.*, 2001) makes it theoretically possible for there to be an infinite number of allele sizes. However, very long microsatellites are rare – generally, microsatellites have a limit to their length, usually not exceeding a few tens of core repeats (Estoup and Cornuet, 1999). Natural selection may prevent too long repeats from being maintained (a reflecting boundary), while an absorbing boundary would be one that induces very long repeats to degenerate into shorter ones through deletions. There may also be a 'centrally directed mutation bias' whereby long repeats contract, but short and medium-length ones expand. Much work favours various aspects of each argument/theory, but no conclusive explanation is yet available (Amos, 1999).

1.12.3 Stepwise Mutation Model (SMM)

Put forward by Kimura and Ohta (1978), this is currently accepted as the predominant mode of microsatellite evolution (Ciofi *et al.*, 1998), and is supported by investigations into the 'birth' of two separate microsatellites in hominoids (humans and apes) and monkeys (Messier *et al.*, 1996).

According to the SMM, short arrays of tandem repeats increase or decrease in number by one or two repeats as a result of replication slippage (Messier *et al.*, 1996) - each mutation is likely to change to its adjacent type. Many experiments have shown the large majority of microsatellite mutations to be the gain/loss of one or two repeat units (Amos, 1999).

New alleles at SSR loci have a similar number of repeats as the alleles from which they developed. The difference in the number of repeats can therefore be used as an indication of the length of time that has passed since the microsatellite loci have shared a common (same length) allele. It is acknowledged, however, that while these distances may not be accurate in every instance, they are useful for recreating phylogenies of relatively well-diverged taxa.

1.12.4 Two Phase Model (TPM)

This explanation of microsatellite evolution concedes that the mutational process of microsatellites in complex genomes is still not well understood, and so combines elements of the preceding two models – IAM and SSM – to form a more comprehensive model.

In the TPM (Di Rienzo *et al.*, 1994), most mutations are single-step changes (SSM), but rare large-scale changes in repeat number can also occur (providing a greater chance of a large number of alleles – IAM). Computer simulations have also been used to determine which theoretical model best fits observed data, and have confirmed that either TPM or the SMM generally fits microsatellite data best.

1.13 Microsatellites between species

Sequences adjacent to SSR motifs are generally highly conserved (Rossetto, 2001) enabling the design of site-specific primers that can be 'shared' between different species (or different populations of the same species) (Crespan and Milani, 2000). SSRs have shown successful cross-species amplification in a limited number of diverse taxa, including fish, rodents, birds and plants (Huang *et al.*, 1998). This usually requires a reduction in annealing temperature and an increase in the amount of template DNA used (Ciofi *et al.*, 1998).

There is often a tendency to select the longest repeats in the source species for use as markers (sampling bias). These are usually more informative than shorter sequences, but can reduce the number of informative markers in more distantly related species, because of the tendency for microsatellite alleles to be shorter in related species than in the species on which they were designed (Huang *et al.*, 1998; Amos, 1999; Cipriani *et al.* 2000; Rossetto, 2001)

However, it is expected that in a highly outcrossing genus like *Vitis* (grapevine) that a high percentage of microsatellite loci will nevertheless be informative in related species within the same genus (Huang, 1998). SSR markers developed on *Vitis vinifera* have amplified successfully on other *Vitis* genomes (Grando and Frisinghelli, 1998).

The possibility of transferring microsatellites from one core species to another in the same genus and even to similar genera would obviate the need to develop new markers for each, thus reducing costs. This would be particularly useful in *Vitis*, as many different species are used as rootstocks, though *V. vinifera* is generally used as the only wine or table grape species (Di Gaspero *et al.*, 2000). SSR primers can also be used in related taxa, a feature not shared by minisatellites (Ciofi *et al.*, 1998).

1.14 Functions of microsatellites

SSRs as markers have the advantage of being selectively neutral – they are not as yet known to confer any genetic advantage to an organism and neither is the presence of a microsatellite known to be disadvantageous to the fitness of an organism. Thus far, no conclusive evidence points to a specific function of microsatellites either (Galau *et al.*, 1976), and any supposed functions have been ascribed to them based on phenotypic defects that result when SSR loci mutate. Most work in this area has focussed on humans and animals, with a dearth of information available for plants.

The fact that SSRs are rarely found in the coding regions of the genome most likely has some evolutionary significance, as mutations in these types of repeats have the potential to result in frame shifts and can destroy otherwise functional genes (Hancock, 1999). This is evidenced by a number of human diseases associated with 'tri-nucleotide repeat expansion' – once a certain SSR in a gene expands beyond a 'healthy' threshold level, potentially fatal diseases result. These include Huntington's disease, Fragile X syndrome, Myotonic Dystrophy. Microsatellite instability is also associated with a number of cancers, the tissues of which show an accumulation of somatic mutations (Di Rienzo *et al.*, 1998; Klug and Cummings, 2000; Sturzeneker *et al.*, 2000). This is the first evidence of microsatellites having an effect on the phenotype of an individual (Hancock, 1999).

Microsatellites have been attributed with a number of functional roles, including sex determination, recombination hotspots and the regulation and expression of genes (Gupta *et al.*, 1994). These conclusions however, were not based on work involving higher-order plants.

It is commonly accepted that the conservation of sequences between species is an indicator of biological function or significance. In light of this fact, it is also interesting to note that microsatellite repeats have been found to have a marked effect on the effectiveness of promoter and enhancer regions in a wide variety of organisms, including *Drosophila*, *Aspergillus, Phythophthora infestans* and *Saccharomyces cerevisiae* (Hancock, 1999). Some short nucleotide motifs found in upstream activation sequences also serve to bind regulatory proteins/molecules. Constructs containing extra copies of repeat sequences often show increased activity when inserted into promoters of expression vectors. The length of the total repeat also affects the rate of transcription, with longer total lengths more effective than shorter lengths of the same repeat. It must be pointed out though that the promoter effect of these microsatellite repeats is generally mild (Hancock, 1999).

1.15 The Vitis Microsatellite Consortium (VMC)

The VMC was founded by researchers at UC Davis in 1997. Initially, there were 18 member laboratories, representative of all major wine-producing countries, but this number has since risen to over 24. Each research group is participating in a collaborative effort to develop a significant number of informative SSR markers on grapevine. These markers are to be used for grapevine genome mapping, and it is hoped that they will also be useful in identifying closely related varieties and cultivars.

There is a moratorium on publicly disclosing any findings until 2005, but member groups can freely share any information gained between each other. The primary aim of this effort is to develop a database of SSR allelic profiles against which any grapevine cultivar could be tested in order to determine its identity. Microsatellites are the ideal marker for this type of international group effort – highly informative, relatively simple and quick to use, reliable and reproducible. Thomas *et al.* (1994) was among the first to suggest the usefulness of this type of database for grapevine researchers. A further advantage is that not only can locus information (primer sequences) be shared between groups in a digital format, but allele lengths can also easily be expressed as quantitative data.

The expense and time-consuming effort required to develop primers was the very reason a 'group effort' approach was adopted – Agrogene SA, a commercial company based in

France specialising in plant diagnostic systems, co-ordinates the VMC and was instrumental in establishing a genomic library of *Vitis* and distributing SSR-enriched clones to VMC participants for sequencing. The cost of sequencing the entire library was therefore shared between all the participants. Raw sequences were then returned to Agrogene SA, who designed all the primers. This was done centrally in order to ensure consistency between loci. Reaction conditions were also recommended in order for the data at different loci to be comparable. These final primer sequences were then returned to the groups who had sequenced the clones from which they came, together with a small group of samples to be tested for polymorphisms.

Prof. Carole Meredith, a founding member of the Vitis Microsatellite Consortium, in here defence of microsatellites as the marker of choice, voiced concerns over the limitations of the AFLP technique. The primary limitation and most important with respect to the VMC is the difficulty in sharing AFLP data between different groups working with different mapping populations (Meredith, 1998). The main aim of the VMC, it must be remembered, was the pooling of information on the grapevine for the benefit of the international viticulture community. Although SSRs are expensive and time consuming to generate, once 'up and running' they are quick and easy to use, and still the only consistently reproducible markers which can easily be shared between groups in digital format (Meredith, 1998). This type of analysis, using discrete allele lengths, is ideal for a collaborative effort like the VMC's (Botta *et al.*, 1995). Additionally, SSRs allow for an increasing number of like markers to be incorporated into the database as the number of profiled cultivars increases. This will also allow for a more accurate determination of allele frequencies within a population (Brady *et al.*, 1996).

1.16 Uses of microsatellites in *Vitis*

The grapevine genome is typical of eukaryotic genomes in that it is rich in microsatellite regions, which are well represented in the grapevine genome despite its small size - approximately 475Mb (Lodhi and Reisch, 1995). The most common repeats in Vitis are $(GA)_n$ and $(GT)_n$, and are dispersed evenly throughout the genome. Longer dinucleotide repeats are genetically more variable than shorter repeats, and are therefore more informative markers. Longer repeats like CAC or GATA are also

present, but at a much lower frequency in the genome (Thomas *et al.* 1993; Thomas and Scott, 1993).

Of course, although the primary aim of the VMC was to establish an international database of SSR profiles to facilitate the identification of cultivars from the field and even within existing germplasm collections, the markers are by no means to be used exclusively for this purpose - having an established set of markers with the characteristics peculiar to microsatellites is a boon to researchers investigating a number of phenomena.

1.16.1 Population genetics

Different levels of polymorphism (number of alleles per locus) can give an indication of the level of diversity between different species, which is affected by genome structure and history of domestication. These markers can therefore give clues as to the structure and breeding history of a population depending on the ploidy levels and degree of polymorphism displayed in allele data (Cipriani *et al.*, 2000). In addition, by examining the tendencies of allele sizes and mutations within populations, information can be gained regarding the model of mutation at work.

1.16.2 Taxonomy and Phylogeny

Microsatellite flanking sequences can also be used to investigate the taxonomy and phylogeny of a species. Mutation processes not only affect the core repeat, but also the adjacent unique sequences, albeit to a lesser degree. It follows then that phylogenetic reconstruction between taxa amplifying the same marker loci is possible, and will give an indication of the accuracy of taxonomic classification of the respective taxa. Base pair changes and deletions can give an indication of genetic distances between species. This approach was used was used by Di Gaspero *et al.* (2000), who argued for the re-examination of the taxonomy of *Vitis*, as the phylogenetic trees which were reconstructed from the data did not agree with the current taxonomy of the *Vitis* genus.

1.16.3 Mapping

Molecular markers of all types are also essential to genome or physical mapping efforts. SSRs are ideal, as they can easily be exchanged between groups even if they are working on different crosses (Thomas and Scott, 1993). Data can also therefore easily be combined to give one unified map – there was at one stage 6 different mapping strategies (Reisch, 2000, Thomas *et al.*, 1994) – whereas this is more difficult when using RAPD and AFLP markers. Once a genome map is established, information on possibly syntenic crops could then be extrapolated. This is exactly the situation in the cereal crops – the rice genome was mapped and found to be syntenic with most other grasses, resulting in much faster progress in work on 'related' species (Hodgkin *et al.*, 2000; Ware *et al.*, 2002).

1.16.4 Parentage determination and Pedigree analysis

The co-dominant pattern of inheritance of microsatellite markers also allows for the reconstruction of crosses. A classic example is that of Bowers and Meredith (1997), where 15 microsatellite loci were used to show that the renowned red wine cultivar, Cabernet Sauvignon, resulted from a cross between Cabernet Franc (red) and Sauvignon Blanc (white). This was independently confirmed by Sefc *et al.* (1997), using more SSR markers. The genetic relationships between several European cultivars, as well as the parentage of a number of common cultivars were subsequently determined, which allowed for a tentative pedigree to be drawn up (Regner *et al.*, 2000a; Sefc *et al.*, 1998c).

Sefc *et al.* (1997) further expanded on previous work (Regner *et al.*, 2000a) and demonstrated the suitability of SSRs for parentage analyses using Austrian cultivars. The usefulness of microsatellites in parentage analyses was confirmed by Crespan and Milani (2000) with work on Italian cultivars. Similar research was also successfully performed using SSRs in Australian *Vitis* germplasm samples (Thomas *et al.*, 1994), and Bowers *et al.* (2000) found the single set of parents of a group of French grapevine varieties using the same techniques.

Pedigrees are especially useful in determining what breeding steps created other cultivars, and are therefore useful to growers as they provide information about the segregation of important traits, as well as emphasising the importance of keeping detailed breeding records (Regner *et al.*, 2000a). They provide useful

information on previously successful hybridisation strategies, as well as an insight into the development of a number of very similar cultivars in a particular region (Regner *et al.*, 2001). Furthermore, microsatellite analysis of a variety of European cultivars has also shown that the geographic origin of a particular cultivar can usually be derived from its genotype (Regner *et al.*, 2000a; Sefc *et al.*, 2000b).

It is clear then, that pedigrees can provide valuable information regarding the origins of cultivars still in use today and which have existed for centuries (Regner *et al.*, 2000a). There are still a number of questions which may be answered on the basis of pedigree analysis, such as whether crosses display a certain 'preference', or what the basis is for the tremendous morphological variation between the thousands of cultivars known today (Regner *et al.*, 2001).

1.16.5 Marker Assisted Selection and Breeding

The characterisation of varieties with respect to a certain trait is important for any selection or breeding programme and the subsequent agricultural/commercial use of such varieties (Monteiro *et al.*, 2000). MAS is ideally suited to long-cycle crops like grapevine - if important genes are associated with a specific molecular marker, early selection of seedlings can take place without waiting for the plant to develop to maturity (approximately 3 years). Pyramiding of multiple genes for a single trait is also possible using molecular markers.

The tying of a marker to a specific trait goes hand in hand with genetic mapping, and as such requires intensive labour and screening practices. Fortunately, there are a number of research groups in France, Australia, Germany and the USA that have mapping projects in progress (Reisch, 2000). Since each is concentrating on different traits, an extensive map is likely once these results are combined. Such molecular markers would facilitate the selection of favourable characteristics like disease resistance, drought hardiness, seedlessness or flower sex, and are therefore of significant agricultural and economic importance.

While any molecular marker that is linked to specific trait can be used in the preferential selection of genotypes, microsatellites are one of the simplest and easiest to use. Dalbo (1998) used SSRs in conjunction with other molecular markers to map Quantitative Trait Loci involved in fungal resistance (powdery mildew) and morphological characters in grapevine. A number of other QTLs have been identified for other traits using different marker systems, like RAPDs (for mildew resistance)(Buck and Zyprian, 2000), and AFLPs (for seedlessness) (Scott *et al.*, 2000a).

These molecular marker maps should facilitate gene cloning, in turn making successful genetic transformations a possibility (Reisch, 2000). Marker maps will also obviously be of great agricultural importance – breeding programmes will be able to advance to the field trial stage far faster than before as a result of seedling selection and specific cultivars will be able to be screened for desirable traits and hybrids created.

However, any newly developed cultivars will still have to be tested in the field over several seasons before they can be verified as safe and their potential proven. There is therefore still a need for classical breeding techniques, as not all complex traits can be dissected using only new molecular approaches – the two approaches should complement each other.

1.16.6 Conservation

A great advantage of SSRs in conservation genetics is their transferability between species and even across taxa. This is valuable in ecology and the protection of endangered species, as it has the potential to reduce the extremely high costs of establishing a new set of markers for a given species from scratch (Monteiro *et al.*, 2000).

For example, while most wine today is produced from *Vitis vinifera* cultivars, very few wines are actually grown on *Vitis vinifera* rootstocks. The use of SSR markers developed on wine grape cultivars in rootstock species is common amongst grapevine researchers (Sefc *et al.*, 2000a), and confirms the effectiveness of these markers between different species. From our work, it also became evident that markers (ZAG67 and ZAG112) developed on the rootstock

species *Vitis riparia* (Sefc *et al.*, 1998c) were equally suitable for use on wine grape cultivars.

The somatic stability of microsatellites has not been thoroughly researched, and little information is available on plant genomes (like the grapevine), which have been vegetatively propagated for long periods. However, over only a few generations microsatellites are more stable than expected. This poses a problem for their use in quite recently diverged cultivars, as they may then be less informative than for cultivars that diverged significantly earlier (Cipriani *et al.*, 2000). This should be kept in mind when SSRs are used for population and diversity analyses.

1.16.7 Germplasm management and Diversity assessment

Another use of molecular markers is in the management of germplasm repositories. These are essentially greenhouses in which, ideally, a representative of each cultivar and clonal selection is maintained. Each individual plant has a complete set of documentation, detailing its breeding history, virus status etc.

Besides the costs associated with obtaining a complete set of information regarding a particular accession's identity and history, there are substantial costs involved in the physical maintenance of such collections – the entire area must be maintained at an optimal temperature, the plants must be protected from fungal and viral infections and each entry must independently be identified as belonging to a specific species, cultivar and clone. This has been a daunting task in earlier times, relying on sometimes questionable ampelographic classifications, as well as being vulnerable to previous, less accurate molecular techniques like isozyme analysis. Often, incomplete paperwork for a particular accession has also resulted in the mis-classification of the vine, which has gone unchecked for many years. This has potentially far reaching repercussions, as members in the user community tend to request vines known for specific qualities (Lamboy and Alpha, 1998).

The following highlights the extent of the problem: There are many instances where the same cultivar is known by several different names (synonyms) especially in Europe, where each region and language area has a different common name for a particular grape. There are also numerous instances of different vines being known by one name (homonymy) (Sefc *et al.*, 2000a; Bowers *et al.*, 1996; Cervera *et al.*, 1998, Borrego *et al.*, 2000). It is not only of commercial importance to clarify such matters, but also essential if the true level of genomic diversity is to be determined. Such information is also useful as far as conservation efforts are concerned – a balance must be found between the financial costs of maintaining lesser-known cultivars that contribute to the diversity of the total vine population, and the monetary costs of retaining or acquiring duplicates of such accessions in germplasm banks.

To this end, SSR markers are especially useful as unique genetic profiles can be obtained. SSRs can therefore be applied in analyses of genetic diversity and redundancy in core collections, where each accession should have a unique genetic profile against which new acquisitions can be compared. Characteristics unique to microsatellites facilitate the sharing of such distinctive profiles between core collections, and can therefore also aid in minimising redundancies between different collections. Furthermore, the importance of maintaining even minor cultivars in collections is emphasised by the fact that the success of pedigree analyses (See 1.15.4) using SSRs depends on the availability of ancient cultivars (Regner *et al.*, 2000a).

1.16.8 Cultivar identification

The viticulture industry is now international, and consumers worldwide should enjoy the same protections regardless of the origins of their wine, grapes or brandy. The need for a user-friendly, reliable method of cultivar identification is thus obvious - it is not enough to simply determine what a vine is *not* (*not* a Pinot, *not* a Chardonnay), there is an urgent need to be able to confidently identify an unknown sample as a specific cultivar or clone. In Europe, strict legislation controls the content and labelling of wines. Nurseries must also be confident in their product classifications, and consumers have a right to know that what they are buying is accurately described on the label.

Besides the clear commercial requirements, Ibañez (2000) highlights the importance of accurate varietal identification as having serious legal implications; if a patented variety is to be protected, the owner/developer must be able to unambiguously identify his plant variety from all others commercially available. Detailed mathematical analyses are available to evaluate the accuracy of such assignments when molecular markers are used. Legal disputes may result when expert ampelographers make such assignments; molecular marker tests are inherently more reliable and trustworthy. Even so, cases like this are fraught with difficulties – if too few markers of any type are used, or if incorrect reference samples are used for comparison, results may be reasonably queried. Again, the importance of an exhaustive germplasm collection is reiterated.

The ability to amplify single loci from DNA of grapevine using marker technology also potentially allows for the monitoring of grapevine for the presence of transgenic sequences (Sefc *et al.*, 2000a). This is especially significant now, as there are currently negative perceptions with respect to genetically modified organisms. Unfortunately, grapevine is susceptible to a variety of fungal and viral diseases, as well as insect parasites. In order to combat these problems and reduce the use of environmentally damaging and expensive pesticides and herbicides, researchers are looking at transgenics. Of course, these altered vines are unlikely to be welcomed by all sectors, and so the tracing and labelling of modified vines and derived products further protect consumer rights and indirectly, producers, from any possible legal actions.

Compared with RAPDs and AFLPs, where the former requires all samples to be analysed in order for comparisons to be made between samples and the latter requires many more steps in generating the pattern of bands required for analysis, microsatellites call for the least number of reactions/reaction steps before distinct genotypes can be discerned, and are thus the best as far as fingerprinting is concerned (Regner *et al.*, 2000b) – 90% of Spanish accessions studied were distinguished on the basis of a single microsatellite locus (Borrego *et al.*, 2000). SSRs have also been found to be the most robust of the three, and

were recommended for the analysis of clonal diversity in grapevine by Merdinoglu *et al.* (2000)

Regner *et al.* (2001) found the 6 most informative microsatellite markers for grapevine, which have served to distinguish between all cultivars analysed thus far. It is therefore possible that previous efforts in distinguishing between cultivars have been hampered by the inclusion of less polymorphic markers in the analyses. Of course, the informativeness of a given marker can only be determined empirically, and any given group would have to screen many markers on a large number of different samples before a conclusion could be reached.

There are conflicting reports on the effectiveness of microsatellites in distinguishing between clones of the same cultivar, but it seems unlikely that genetic differences that result in nearly identical clones of a cultivar will be routinely detected by such a system, which detects genetic changes in very specific (non-coding) regions of the genome. Theoretically, it makes more sense that techniques that screen a larger portion of the genome (AFLPs and ISSRs) are more likely to detect such miniscule changes, and this logic seems to bear out in practice.

Although Vignani *et al.* (1996) separated several clones of 'Sangiovese', and five clones cultivar 'Fortana' were distinguished by Silvestroni *et al.* (1997), Sánchez-Escribano *et al.* (1995) were unable to detect any clonal differences in 'Napoleon'. This could be explained by the fact that the two former cultivars are both ancient and probably derive from closely related zygotically derived genotypes – i.e. they are of polyclonal origin. 'Napoleon' was an autochthonous cultivar (all clones originated in the area in which they were eventually found) and any somatic mutations were probably undetectable using SSRs, a conclusion supported by Cipriani *et al.* (1994).

The detection of clonal differences on the basis of SSR polymorphisms alone seems to be the exception rather than the rule. (Monteiro *et al.*, 2000) used microsatellites to distinguish between multiple clones of 31 Portuguese varieties, but while the varieties could be separated; the three clones of each could not.

Sefc *et al.* (2000a) also used microsatellites, and were slightly more successful – only two rootstock cultivar clones of 30 were indistinguishable. In addition, of nearly 100 wine cultivars, only the (notoriously difficult to distinguish) three Pinot colour variants (Pinot blanc, Pinot noir and Pinot gris) were impossible to tell apart. Interestingly enough, even after these were analysed using AFLPs, their genotypes were still identical. This confirms previous work on this group of clones, which have not yet been distinguished using any molecular system (Ye *et al.*, 1998).

Berry colour types amongst the Pinot group, and between the 'Verdelho' grown in Madeira and the Azores have proved to be indistinguishable on the basis of microsatellite analysis. It is suspected very small somatic mutations are responsible for these differences, most likely in the coding region of the genome (Lopes *et al.*, 2000; Thomas and Scott, 1993).

While a combination of ampelography, isozymes and microsatellite analysis has proved to be highly effective when used together, as they each have different strengths and so together give an accurate indication of the situation at varietal level (Weihl and Dettweiler, 2000), this is not the case at the clonal level. Crespan and Milani (2000) used isozymes, RAPDs and microsatellites to distinguish twelve seeded and seedless cultivars and none was able to distinguish between clones.

Having said that, it is seems only sensible to use more than one marker system in the identification and classification of samples, especially if a legal issue may plausibly arise, as then any shortcomings with one in the identification of some samples will be ameliorated by the use of another, more informative marker (Hodgkin *et al.*, 2000). Moreover, markers should not be seen as independent of and always superior to traditional techniques, but rather be used in conjunction with ampelography (Borrego *et al.*, 2000).

1.17 Project background

The Institute for Wine Biotechnology (IWBT), University of Stellenbosch became a member of the VMC in 1997. The gDNA library clones were processed as described in

1.14 by Ms Kathy Wilsen in 1998, and by 1999 five microsatellite loci were 'owned' by the IWBT.

Agrogene SA required that these primers be screened in seven gDNA samples they provided, and questions regarding the polymorphism levels of the markers answered. This work was done through collaboration between the IWBT and the Department of Genetics, Stellenbosch. It was subsequently decided that this simple and efficient method be expanded to include cultivars of particular relevance to South African industry. Consequently, samples were obtained from the core germplasm collection of KWV, a major distributor of grapevine propagation material in South Africa, as well as an international exporter of wines and brandies. These included several wine and rootstock cultivars and clones. An additional five primers were also selected from the total available to all members of the VMC, in order to test polymorphism at a significant number of loci.

1.18 Project aim

The aim of this project was to screen all the samples supplied by KWV and to determine whether they were distinguishable on the basis of the selected microsatellite markers.

Ideally, a unique 'fingerprint' for each cultivar was sought. The hope that clonal differences were also detectable was not a primary concern, but became more important once differences at a few loci in a few samples were detected in preliminary screenings. Eventually, the scope of the project expanded to include this (ambitious) aim. In the final analysis, we sought a database of allelic profiles for each sample at the loci considered, which would allow us to compare an 'unknown' sample with the combined profiles and make a suggestion as to its identity.

Further, we wanted to derive a sub-set of the primers used that would allow for the distinction of as many cultivars as possible from our sample set. This aim was particularly important, but an additional benefit would also have been to find a set of primers that would allow us to distinguish between *clones* of a specific cultivar (this as a result of our preliminary findings).

Finally, we wanted to ensure that our methods yielded the same results for common marker/sample combinations that were used by other groups involved. This would prove the reliability of the marker system, and confirm the reproducibility of results between the different laboratories participating in the VMC.

Chapter 2

Materials and Methods

2.1 Plant Material

The bulk of the plant material for this study was supplied by the KWV ('Kooperatiewe Wynbouers Vereeniging) Pty Ltd., South Africa. The plants were selected from their core germplasm collection in Paarl, Western Cape, by Mr. Tobie Oosthuizen and were chosen as being of particular interest to KWV as far as their identities were concerned.

Five samples were supplied (as extracted gDNA) by Agrogene SA, Moissy Cramayel, France as part of our collaborative effort as member of the Vitis Microsatellite Consortium (VMC).

Three samples were supplied as extracted gDNA by the Institute of Wine Biotechnology, University of Stellenbosch.

Altogether, these samples comprised cultivars and clones of *Vitis vinifera*, *V. riparia* and *Muscadinia* species, as shown in Table 1 (pp. 54, 55).

2.2 Genomic DNA extraction and quantification

Approximately 5 young, fully opened leaves for each plant were collected from the KWV core germplasm collection during the early afternoon, and stored in 15x15cm plastic sandwich bags, with the KWV planting codes as identifiers. These bags were transported to the laboratory in cooler bags with ice packs, where they were stored at –20°C until needed.

DNA extractions were performed according to the following protocol by Kim *et al.* (1997), with minor modifications, as other protocols yielded little DNA or DNA of inferior quality (Edwards *et al.,* 1991; Lodhi *et al.,* 1994). This problem has previously

been reported by several groups, including Porebski *et al.* (1998), Bowers *et al.* (1993) and Kim *et al.* (1997). As yet, there is still no standard DNA extraction protocol for grapevine.

Four or five leaf discs were punched out of each leaf using the lid of a 1.5ml Eppendorf tube. Five microlitres of 1% (v/v) 2-mercaptoethanol and a pinch of silicon carbide C400 powder ('carborundum') was added and the leaf discs were then manually ground to a pulp with plastic conical grinders. Next, 250µl of Extraction buffer (200mM Tris-HCl (pH 8.0 at 20°C), 250mM NaCl, 25mM EDTA and 0.5% SDS) was added and mixed well with the pulp. A further 250µl extraction buffer was then added, together with 2% (w/v) final volume polyvinylpyrrolidone (PVP) (MW = 10 000). This was mixed together well, then incubated in a 60°C water bath for at least 1 hour.

Samples were removed from the heat and allowed to cool for 5 min at room temperature. One half volume of 7.5M NH_4OAc was then added to the emulsion, which was left to incubate at –20°C for at least 30 min., then centrifuged at 4°C, 10 000g for 10min. The supernatant was then transferred to a new tube and an equal volume of isopropanol (iso-propyl alcohol) was added to the supernatant. This was left at –20°C overnight, to precipitate the DNA out of solution.

The suspension was then centrifuged at 4°C, 10 000g for 15 min, the supernatant discarded and the pellet left to air-dry. Pellets were re-suspended in 50µl ddH_2O, 5µl of which was run on a 1% agarose gel, with 1 x TBE running buffer at 80V for 30 min to check whether sufficient DNA had been extracted thus far. The gel was stained with ethidium bromide and viewed using UV light. If too little DNA was present at this stage, the extraction was re-done for that sample; otherwise extractions were continued as follows:

The remaining 45µl was treated with RNase A (0.004mg/ml), (Roche) and incubated at 37°C for 15 min. One volume of chloroform: isoamyl alcohol (24:1) was added to each sample, and gently shaken. The mixture was then centrifuged at 4°C, 10 000g for 10 min and the supernatant transferred to a new tube. One volume of isopropanol was again

added to this, and precipitated for at least 30 min at –20°C. This was centrifuged at 4°C, 10 000g for 10 min. The supernatant was poured off, and the pellet washed once with 1ml ice-cold 80% ethanol. The pellet was then left to air dry, before re-suspending in 50μl ddH_2O. DNA was stored at 4°C.

The extracted DNA was quantified on an agarose gel (1.8%, 1xTBE), using a known concentration gradient of lambda DNA for comparison. The gel was stained with ethidium bromide (EtBr) (0.005mg/ml), and viewed using ultraviolet light. The approximate DNA concentrations were determined, and then standardised to 25ng/μl per sample.

2.3 Primers

Primer sequences for this project were supplied by Agrogene, France. They designed all primers from sequences obtained from each VMC member laboratory, which in turn were derived from a microsatellite-enriched library based on *Vitis vinifera* cv. Sultana. Our laboratory had the primers manufactured according to their speçifications, by Gibco BRL. Primer sequences are as given in Table 2.

All polymerase chain reaction conditions were optimised using unlabelled primers and visualized on 1.8% agarose gels in 1 x TBE buffer with EtBr staining, and viewing through UV light. These conditions were then maintained when working with ^{33}P.

2.3.1 Primer labelling

Three hundred nanogrammes of each forward primer was end-labelled with 0.5μl γ-^{33}P-ATP (NEN™, Boston, USA Life Science Products, Inc.) in a 10μl reaction, containing 1μl 10X One-Phor-All reaction buffer (250mM imidazole (pH 6.4), 60mM $MgCl_2$, 5mM 2-mercaptoethanol, 350μM ADP), 0.5 10X T4 polynucleotide kinase (9 800 U/ml) (USB), and made up to 10μl with ddH_2O. This was incubated at 37°C for 1 hour. The reaction was then stopped by incubating at 65°C for 10min. Primers were stored in a lead container at 4°C until needed.

Table 2. Primer details, with amplification conditions used in 33P analysis

Locus	Sequence	T_A (°C)	$MgCl_2$ (mM)	Expected size (bp)	Repeat motif
1a2	(F) 5'-TAAAATGTAGGGCGGCCACC	61	1.50	128	$(GA)_{19}(GATA)_6$
	(R) 5'-AACATAAATGGCCACCAGGG				
1a5	(F) 5'-GGCTACAACAACTCTATTACCT	57	1.50	203	$(GA)_{19}$
	(R) 5'-AATATTCTCATTTTTTCCCTC				
1g3.2	(F) 5'-GATAGTTACCATACTTAGTCGGA	51	1.00	131	$(CA)_6(GA)_{17}$
	(R) 5'-ACTTAGCTTCAGAAGAAAATAGA				
1h5	(F) 5'-AAAAACCTCACCAAAACTCTTC	57	1.50	126	$(GA)_{14}$
	(R) 5'-TAGAGTAGTTTGCTCTCTCCCC				
3c7	(F) 5'-AATTTGGCTAAGAAAGGA	58	2.00	123	$(GA)_{16}$
	(R) 5'-AATATTCAGAAAATTTGTGTC				
2e'8	(F) 5'-TCTTACAATCAGGTTGTGTTAC	54	2.50	104	$(GA)_{29}$
	(R) 5'-TCATCTTCCTTAAAACTTATCC				
3d8	(F) 5'-AAACCAAACGGAAAAAT	51	2.00	63	$(GA)_{14}$
	(R) 5'-ACCTTCCCTTTCAATCA				
16a1	(F) 5'-AATTAGTTTCTAATAATGCAGGA	51	2.00	90	$(TC)_{20}$
	(R) 5'-GTGAGAGAACAGGATGGTAA				
4c6	(F) 5'-CTCCATCCCTATCTCATCAG	50	1.00	175	$(GCT)_{11}$
	(R) 5'-CTCTAACACCCAATCTCACA				
3g11	(F) 5'-TTTCCAGAATTTTCAGAGAGTC	60	1.00	206	$(TC)_{23}$
	(R) 5'-TAGAGGCAAAGAAAAGCTACAA				
3f12	(F) 5'-GCTTTGAATGAACACTGTTATG	60	1.00	248	$(TC)_n$*
	(R) 5'-ATGAAGCTGGAAAAGAATAGAA				
16c1	(F) 5'-CGCATTACATATTCAATTTCCT	58	2.50	146	$(TC)_{29}$
	(R) 5'-TGAAGTGCTGTTTGAAGAGAGT				
4d3	(F) 5'-TTAGTTTCTAATAATGCGGGA	59	2.50	121	$(TC)_{19}$
	(R) 5'-TCTTATGATTATGATGCACCA				
4g6	(F) 5'-CCTTGAAGAGATGAGTTTGCTA	61	1.00	136	$(GA)_{17}$
	(R) 5'-TATTTAACTTTGTGCCTCTGCT				
ZAG 67†	(F) 5'-ACCTGGCCCGACTCCTCTTGTATGC	49	1.00	121 - 140**	$(GA)_{18}(TA)_{10}$
	(R) 5'-TCCTGCCGGCGATAACCAAGCTATG				
ZAG 112†	(F) 5'-TGGCTCCATACTGCTTCACGTAGGC	49	1.50	243	$(TC)_{16}$
	(R) 5'-CGTTTAAAGCCAGCTGAATCTTGGG				
5g6.1	(F) 5'-TTCTAAGACAGAATTGCTTGGC	63	2.00	147	$(TC)_{24}$
	(R) 5'-TTATCTGTAGCTTTCACACCCC				
2f10	(F) 5'-AGATTCTTCTGATGGTGTTGGG	55	2.00	101	$(GA)_{15}$
	(R) 5'-ATCAGAGCTCCTCTTTCCTTCC				

†= Developed on *V.riparia* (Sefc *et al.*, 1997), not as part of the VMC
*= Number of repeats unknown
**=Empirically determined in our laboratory

2.3.2 Ladder labelling

A 10bp ladder (Gibco BRL) was used, with bands at 10bp increments from the 10bp to 330bp positions. The 100bp band stains more intensely than the other bands to allow for internal orientation. This was labelled as follows: 4μl of ladder (0,1μg/ul) was used in a 9μl reaction with 1μl 10X One-Phor-All buffer, 2μl T4 PNK (9 800U/ml) and 2μl γ-^{33}P-ATP (10mCi/ml). The labelling reaction was incubated at 37°C for 1 hour. The reaction was then stopped by incubating at 65°C for 10min. and stored in a lead container at 4°C until used.

2.3.3 PCR amplification

Each primer had previously been optimised for annealing temperature and $MgCl_2$ concentration on agarose gels. These conditions are as given in Table 2.

Twenty-five nanogrammes of template DNA was used in each 25μl amplification reaction, with 2.5μl 10X NH_4 reaction buffer (100mM Tris-HCl (pH 8.3), 500mM KCl), 50μM each dNTP, 30ng each primer and 1U *Taq* DNA polymerase (AmpliTaq Gold™, Perkin Elmer Applied Biosystems, Inc., Germany). Reactions were made up to 25μl using ddH_2O.

Initially, AmpliTaq Gold™ was used, as it has no residual activity at room temperature (manufacturer's accompanying literature), therefore giving fewer problems with non-specific primer annealing and non-templated A-extension. However, since the first few gels analysed showed many stutter and over-stutter bands (typical of microsatellites, especially dinucleotide repeats), (Ciofi *et al.*, 1998; Brady *et al.*, 1996), it was decided to continue the work using BIOTAQ™ DNA *Taq* polymerase (Perkin Elmer Applied Biosystems, Inc., Germany), which was much cheaper and gave results of comparable quality.

All amplifications were performed according to the following programme, on an Eppendorf Mastercycler Gradient 5331 thermal cycler. Amplifications were replicated at least twice for each sample/primer combination.

95°C 7 min

94°C	1 min	x35
T_A	1 min	
73°C	1 min	

72°C	5 min
4°C	hold

PCR products were stored at 4°C until electrophoresed.

2.3.4 Electrophoresis

Five microlitres of loading buffer (95% formamide, 0.01N NaOH, 0.02M EDTA (pH 8.0), 0.025% bromophenol blue and 0.05% xylene cyanol) was added to each sample. Loading buffer was added to the ladder in a 1:1 ratio. The DNA samples and ladder were then denatured together with the samples at 95°C for 5 min and immediately transferred to ice.

Prior to loading the samples on the gel, the gel was pre-run at 60W for 1 hour to equilibrate.

Five microlitres of each sample was loaded on to a 30cm x 40cm 4% denaturing polyacrylamide gel (76g polyacrylamide: 4g bis-acrylamide) containing 7M urea (M.W. = 60.06, Promega) and run with a 1xTBE buffer in a vertical gel tray unit (Model S2001, Life Technologies) at a constant power of between 60-80W, until the bromophenol blue dye band had just run off the gel. Five microlitres of the 10bp ladder was loaded in the first lane, the last lane and at approximately 10 lane intervals between the samples in the gel for sizing purposes, in agreement with the methods of Monteiro *et al.* (2000).

The gel was then transferred to 3mm thick blotting paper (Machery Nagel, Germany), covered in household cling film, and dried on a Savant Slab Gel Drier for 90min at 80°C. The dried gel was then exposed to a 30cm x 40cm Kodak™ Biomax MR film for between 2 and 5 days in 35cm x 43cm Hypercassettes™ (Amersham Life Sciences, England),

depending on level of radioactive signal emitted. The films were then developed in First Graphics Imaging Systems developing solutions according to manufacturer's instructions, and air-dried.

2.3.5 Data analysis

Autoradiographs were manually scored using the ladders on either side of any given sample. Alleles were scored as base pair lengths (Thomas *et al.,* 1994). When there was any doubt as to homo- or heterozygosity for a null allele, samples were scored as heterozygotes, according to the convention adopted by Thomas *et al.* (1994). Results were stored in a Microsoft Excel spreadsheet.

Table 1. Species and cultivars supplied from designated sources for this project

Code	Species	Cultivar and clone no.	Source	Supplied as
A1	*V. vinifera*	Alicante Bouschet 2	KWV	Plant Material
A2	*V. vinifera*	Auxerrois 6	KWV	Plant Material
A3	*V. vinifera*	Bianca	KWV	Plant Material
B4	*V. vinifera*	Ferdinand de Lesseps 2	KWV	Plant Material
B8	*V. vinifera*	Jacquez 6	KWV	Plant Material
C4	*V. vinifera*	Pinotage 48	KWV	Plant Material
D4	*V. vinifera*	Roobernet 1	KWV	Plant Material
A4	*V. vinifera*	Cabernet Franc 1	KWV	Plant Material
A5	*V. vinifera*	Cabernet Franc 623	KWV	Plant Material
B9	*V. vinifera*	Malbec 1	KWV	Plant Material
B10	*V. vinifera*	Malbec 71	KWV	Plant Material
D2	*V. vinifera*	Ruby Cabernet 1	KWV	Plant Material
D3	*V. vinifera*	Ruby Cabernet 170	KWV	Plant Material
D5	*V. riparia*	Richter 99, 13	KWV	Plant Material
D6	*V. riparia*	Richter 99, 179	KWV	Plant Material
E1	*V. vinifera*	Shiraz 21	KWV	Plant Material
E2	*V. vinifera*	Shiraz 99	KWV	Plant Material
E8	*V. vinifera*	Viognier 1	KWV	Plant Material
E9	*V. vinifera*	Viognier 642	KWV	Plant Material
A6	*V. vinifera*	Cabernet Sauvignon 163	KWV	Plant Material
A7	*V. vinifera*	Cabernet Sauvignon 27	KWV	Plant Material
A8	*V. vinifera*	Cabernet Sauvignon 33	KWV	Plant Material
B5	*V. vinifera*	Semillon 1	KWV	Plant Material
B6	*V. vinifera*	Semillon 121	KWV	Plant Material
B7	*V. vinifera*	Semillon 14	KWV	Plant Material
C1	*V. vinifera*	Merlot 182	KWV	Plant Material
C2	*V. vinifera*	Merlot 346	KWV	Plant Material
C3	*V. vinifera*	Merlot 447	KWV	Plant Material
D7	*V. vinifera*	Sauvignon Blanc 1	KWV	Plant Material
D8	*V. vinifera*	Sauvignon Blanc 159	KWV	Plant Material
D9	*V. vinifera*	Sauvignon Blanc 2	KWV	Plant Material
D10	*V. vinifera*	Sauvignon Blanc 316	KWV	Plant Material
A9	*V. vinifera*	Chardonnay 3	KWV	Plant Material
A10	*V. vinifera*	Chardonnay 277	KWV	Plant Material
B1	*V. vinifera*	Chardonnay 55	KWV	Plant Material
B2	*V. vinifera*	Chardonnay 76	KWV	Plant Material
B3	*V. vinifera*	Chardonnay 95	KWV	Plant Material
E3	*V. vinifera*	Chenin Blanc 1	KWV	Plant Material
E4	*V. vinifera*	Chenin Blanc 1061	KWV	Plant Material
E5	*V. vinifera*	Chenin Blanc 24	KWV	Plant Material
E6	*V. vinifera*	Chenin Blanc 550	KWV	Plant Material
E7	*V. vinifera*	Chenin Blanc 9	KWV	Plant Material

Table 1. (Cont.)

Code	Species	Cultivar	Source	Supplied as
C5	*V. vinifera*	Pinot Noir 111	KWV	Plant Material
C6	*V. vinifera*	Pinot Noir 113	KWV	Plant Material
C7	*V. vinifera*	Pinot Noir 115	KWV	Plant Material
C8	*V. vinifera*	Pinot Noir 459	KWV	Plant Material
C9	*V. vinifera*	Pinot Noir 667	KWV	Plant Material
C10	*V. vinifera*	Pinot Noir 777	KWV	Plant Material
D1	*V. vinifera*	Pinot Noir 9	KWV	Plant Material
S	*V. vinifera*	Sultana	IWBT	Extracted DNA
RG	*V. vinifera*	Red Globe	IWBT	Extracted DNA
D	*V. vinifera*	Dauphin	IWBT	Extracted DNA
Cab Sauv	*V. vinifera*	Cabernet Sauvignon	Agrogene	Extracted DNA
Ries	*V. vinifera*	Riesling	Agrogene	Extracted DNA
Rip	*V. riparia*	unknown	Agrogene	Extracted DNA
Sult	*V. vinifera*	Sultanine	Agrogene	Extracted DNA
Musc a	*Muscadinia*	unknown	Agrogene	Extracted DNA

Chapter 3

Results and Discussion

3.1 Samples studied

Initially, we received 5 samples from Agrogene, which we were to screen using 'our' primers – 1a2, 1a5, 1g3.2 and 1h5. These were obtained from the sequences of the clones we were assigned by Agrogene at the start of our involvement with the VMC. In an effort to standardise the methods and results obtained by VMC members, Agrogene dictated much of the approach to practical work (see chapter 2). We also had clear directives from Agrogene concerning the interpretation of our results. Products could be separated using any method, as long as allele sizes were provided. In addition, main questions to be answered included whether a given locus was polymorphic; whether there was more than one major PCR product visible; whether the allele pattern was clearly scorable and finally, whether there was more than one polymorphic locus present (P. Isaac, pers. comm. 1999).

Following this analysis on only 5 samples, we decided to broaden the study to include a number of locally grown cultivars and also to increase the number of loci considered. The remainder of this chapter deals with the results obtained from this investigation, and how they compare with data obtained by other groups, as well as their applicability to South African viticulture.

The samples under study included wine, table and rootstock cultivars, as well as a number of clones. Samples were grouped according to their origin. As explained in the text, these three groups were alternately pooled or separated, depending on the parameter measured. Table 1 (pp. 54, 55) is a complete list of all the samples used.

3.1.1 IWBT samples

Three samples from the table grape cultivars Sultana, Red Globe and Dauphin were obtained from the IWBT, and placed in a separate group to the other rootstock and wine cultivars.

3.1.2 VMC samples

Five samples were obtained from Agrogene and constituted our part in the VMC effort. At the time, only four marker loci had to be tested so we received only approximately 20µl of extracted DNA per sample. The result was that a total of only 11 markers could be tested. Comparisons with the other sample sets were made accordingly.

3.1.3 KWV samples

These constituted the bulk of the work (49 of 57 samples), and form part of KWV's core germplasm collection, which supplies material that is to be vegetatively propagated for the establishment of vineyards and the like. Our samples were specifically selected for investigation because of their particular significance to KWV; they can therefore not be considered representative of either the collection or any wild population. An important implication of this is the limitation placed on the extrapolation of data from this study to other collections and to wild populations.

3.2 DNA Extraction

Although DNA was finally extracted using a protocol modified from Kim *et al.* (1997), various other extraction methods were tried, including those of Edwards *et al.* (1991) and Lodhi *et al.* (1994). This confirmed that the extraction of significant amounts of good quality, clean DNA from grapevine is no easy task. This is primarily the result of high quantities of polysaccharides and polyphenols present in grapevine, but can also be complicated by using older leaf material (Lodhi *et al.*, 1994; Wang *et al.*, 1996; Porebski *et al.*, 1998). Fortunately, large quantities of DNA, while necessary for work involving restriction enzyme digestion (like AFLPs or minisatellites), is not a prerequisite for PCR-based marker systems like microsatellites (Balding, 1999). Amplifications were therefore successful despite extracting quantities of DNA from

between 100ng/µl and 750ng/µl per sample. Yields were however not consistent between samples. For example, Chenin Blanc typically yielded more, better quality DNA than either Chardonnay or Cabernet Sauvignon for any extraction method used.

3.3 Aims

The primary focus of this project was the typing of KWV cultivars at loci obtained via the VMC. The hope was that a database could be established which would allow for the detection of duplicates in the collection, as well as the identification and classification of unknown cultivars by comparison with known cultivar fingerprints. KWV would obviously benefit from any findings – germplasm collections are expensive to maintain, and the identification of grapevine cultivars by other methods is tedious and often unreliable (Hodgkin *et al.*, 2001; Lamboy and Alpha, 1998).

3.4 Method

As Ulanowsky *et al.* (2001) pointed out, the choice of marker used for a project depends not only on the purpose of the project, but is also affected by budgetary and time constraints, ease of use and technical limitations, such as the availability of facilities. Microsatellites are known for their capacity to discriminate between varieties (Borrego *et al.*, 2001), and are today the method of choice for DNA fingerprinting across a wide range of species because of their high level of polymorphism, repeatability and ease of use (Hinrichsen *et al.*, 2000). Furthermore, since most grapevine cultivars are highly heterozygous, the number of alleles at a given locus continues to increase. This has the result of making SSR loci progressively more informative within a breeding population, and therefore increasingly useful in cultivar identification and pedigree analyses.

Of the 48 grapevine gDNA clones we received from Agrogene at the start of our involvement with the VMC, only 5 yielded viable primers. This was either because the microsatellite repeat contained therein was too short to be informative (less than 10 core repeat units) (Brady *et al.*, 1996), or too close to the end of the sequence for good primers to be developed. It is also likely that the clones that were unsuccessfully sequenced (4 of 48) actually contained very large microsatellites – such clones do not typically grow very vigorously. Furthermore, sequences that amplified duplicate loci would also have been rejected – the majority of our 48 supplied clones (30) fell into this

category (pers. comm. P Isaac, 1999). As a result, only 5 of the 16 primer pairs used were derived from sequence data generated by *us*, while the other 11 primer pair sequences were obtained from the large pool of primers available to VMC members.

All 57 samples were investigated at a minimum of 11 independent microsatellite marker loci, and the KWV and IWBT samples were analysed at a further 5 loci – 16 in total. Results were manually scored, and stored in an Excel spreadsheet. This raw data is presented in Table 3. Statistical analyses were performed using IDENTITY 1.0 (Wagner and Sefc, 1999). This programme was written specifically for the analysis of microsatellite data, and was created during a study of grapevine cultivars by the authors (Sefc *et al.*, 1997).

3.5 Detection

Although today fluorescent detection methods are favoured over manual sizing methods, especially in projects involving a large number of different cultivars, ^{33}P analysis using polyacrylamide gel electrophoresis systems is still an acceptable method in studies such as ours involving only a small number of samples (Sánchez-Escribano *et al.*, 1999). A few salient points are worth mentioning here, though.

3.5.1 Stutter bands

'Stutter' or 'ghost' bands are typical of dinucleotide repeat markers, and result from polymerase slippage (Kijas *et al.*, 1995), which results in a number of fragments differing by the core repeat length. These show up on an autoradiograph as a ladder pattern *ahead* of the true allele length, and can make the scoring of alleles difficult especially when a stutter overlaps a true allele (See Fig. 1). The appearance of stutter bands can be minimised by increasing the annealing temperature as well as reducing the number of amplification cycles.

In scoring the data, if there was more than one equally intense band, the allele pair with the highest molecular weight was scored, as per Lin and Walker (1998). Despite these drawbacks, long dinucleotide-repeat markers are the most informative, so we persevered. (As Ciofi *et al.* (1998) points out, stutter bands

Table 3. Amplified alleles across all loci at each sample, scored in base pair lengths

	3c7	3c7	2e8	2e8	3d8	3d8	16a1	16a1	4c6	4c6	3g11	3g11	3f12	3f12	4d3	4d3
A1	128	164	0	0	59	69	72	74	155	161	0	0	115	117	0	116
A2	122	124	0	0	56	0	72	74	118	124	0	0	0	0	0	0
A3	112	124	0	0	56	0	74	76	112	0	0	0	0	0	110	0
B4	110	138	0	0	69	75	68	0	118	0	0	0	113	117	0	116
B8	117	119	0	0	0	94	68	0	112	127	0	0	0	0	0	0
C4	122	128	0	0	56	66	72	0	0	127	0	0	117	138	0	0
D4	118	122	0	0	59	63	0	0	0	127	0	0	117	0	0	116
A4	102	108	117	148	0	0	160	185	164	0	0	206	0	180	0	0
A5	102	108	117	148	64	70	160	185	164	0	0	206	0	180	0	0
B9	100	116	72	78	225	237	118	158	161	173	0	217	117	180	0	0
B10	100	116	72	78	225	237	118	158	0	0	0	217	117	180	0	0
D2	123	125	72	76	63	0	0	0	165	167	0	195	129	179	0	0
D3	123	125	72	76	63	0	0	0	165	167	0	195	129	179	127	0
D5	103	111	65	63	0	0	185	188	155	0	200	223	127	179	127	0
D6	103	0	65	63	0	0	0	0	155	0	200	223	127	179	127	0
E1	111	125	0	0	0	0	154	158	158	176	188	0	128	178	0	123
E2	111	125	0	0	0	0	154	158	158	176	188	0	128	0	0	123
E8	116	124	74	78	222	0	154	158	155	164	0	0	0	0	0	0
E9	112	122	70	0	222	236	154	158	158	164	188	0	126	0	0	0
A6	106	114	0	0	0	0	118	160	166	0	0	0	0	0	0	0
A7	106	114	0	0	0	0	118	160	166	0	0	0	0	0	0	0
A8	106	114	0	0	0	0	118	160	166	0	0	0	0	0	0	0
B5	100	114	71	73	64	71	153	173	0	0	0	0	115	117	0	0
B6	100	114	71	73	64	71	153	173	0	0	0	0	115	117	104	123
B7	100	114	71	73	64	71	153	173	0	0	0	0	115	117	104	123
C1	0	0	67	73	0	0	0	0	164	0	190	208	170	180	157	0
C2	0	0	67	73	0	0	0	0	164	0	190	208	170	180	157	0
C3	0	0	67	73	0	0	0	0	164	0	190	208	170	180	157	0
D7	113	115	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D8	113	115	0	0	0	0	0	0	0	0	195	198	117	180	0	0
D9	113	115	0	0	0	0	0	0	0	0	195	198	117	180	0	0
D10	113	115	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A9	122	124	0	0	0	0	116	158	158	164	190	196	116	118	111	123
A10	122	124	0	0	0	0	114	158	158	164	190	196	116	118	111	123
B1	0	0	0	0	0	0	0	0	158	164	0	0	116	118	111	123
B2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B3	0	0	0	0	0	0	0	158	158	164	0	0	0	0	0	0
E3	115	123	0	0	0	0	117	0	0	0	194	198	115	117	0	0
E4	115	123	0	0	0	0	0	0	164	176	0	0	115	117	104	112
E5	115	123	0	0	0	0	0	0	164	176	0	0	115	117	0	0
E6	115	123	0	0	0	0	0	0	164	176	194	198	115	117	104	112
E7	115	123	0	0	0	0	0	0	0	0	194	198	115	117	104	112
C5	124	0	0	0	0	0	0	0	0	164	0	194	118	128	113	123
C6	124	0	0	0	0	0	0	0	0	164	0	194	118	128	113	123
C7	124	0	0	0	0	0	0	0	0	164	0	194	118	128	113	123
C8	124	0	0	0	0	0	158	183	0	164	0	194	118	128	113	123
C9	124	0	0	0	0	0	158	0	0	164	0	194	118	128	113	123
C10	124	0	0	0	0	0	157	183	0	164	0	194	118	128	113	123
D1	124	0	0	0	0	0	0	0	0	164	0	194	118	128	113	123
S	165	211	158	184	67	0	75	145	112	0	173	199	117	119	113	116
RG	152	158	158	184	78	0	79	145	118	120	0	199	119	171	113	116
D	0	164	158	184	67	0	73	0	112	116	197	203	117	119	113	116
Cab. Sauv.	165	177	0	0	0	0	0	0	0	0	200	205	117	180	114	0
Riesling	149	0	183	0	0	0	0	0	104	0	171	200	113	117	98	110
V. riparia	153	159	0	0	0	0	0	0	117	119	188	200	224	228	0	0
Sultanine	154	160	0	0	0	0	0	0	107	0	173	200	117	180	112	116
Muscadinia a	157	167	0	0	0	0	0	0	116	0	171	177	115	0	98	0

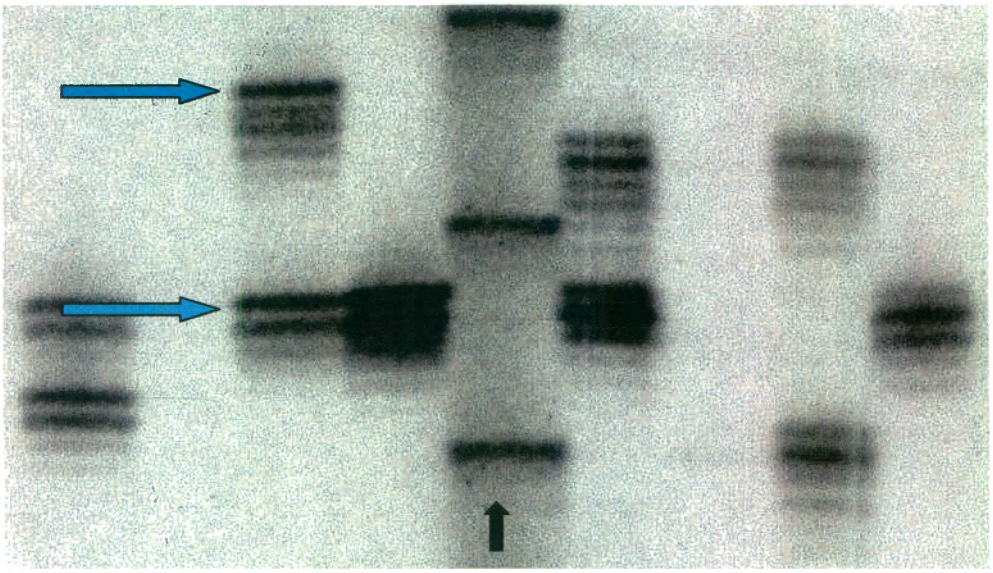
Grey areas indicate the primer/sample combinations not analysed, as a result of insufficient DNA. These only included the original VMC samples, for which we received DNA sufficient only for reactions with 11 primer pairs.

Table 3. Cont.

	4g6	4g6	ZAG67	ZAG67	ZAG112	ZAG112	5g6.1	5g6.1	1a2	1a2	1a5	1a5	1g3.2	1g3.2	1h5	1h5
A1	122	0	0	139	230	240	133	0	0	0	0	0	0	135	121	123
A2	124	0	135	139	228	240	135	147	0	0	0	0	133	0	113	119
A3	0	206	127	0	230	240	123	147	108	134	180	196	0	0	104	0
B4	0	206	127	0	240	244	122	124	106	140	171	184	120	128	100	122
B8	0	0	0	0	228	240	0	141	108	0	0	195	138	0	0	208
C4	124	206	135	139	234	240	118	0	114	0	0	0	130	0	119	121
D4	0	206	135	139	218	226	139	141	105	0	0	0	130	0	116	0
A4	0	0	135	139	228	242	141	165	128	130	176	190	133	0	121	127
A5	0	0	135	139	228	242	141	0	109	115	190	205	133	0	121	123
B9	0	0	127	139	235	243	118	0	108	0	0	195	138	0	208	0
B10	0	0	127	139	235	243	118	0	108	0	0	0	138	0	208	210
D2	138	166	127	139	230	240	130	132	105	0	0	0	130	0	114	0
D3	138	166	127	139	230	240	0	0	105	0	0	0	130	0	0	0
D5	0	0	0	143	227	247	131	133	0	0	0	0	130	0	0	0
D6	138	0	0	143	227	247	131	133	110	0	0	0	152	0	150	198
E1	135	0	127	149	226	240	117	0	103	120	0	184	128	0	124	0
E2	135	0	127	149	226	240	117	0	0	120	0	184	128	0	124	0
E8	126	135	135	139	231	239	136	140	95	103	178	196	0	134	119	121
E9	126	135	135	139	227	239	136	140	0	0	178	196	0	134	101	107
A6	0	0	0	0	228	234	0	0	104	150	0	0	0	133	103	127
A7	125	131	0	0	228	234	0	0	86	108	0	0	0	0	0	0
A8	125	131	0	0	228	234	135	137	0	106	0	180	120	128	112	0
B5	0	0	0	0	232	238	0	141	0	108	0	0	138	0	124	0
B6	0	0	0	0	232	238	0	141	106	108	0	189	138	0	124	0
B7	0	0	0	0	232	238	0	141	118	120	0	189	138	0	124	0
C1	0	0	133	137	0	0	138	140	114	0	0	207	138	0	208	0
C2	0	0	133	137	0	0	138	140	114	0	0	207	130	0	103	107
C3	0	0	0	137	227	239	138	140	114	0	0	207	130	0	103	107
D7	0	0	126	149	233	239	138	140	110	0	184	189	134	0	116	124
D8	0	0	126	149	0	0	138	140	110	0	184	189	134	0	0	124
D9	0	0	126	149	0	0	138	140	110	0	184	189	134	0	0	124
D10	0	0	126	149	0	0	0	0	110	0	184	189	134	0	0	124
A9	119	125	138	140	0	238	117	145	127	0	0	196	120	128	0	135
A10	119	125	138	140	0	238	117	145	127	0	0	196	120	128	106	200
B1	0	125	138	140	0	228	117	0	127	0	0	196	120	128	106	116
B2	0	125	138	140	180	238	117	0	127	0	0	196	120	128	106	124
B3	0	125	138	140	180	238	117	0	127	0	0	196	120	128	106	124
E3	124	126	134	150	0	232	131	133	0	0	178	196	128	134	119	121
E4	124	126	134	150	0	232	131	133	0	0	178	196	128	134	119	121
E5	116	126	134	150	0	232	131	133	0	0	178	196	128	134	119	121
E6	122	126	134	150	0	232	131	133	0	0	168	184	128	134	119	121
E7	122	126	134	150	0	232	131	133	0	0	178	196	128	134	119	121
C5	126	0	125	127	0	240	0	0	0	142	0	0	124	130	119	121
C6	126	0	125	127	0	240	116	118	103	0	0	0	124	130	119	121
C7	126	0	125	127	0	240	116	118	0	0	0	0	124	130	119	121
C8	126	0	125	127	0	240	116	118	103	144	0	0	124	130	119	121
C9	126	0	125	127	0	240	116	118	105	0	0	0	124	130	119	127
C10	126	0	125	127	0	240	116	118	105	0	0	0	124	130	119	127
D1	126	0	125	127	0	240	116	118	105	0	0	0	124	130	119	127
S	123	127	0	0	228	260	136	138	111	230	121	126	126	204	103	0
RG	123	127	0	0	228	232	136	138	113	145	121	126	126	132	103	107
D	123	127	0	0	0	238	136	138	104	106	121	126	126	132	100	102
Cab. Sauv.	126	130	0	0	0	0	0	0	98	106	0	187	129	135	103	110
Riesling	124	128	0	0	0	0	0	0	107	127	0	204	127	129	110	124
V. riparia	0	0	0	0	0	0	0	0	170	170	170	187	127	129	0	200
Sultanine	122	126	0	0	0	0	0	0	0	143	186	188	123	141	0	200
Muscadinia a	0	137	0	0	0	0	0	0	106	160	182	194	134	0	0	200

and 'over-stutters' (see below) are usually fainter than the true allele, and with practice scoring becomes easier).

Figure 1. Section of a sequencing gel illustrating allele lengths and associated stutter bands.



Blue arrows indicate true allele lengths, scored in relation to the DNA size standard in the middle (black arrow).

3.5.2 Non-templated A-addition

Also known as (n+1) morphology, this is as a result of the addition of a single Adenosine residue to the end of an amplified fragment (terminal transferase activity) (Brady *et al.*, 1996). This occurs spontaneously and is a feature of many *Taq* DNA polymerases commonly used in laboratories. The result is that amplified fragments are inconsistently 1bp longer than the actual template sequence – A-addition is affected by the nature of the primer itself (Brownstein *et al.*, 1996). The effect appears similar to stutter on an autoradiograph, except the extra band runs *behind* the true allele (since it is larger). Different types of *Taq* enzymes, like DeepVent polymerase (New England BioLabs) can be used to minimize the appearance of both (n+1) morphology and stutter bands (Lamboy and Alpha, 1998). Also, primers can be 'pigtailed' (specific sequences added to them) which ensures A-addition (Brownstein *et al.*, 1996), or PCR cycling conditions can be modified (Harker, 2001).

3.5.3 Null alleles

Null alleles are not actually alleles at all; rather they are non-amplifications (no band) at a specific marker locus. They arise most often from primer binding-site errors (Brookfield, 1996), but can also be due to the site being absent in that particular sample, technical problems like poor DNA quality, or mutations in the area like insertions or deletions (Thomas *et al.*, 1994). The inheritance of a null allele can be determined by tracing 'non-amplification' at a particular locus in a pedigree (Callen *et al.*, 1993).

3.6 Scoring of data

Alleles were scored as base pair lengths; and recorded in a rectangular spreadsheet form. Some groups like (Brady *et al.*, 1996), assigned letters to the different alleles, and the combination of alleles (profile) was entered into a database. When only one band was visible for a locus, the sample was scored as a heterozygote at that locus (heterozygous for a null allele), according to the precedent set by Thomas *et al.* (1994). Consistently scoring single-banded loci as heterozygotic has the effect of over-estimating the level of heterozygosity. Conversely, consistent scoring of a single band as homozygotic would underestimate the actual heterozygosity in a population. A more accurate estimation of this parameter would therefore lie somewhere between the two extreme values. Heterozygosities for the combined data set 'VMC and KWV samples' were calculated using IDENTITY 1.0.

At least two replicates of each sample/locus combination were performed, and results checked for consistency between amplifications and detections, as was done by (Lamboy and Alpha, 1998). Consistent patterns were obtained for each sample/locus combination, regardless of the extraction method or time of sample collection (February or November).

3.7 Database establishment and Use

Although this work was done using samples from a commercial source with the intention that the results benefit them in some way, findings will help to enlarge the databases established as a result of the VMC.

Not all laboratories involved in this effort used the same detection methods as we did. Before the VMC was established, SSR markers had been developed by two groups - Bowers *et al.* (1996), Thomas and Scott (1993). Each used different methods of separation, but today their results are compared with little trouble. However, it is important to bear in mind that consistent size differences have been observed between radio-isotope and automated detection methods, as well as between gel and capillary methods of separation (Ciofi *et al.*, 1998; Sefc *et al.*, 1998b). Alleles sized using gel electrophoresis and radio-isotopes are 1-2bp longer than the same alleles using automated sizing software (Cipriani *et al.*, 2001). Also, bands detected using silver staining were consistently 1-2bp larger than the same bands detected in a fluorescent system. Such 1or 2bp differences are common when different techniques are used (Lin and Walker, 1998), and these errors fall within the bounds of acceptable sizing error (Ciofi *et al.*, 1998).

In order to avoid confusion and discrepancies in allele lengths, it is recommended that DNA size markers be run in each gel (Ciofi *et al.*, 1998; Cipriani *et al.*, 2001). In our case, we used the same 10bp ladder in each gel. As an example of the effectiveness of this approach, when SSR markers were used for variety identification of tomato and wheat cultivars, several laboratories were involved and a variety of detection methods were employed. Despite these differences, allele scoring was consistent between the groups and allele classifications were the same for the 16 varieties tested (Vosman *et al.*, 2001).

Vosman *et al.* (2001) further points out that the scoring is heavily dependent on the marker used, as well as equipment. It is therefore important to verify that a database (or a contribution to one like the VMC) is accurate and reliable. For example, while some markers in our set displayed quite severe stutter, others showed very little. The same can be said for (n+1) morphology.

Furthermore, samples analysed at loci common to other groups/investigation (ZAG67 and ZAG112) were compared wherever possible. Our data at these sample/locus combinations are consistent with that found both in the literature as well as on the Internet (Regner *et al.*, 2000; Grando *et al.*, 2002).

Once all samples had been analysed at a minimum of 11 loci (common to all three groups), we had the beginnings of a database. It will become more informative and reliable in its estimations of identity as the number of samples and loci considered increases.

The combined genotypes across all the loci represent the specific DNA profile of a given cultivar (Botta *et al.*, 1995). We analysed the profiles of each sample in our database, and determined estimates for a number of parameters. Where appropriate, comparisons were made between different sample groups.

3.8 Statistical analyses

3.8.1 Identical genotypes

By merely looking at the data, it was clear that the three table grape samples from IWBT were distinct. It was subsequently discovered that there were neither similarities between these three cultivars, nor any samples from the VMC and KWV groups.

The VMC samples were analysed separately from the KWV group, and then together. No matching genotypes were uncovered either within the group, or between the two groups.

Finally, the KWV samples were analysed at all 16 loci. As mentioned previously, this group contained a number of clones; a large number of duplicates (synonymies) were therefore expected. Surprisingly, this was not the case – only one pair of cultivars was identical at all loci examined. These were two of the four clones of Sauvignon Blanc, samples D8 and D9. Their profiles are shown in comparison with the other two 'clone' profiles in Table 4 overleaf.

This means that except for these two samples that showed a 100% match at all loci, each accession in our fledgling database has a unique DNA 'fingerprint', differing from the other at a minimum of one locus.

However, the issue of how many differences are allowed before two samples are considered to be separate cultivars remains. Whether there is a 'threshold' number at which differences between cultivars become significant, or whether even a single alteration is reason enough to maintain all variants in a germplasm collection will have a bearing on the maintenance of a diverse, yet non-redundant repository.

Table 4. Comparison of profiles between four Sauvignon Blanc 'clones'*. Alleles are given in base pair lengths.

	D7		D8		D9		D10	
	SB 1		SB 159		SB 2		SB 316	
3c7	113	115	113	115	113	115	113	115
3g11	-	-	190	196	190	196	-	-
3f12	-	-	116	118	116	118	116	118
ZAG67	126	149	126	149	126	149	126	149
ZAG112	233	239	-	-	-	-	-	-
5g6.1	138	140	138	140	138	140	-	-
1a2	110	-	110	-	110	-	110	-
1a5	184	189	184	189	184	189	184	189
1g3.2	134	-	134	-	134	-	134	-
1h5	116	124	-	124	-	124	-	124

* Only informative loci are shown; the area shaded grey highlights the identical profiles of Sauvignon Blanc clones 159 and 2.

Furthermore, when determining the identity of a sample merely by the comparison of allelic profiles, many factors can have an impact on the results obtained. For example, the relative informativeness of the markers used, or technical factors such as DNA quality, amplification conditions or method of separation and detection. As a result, confidence in a result diminishes if it is known that the markers are of average informativeness (low polymorphism levels) or frequently amplify null alleles (often due to poor primer design/ annealing).

A second point to note here is that since all single amplified bands were scored as heterozygotes, there is a high incidence of null alleles making up a large proportion of the total number of loci at which samples differ. As previously mentioned, constant scoring of heterozygotes gives skewed results inasmuch as

the level of heterozygosity is over-estimated. Table 5 shows this trend, by showing that a large number of clonal selections apparently differ at a significant proportion of the loci examined purely because of null alleles. Had we scored all single bands as homozygotes, we would interpret no variation as a result of null alleles. A null allele can result from a mutation in the primer binding site, or if the actual DNA at that particular locus has undergone a mutation. Another explanation for the presence of null alleles is the absence of that particular locus in the sample's genome (Thomas *et al.*, 1994; Pemberton *et al.*, 1995; Paetkau and Strobeck, 1995). Clearly, these two situations represent the extremes, and the true estimation lies somewhere between the two.

Table 5. Number of loci at which clonal groups amongst KWV cultivars differ.

Cultivar	# Clones	# Loci at which they differ	% Loci due to null alleles
Cabernet Franc	2	2	100%
Malbec	2	3	100%
Ruby Cabernet	2	3	100%
Richter 99	2	5	100%
Shiraz	2	2	100%
Viognier	2	8	62.5%
Cabernet Sauvignon	3	6	83%
Semillon	3	3	100%
Merlot	3	2	50%
Sauvignon Blanc	4	5	100%
Chardonnay	5	10	83%
Chenin Blanc	5	6	67%
Pinot Noir	7	5	80%

Another point of interest involves the Pinot Noir clones. The Pinot group of cultivars is notoriously difficult to distinguish (Sefc *et al.*, 2000a; Lopes *et al.*, 2000; Thomas and Scott, 1993), so the variation here is noteworthy. However, null alleles are the source of variation at 80% of the markers showing polymorphism. This seems unlikely, and could be due to poor primer annealing, or the presence of polyphenols or polysaccharides in the DNA (Monteiro *et al.*,

2000). It is also possible that the variable loci are actually homozygotic for the single amplified bands.

3.8.2 Parent/offspring combinations

The only possible parent/offspring combinations were detected amongst the Pinot Noir clonal group. Likelihood ratios for each combination were determined, and the most probable combination was determined as being C10xD1 → C9, representing Pinot Noir clones PN777 x PN 9 → PN 667. Other possible combinations together with their likelihood ratios are listed in Table 6.

Table 6. Possible parentage combinations found amongst Pinot Noir clones in the KWV group of samples

Possible parent/offspring combination	Likelihood ratio considering observed allele frequencies	Likelihood ratio incl. 95% upper confidence limits
C7 = C6 x C9	3.86×10^{11}	3.61×10^{7}
D1 = C6 x C9	2.88×10^{12}	1.42×10^{8}
C9 = C6 x C10	5.77×10^{12}	2.17×10^{8}
C7 = C6 x D1	7.72×10^{11}	7.22×10^{7}
D1 = C7 x C9	5.77×10^{12}	2.84×10^{8}
C9 = C7 x C10	1.15×10^{13}	4.34×10^{8}
C9 = C10 x D1	**2.31×10^{13}**	**8.67×10^{8}**

This cross (C10xD1 → C9) was compared with the likelihood of 4 other cross combinations:

- The probability of C9 resulting from two random, unknown cultivars.
- The probability of C9 resulting from a cross between C10 and an unknown cultivar.
- The probability of C9 resulting from C10 and a relative of D1.
- The probability of C9 resulting from a random cultivar crossed with D1.
- The probability of C9 resulting from cross between a close relative of C10 and D1.

The probabilities for the random cultivars were calculated from the allele frequencies within the population. Because of this, two values are given – in small samples such as ours, certain alleles may be under-represented resulting in allele frequencies lower than the actual population frequency. To allow for such errors, IDENTITY 1.0 also calculates likelihood ratios from the 95% upper confidence limits of the observed frequencies. It is clear from the table that the proposed parent/offspring combination is most likely.

The fact that four of the seven Pinot Noir clones are thought to somehow be involved in a parent/offspring combination is easily believed when one considers a point mentioned above – that the Pinot group is notoriously difficult to differentiate using current molecular markers (incl. RAPDs, AFLPs and SSRs). Furthermore, allowing for the presence of null alleles, it seems that the Pinot group's high degree of relatedness (over centuries of breeding) could account for the inability to separate cultivars using microsatellites.

Such information could potentially be very useful to KWV, were they to be interested in a breeding programme involving any of these cultivars. Additionally, this result confirms that SSRs are suitable for pedigree determination. By increasing the number of informative markers, they may potentially be useful in an entire collection, as has proven to be the case in a Portuguese grapevine set (Lopes *et al.*, 1999).

3.8.3 Informativeness of Markers

In order to determine the usefulness of the various markers used, a number of determinations had to be made. These included the number of alleles per locus, the percentage heterozygosity detected and the Polymorphism Information Content of each marker. This information is summarised in Table 7.

Clear amplification products were obtained for most samples at most of the loci tested. Only primers 2f10 and 16c1 consistently failed to amplify in the majority of the cultivars and were therefore omitted from all calculations.

3.8.3.1 Allele frequencies

Since SSRs are a co-dominant system and therefore allow both allelic states to be detected, allele frequencies were obtained simply by counting the proportion of each in the collection. All loci were multi-allelic, with an average number of 13 alleles per locus. This is slightly higher than what was found by Lopes *et al.* (1999) and Sefc *et al.* (2000a), who found an average of seven and nine.

Table 7. Relative informativeness of 16 markers used to analyse KWV sample group.

Locus	Repeat type	# Alleles	% Heterozygosity expected	% Heterozygosity observed	PI (%)	Freq. (%) Null alleles	Paternity exclusion probability (%)
3c7	Perfect	18	88%	88%	5%	0%	77%
2'e8	Perfect	11	52%	33%	27%	13%	34%
3d8	Perfect	13	48%	35%	30%	9%	32%
16a1	Perfect	16	73%	57%	12%	9%	56%
4c6*	Perfect	13	74%	78%	15%	-2%	55%
3g11	Perfect	11	63%	57%	19%	3%	45%
3f12	Perfect	11	84%	78%	8%	4%	69%
4d3	Perfect	8	64%	55%	21%	5%	44%
4g6	Perfect	11	66%	65%	21%	1%	46%
ZAG67	Compound	11	86%	86%	7%	0%	72%
ZAG112	Perfect	14	85%	90%	7%	-2%	71%
5g6.1	Perfect	14	87%	90%	6%	-2%	74%
1a2	Compound	18	73%	80%	11%	-4%	57%
1a5	Perfect	11	67%	61%	18%	3%	47%
1g3.2	Compound	9	81%	96%	10%	-8%	64%
1h5	Perfect	19	88%	94%	5%	-3%	76%

* = The only trinucleotide repeat; all other markers amplify dinucleotide repeats.

3.8.3.2 Heterozygosity

The observed heterozygosity was calculated from direct counts, and compared with the expected heterozygosity (what can be expected based on the observed allele frequencies), calculated according to Nei (1973).

The expected heterozygosity ranged from between 48% to 88%, with an average of 74%, which is approximately the level observed in other studies on grapevine (Lamboy and Alpha, 1998; Sánchez-Escribano *et al.*, 1999; Hinrichsen *et al.*, 2000). It is also on par with the average observed heterozygosity of 71%.

The observed heterozygosity was higher than the expected heterozygosity at 6 of the 16 loci. This may be explained by the fact that vines used in breeding or for commercial production are usually selected for quality and yield. These hardier plants are typically highly heterozygous in grapevine (which suffers severe inbreeding depression (Olmo, 1976)). In addition, markers 2e8 and 3d8 have significantly lower observed heterozygosity values than to be expected. A likely reason for this is the high incidence of null alleles at these loci. The frequency of null alleles at these two loci was calculated as being 13% and 32% respectively (according to Brookfield, 1996). Values such as these will consequently have a negative effect on the observed heterozygosities at such loci. While inbreeding also typically results in a decrease in heterozygosity (Lopes *et al.*, 1999), this is not likely to be the reason here, since these samples constitute source material for the establishment of vineyards. As such, they will be selected specifically *for* heterozygosity, since such plants are typically more vigorous (Lamboy and Alpha, 1998).

The observed and expected heterozygosities were very closely matched, high values for 3c7, 4c6, 4g6, ZAG67, ZAG112 and 1a5, and indicate that they accurately represent the characteristics of grapevine at a molecular level.

3.8.3.3 Probability of Identity (PI)

This can be defined as the probability of a given marker displaying identical profiles (the same allelic combinations) for samples in a collection (Lopes *et al.*, 1999).

The alleles amplified by most of the markers used are evenly distributed with respect to *frequency*, resulting in the relatively low probabilities of identity observed. However, the high PI values for markers 2e8 and 3d8 indicate an uneven distribution of allele frequencies, even though they give an average number of alleles (11 and 13, respectively). For example, comparing marker 2e8 with marker ZAG67 (11 alleles), one can see a marked difference in informativeness, and a significantly lower chance of finding identical profiles at this locus. 2e8's most common allele has a frequency of approx. 60%, with the other ten alleles making up the remaining 40%, whereas ZAG67's most common allele has a frequency of 20%.

Markers 3c7 and 1h5 are the least likely to yield identical profiles for a given sample, as they have the lowest PI values of 5% each (most common alleles' frequencies approximately 20% each).

The probability of exclusion describes the chance of being able to accurately identify a sample as being a possible parent on the basis of the allelic profile at a given locus (by 'excluding' other individuals). The higher the value of this measure, the more informative the marker is in this regard. There is an inverse relation between the probability of exclusion and the probability of identity – the likelihood of being identical at a locus must be small if the chance of excluding another individual as a relative is to be high.

Overall, the total exclusion probability is ~100% (99.99%), and the total probability of identity is 1.029389×10^{-15}. This means that using all of these markers in this 'population', one is virtually guaranteed to distinguish between each sample. Furthermore, the chance of two non-identical samples having the same allelic profiles is vanishingly small. The effect of an increase in sample size on the PI value of a marker can be seen by comparing Tables 8, 9 and 10 – as sample size increases,

there is a significant reduction in the likelihood of their profiles being identical at any particular locus.

Table 8. The 5 most polymorphic loci observed in the IWBT cultivars

Locus	# Alleles	Probability of Identity (PI)* (%)
1a2	6	10%
3c7	5	15%
3g11	5	15%
4c6	5	57%
1h5	4	25%

* The 'Probability of Identity' is the probability of two randomly chosen cultivars displaying the same SSR profile (30). This index is included here to illustrate how the PI value drops as the number of cultivars considered increases (see tables 9 and 10).

Table 9. Loci analysed in the VMC samples, with number of amplified alleles

Locus	**# Alleles**	**Probability of Identity (%)**
3c7*	**10**	**4%**
1a2	**8**	**7%**
1a5	7	10%
1g3.2	7	10%
4g6	7	10%
3f12	**8**	**8%**
4c6	6	17%
4d3	**6**	**15%**
3g11	6	15%
1h5	5	18%
2e8	**2**	**72%**

bold* = The most informative loci from Table 8 - table grape cultivar analysis (3 samples). Note the general trend in PI values, dropping as the number of cultivars increases (here, 5 samples were analysed).

Table 9. IWBT and VMC samples analysed together

Locus	# Alleles	Probability of Identity (%)
3c7	11	4%
1a2	13	3%
1a5	9	7%
1g3.2	10	5%
3g11	9	9%
3f12	10	7%
4c6	9	9%
4g6	7	11%
1h5	7	11%
4d3	6	14%
2e8[b]	**3**	**37%**

[b] = Marker 2e8 displays the largest decrease in PI from the combination of data, even though the number of alleles increases by only one.

3.8.4 Specific markers

Two markers of interest in this project were developed on the rootstock species *Vitis riparia* and were not part of the VMC (Sefc *et al.*, 1999). It has been reported that markers used on non-native species often display reduced polymorphism as amplified products tend to be shorter in the foreign species than in the species for which the marker was developed (Cipriani *et al.*, 2001). It is therefore ironic that these two markers are among the most informative in this collection.

PCR conditions were adapted in order to obtain amplification products. This primer pair has the lowest annealing temperature (49°C) of all used. This was the only significant difference between these two and the remaining 14 markers used. Also, results obtained were in good agreement with other groups who also applied these markers to *Vitis vinifera* cultivars, in the sense that levels of heterozygosity were similar, and amplified bands fell into the same size range (*et al.*, 1999; Regner *et al.*, 2000b; Sefc *et al.*, 2000a), with comparable profiles for common samples.

Additionally, our work confirms the use of microsatellite markers between related species – here ZAG67 and ZAG112, developed on *Vitis riparia*, were successfully applied to *Vitis vinifera* cultivars. This has the potential to reduce the (often prohibitive) costs of SSR development for other research groups, thereby making this type of research available to a wider range of researchers.

3.8.5 Specific samples

To satisfy our own curiosity, we compared the two individuals mentioned in Chapter 1 – Auxerrois and Chardonnay – which had previously been identified (incorrectly) as being identical. We found them to differ at 8 of the 16 loci investigated, confirming what is already known. Although they are virtually indistinguishable at the phenotypic level, at the molecular level Auxerrois is completely distinct from Chardonnay (see chapter 1).

3.9 Conclusion

A primary application of this work is the identification of duplicate accessions in this sample set for the possible extrapolation of any successes to include larger data sets. This has successfully been done by Regner *et al.* (2001), Lopes *et al.* (2000) and Borrego *et al.* (2001) for specific collections around Europe, the Azores and Spain respectively.

As mentioned beforehand, grapevines are usually classified according to ampelographic characteristics so unearthing these apparent differences at a molecular level is to be expected. It is likely that although these plants may be genetically dissimilar, they are still for all practical purposes identical – microsatellites are typically in non-coding regions of the genome; molecular differences at these marker loci may not have any real effect on the phenotype of the plant, berry size or quality of wine produced. Our results show that the likelihood of any randomly selected two plants having the identical genotype at all 16 loci is slim. What a finding such as this does provide is a means to identify duplicates in a collection, and so eliminate redundancies. Furthermore, Regner *et al.* (2001) has determined the most informative markers to be six others from the VMC collection – VVS2, VVMD5, VVMD7, VVMD27, VRZAG62 and VRZAG79. They are apparently capable of distinguishing between all the cultivars analysed thus

far. These would doubtless be informative if applied to cultivars in the KWV germplasm collection, which houses a large proportion of cultivars already tested and reviewed in the literature.

Chapter 4

Conclusion

Our involvement in the characterisation of grapevine started with our membership in the Vitis Microsatellite Consortium. The apparent success of this international effort prompted our attempts to apply the same markers used to a larger group of grapevines, which were of specific relevance to the South African viticulture industry. We approached KWV in order to obtain plant material for this study, and were fortunate enough to have samples chosen for us whose genetic identity KWV were particularly interested in investigating.

We successfully characterised 25 different grapevine cultivars (including a number of clones) at a minimum of 11 microsatellite loci. A genetic 'profile' for each sample was obtained, and a potentially highly informative database was established. In addition, we determined the relative informativeness of each marker used. In so doing, we realised that a number of our markers were less than ideal for the sort of analysis that is anticipated in this area in the future. For example, if legal issues are to be resolved, or unknown vines accurately and conclusively identified, one cannot use markers of reduced informativeness (like our 2e8 and 3d8; see table 6) when more suitable markers are known (Regner *et al.*, 2001).

On the other hand, we also confirmed that a few markers in our possession were both informative and reliable, and could well be applied to the characterisation of grapevines in germplasm collections, as well as in the detection of synonymies; such markers include ZAG67, ZAG112, 3c7, 5g6.1, 1g3.2 and 1h5. These six markers display a high level of heterozygosity as well as having low null allele frequencies. They are also informative inasmuch as parentage analysis is concerned, with paternity exclusion probabilities of approximately 70%.

On the basis of these markers in conjunction with the remaining, somewhat less informative markers, each cultivar was profiled. It was subsequently shown that two accessions of the KWV germplasm bank appeared to be identical at all loci examined. It would therefore seem that these clones are in fact duplicates. The remaining samples all appear to have unique profiles.

However, before any samples are ejected from the collection or re-classified, it would be prudent not only to test further using microsatellite markers known to be more informative, but also to use different molecular methods such as AFLPs, which offer far wider genome coverage and which have been successfully applied in a number of grapevine analyses. Prof. Carole Meredith's reservation notwithstanding (Meredith, 1998), AFLPs are a promising marker in the field of grapevine identification and analysis (Arnau *et al.*, 2001; Scott *et al.*, 2000a; Martínez-Zapater *et al.*, 2000).

One should also not ignore the valuable contribution older technologies can make to the identification of vines – ampelography and isozyme analysis still have a place, and can serve as tools to 'narrow the search' for an identity of a vine, as well as allowing groups with less technology at their disposal to continue work which will finally contribute to the greater mass of knowledge. Examples where older, more 'traditional' methods of analysis have successfully been combined with newer approaches (like SSRs and AFLPs and ISSRs) include work by Botta *et al.* (2000) and Borrego *et al.* (2001). It is also important to remember that molecular markers alone may not adequately reflect the patterns of variation related to characteristics subject to differential selection like the agronomically important ones in crop species (Hodgkin *et al.*, 2001).

As far as the aims of the project are concerned, the samples at our disposal have been characterised using the microsatellite markers available, whose informativeness and future applicability has been determined. The expansion of this project to include more samples in order to enlarge upon the fledgling database established would require the use of more, powerful markers. The use of such a database for the identification of duplications in the collection, characterisation of unknown cultivars or the settling of legal disputes would then be viable. An increase in the number of samples and markers used would further add weight to any statistical analyses,

resulting in them being more informative and representative of the population at large. Also, as Bowers *et al.* (1996) points out, as the number of markers increases so too does the likelihood of finding a set of primers (like VVS2, VVMD5, VVMD7, VVMD27, VRZAG62 and VRZAG79 – Regner *et al.*, 2001), which could be used to conclusively identify any cultivar from any collection worldwide. Our study, and our empirical evaluation of the markers in our possession, has confirmed that microsatellites are highly informative, reliable markers. We are confident that this body of work can form the base of a useful and reliable database for South African grapevines.

References

1. Amos, W. 1999. A comparative approach to the study of microsatellite evolution, in Goldstein, D.B. and Schlötterer, C. (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press.
2. Arnau, G., Lallemand, J. and Bourgoin, M. 2001. Are AFLP markers the best alternative for cultivar identification?, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 301-306. Acta Horticulturae 546, ISHS.
3. Arroyo-García, R. and Martínez-Zapater, J.M. 2000. Characterisation of new polymorphic simple sequence repeat loci and chloroplast microsatellites in grape (*Vitis vinifera* L.). Paper presented at Plant and Animal Genome VIII Conference, Town and Country Hotel, San Diego, USA.
4. Balding, D. 1999. Forensic applications of microsatellite markers, in Goldstein, D.B. and Schlötterer, C. (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press.
5. Beckmann J.S. and Soller, M. 1990. Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. *Bio/Technology*, 8: 930-932.
6. Bellin, D., Velasco, R. and Grando, M.S. 2001. Intravarietal DNA polymorphisms in grapevine (*Vitis vinifera* L.), in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 343-349. Acta Horticulturae 546, ISHS.
7. Birky, C.W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 92: 11331-11338.

8. Blair, M.W., Panaud, O. and McCouch, S.R. 1999. Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). *Theoretical Applied Genetics*, 98: 780-792.

9. Böhm, A. and Zyprian, E. 2000. First steps to physical mapping of the grapevine genome by pulsed gel electrophoresis, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 271-272. Acta Horticulturae 528, ISHS.

10. Borrego, J., Rodriguez, I., de Andres, M.T., Martin, J., Chavez, J., Cabello, F. and Ibañez, J. 2001. Characterisation of the most important Spanish grape varieties through isozyme and microsatellite analysis, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 371-375. Acta Horticulturae 546, ISHS.

11. Boselli, M., Iannini, C., Corso, C. and Monaco, A. 2000. Analysis of variability in the Aglianico grapevine (*Vitis vinifera)*, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 45-50. Acta Horticulturae 528, ISHS.

12. Botstein D., White, R.L., Skolnick, M. and Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32: 314-331.

13. Botta, R., Schneider, A., Akkak, A., Scott, N.S. and Thomas, M.R. 2000. Within cultivar grapevine variability studied by morphometrical and molecular marker based techniques, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 91-96. Acta Horticulturae 528, ISHS.

14. Botta, R., Scott, N.S., Eynard, I. And Thomas, M.R. 1995. Evaluation of sequence-tagged site markers for characterising *Vitis vinifera* cultivars. *Vitis*, 34 (2): 99-102.

15. Bourquin, J.C., Otten, L., and Walter, B. 1995. PCR-RFLP analysis of *Vitis*, *Ampelopsis* and *Parthenocissus* and its application to the identification of rootstocks. *Vitis*, 34 (2): 103-108.

16. Boursiquot, J.-M. 2000. Development of methods for the conservation and the management of grape genetic resources, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 31-36. Acta Horticulturae 528, ISHS.

17. Bowers, J.E. and Meredith, C.P. 1997. The parentage of a classic wine grape, Cabernet Sauvignon. *Nature Genetics*, 16: 84-87.

18. Bowers, J.E., Bandman, E.B. and Meredith, C.P. 1993. DNA Fingerprint Characterisation of some wine grape cultivars. *American Journal of Enology and Viticulture*, 44 (3): 266-274.

19. Bowers, J.E., Dangl, G.S., Vignani, R. and Meredith, C.P. 1996. Isolation and characterisation of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). *Genome*, 39: 628-633.

20. Bowers, J.E., Siret, R., This, P., Boursiquot, J.M. and Meredith, C.P. 2000. A single pair of parents proposed for a group of grapevine varieties in North-eastern France, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 129-132. Acta Horticulturae 528, ISHS.

21. Brady, J.L., Scott, N.S. and Thomas, M.R. 1996. DNA typing of hops (*Humulus lupulus*) through application of RAPD and microsatellite marker sequences converted to sequence tagged sites (STS). *Euphytica*, 91: 277-284.

22. Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, 5: 453-455.

23. Brownstein, M.J., Carpten, M.J. and Smith, J.R. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques*, 20: 1004-1010.

24. Buck, S. and Zyprian, E. 2000. First approaches of molecular mapping in a model population derived from the crossing of the grapevine

varieties 'Regent' X 'Lemberger', in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 203-208. Acta Horticulturae 528, ISHS.

25. Büscher, N., Zyprian, E. and Blaich, R. 1993. Identification of grapevine cultivars by DNA analyses: Pitfalls of random amplified polymorphic DNA techniques using 10mer primers. *Vitis* 32: 187-188.
26. Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C. and Sutherland, G.R. 1993. Incidence and Origin of "Null" alleles in the (AC)n Microsatellite Markers. *American Journal of Human Genetics*, 52: 922-927.
27. Cangelosi, J. 2001. *The Vermont Flora Project: Vitaceae* [Online] Available: http://strimbeck.faculty.jsc.vsc.edu/FamilyPages/Vitaceae/Vitaceae.htm.
28. Cervera, M.-T., Cabezas, J.A., Sancha, J.C., Martinez de Toda, F. and Martinez-Zapater, J.M. 1998. Application of AFLPs to the characterisation of grapevine *Vitis vinifera* L. genetic resources. A case study with accessions from Rioja (Spain). *Theoretical Applied Genetics*, 97: 51-59.
29. Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J. and Deka, R. 1987. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proceedings of the National Academy of Sciences of the United States of America*, 94: 1041-1046.
30. Ciofi, C., Funk, S.M., Coote, T., Cheesman, D.J., Hammond, R.L., Saccheri, I.J., and Bruford, M.J. 1998. Genotyping with microsatellite markers, in Karp, A., Isaac, P.G. and DS Ingram (Eds.). *Molecular Tools for Screening Biodiversity*. London: Chapman and Hall: 195-201.
31. Cipriani, G., Frazza, G., Peterlunger, E. and Testolin, R. 1994. Grapevine fingerprinting using microsatellite repeats. *Vitis*, 33: 211-215.
32. Cipriani, G., Lot, G., Huang, W.-G., Marrazzo, M.T., Peterlunger, E. and Testolin, R. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: Isolation, characterisation and cross-species amplification in *Prunus*. *Theoretical Applied Genetics*, 99: 65-72.

33. Cipriani, G., Marrazzo, M.T., Di Gaspero, G. and Testolin, R. 2001. DNA Microsatellite in Fruit Crops: Isolation, Length polymorphism, Inheritance, Somatic stability and cross-species conservation, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 145-150. Acta Horticulturae 546, ISHS.

34. Classification of vine varieties in the European Community and common organisation of the market in wine, COM (93) 9282670325, final.

35. Crespan, M and Milani, N. 2000. Application of various molecular methodologies to the characterisation of rootstocks and table grapevines, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 97-105. Acta Horticulturae 528, ISHS.

36. Crow, J.F. and Kimura, M. 1970. *An Introduction to Population Genetics Theory*. New York: Harper and Row.

37. Dalbó, M.A., G.N. Ye, Weeden, N.F., Wilcox, W.F. and Reisch, B.I. 2001. Marker-assisted selection for powdery mildew resistance in grapes. *Journal of the American Society of Horticultural Science*, 126 (1): 83-89.

38. De Woody, J.A. and Avise, J.C. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, 56: 461-473.

39. Di Gaspero, G., Peterlunger, E., Testolin, R. and Cipriani, G. 2001. Characterisation of microsatellite flanking sequences in *Vitis* species, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 199-204. Acta Horticulturae 546, ISHS.

40. Di Rienzo, A., Donnelly, P., Toomajian, C., Sisk, B., Hill, A., Petzl-Erler, M.L., Haines, G.K. and Barch, D.H. 1998. Heterogeneity of Microsatellite mutations within and between loci, and implications for human demographic histories. *Genetics*, 148: 1269-1284.

41. Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M and Freimer, N.B. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America,* 91 (8): 3166-3170.

42. Edwards, K., Johnstone, C. and Thompson, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 19: 1349.

43. Eimert, K. and Schroder, M. 2001. Molecular analysis of the resistance of the "Borner" grapevine rootstock variety to phylloxera. *Mediterranean Agronomic Institute of Chania* [Online]. Available: http://www.maich.gr/sites/cost844/topics.htm.

44. Eisen, J.A. 1999. Mechanistic basis for microsatellite instability, in Goldstein, D.B. and Schlötterer, C. (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press.

45. Estoup, A. and Cornuet, J.M. 1999. Microsatellite evolution: inferences from population data, in Goldstein, D.B. and Schlötterer, C. (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press.

46. F. Lahogue, P. This, A. Bouquet. 1998. Identification of a co-dominant scar marker linked to the seedlessness character in grapevine. *Theoretical Applied Genetics*, 97: 950-959.

47. Ferguson, A., Hynes, R.A., Prodöhl, P.A. and Taggart, J.B. 1995. Molecular Approaches to the study of genetic variation in Salmonid fishes. *Nordic Journal of Freshwater Research*, 71: 23-32.

48. Galau, G.A., Chamberlain, M.E., Hough, B.H., Britten, R.J. and Davidson, E.H. 1976. Evolution of repetitive and non-repetitive DNA, in Ayala, F.J. (Ed.). *Molecular Evolution*. Sunderland, Massachusetts: Sinauer Associates, Inc.: 200-224.

49. Galet, P. 1979. *A practical ampelography*. Ithaca, New York: Comstock Publishing Association.

50. Gorgorcena, Y. Arulsekar, S., Dandekar, A.M. and Parfitt, D.E. 1993. Molecular Markers for grape characterisation. *Vitis*, 32: 183-185.

51. Grando, M.S. and Frisinghelli, C. 1998. Grape microsatellite markers: Sizing of DNA alleles and genotype analysis of some grapevine cultivars. *Vitis*, 37 (2): 79-82.

52. Grando, M.S., Costantini, L., Madini, L. and Segala, C. 2002. *Grape Microsatellite Collection* [Online]. Available: http://relay.ismaa.it:12164/genetica/gmc.html

53. Grando, M.S., Frisinghelli, C. and Stefanini, M. 2000. Genotyping of local grapevine germplasm, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 183-188. Acta Horticulturae 528, ISHS.

54. Grando, M.S., Frisinghelli, C. and Stefanini, M. 2000. Polymorphism and distribution of molecular markers in a segregating population derived from the cross 'Moscato Bianco' X *Vitis riparia*, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 209-214. Acta Horticulturae 528, ISHS.

55. Guerra, B. and Meredith, C.P. 1995. Comparison of *Vitis berlandieri* x *Vitis riparia* rootstock cultivars by restriction fragment length polymorphism analysis. *Vitis*, 34 (2): 109-112.

56. Gupta, M., Chyi, Y.-S., Romero-Severson, J. and Owen, J.L. 1994. Amplification of DNA markers from evolutionary diverse genomes using single primers of simple-sequence repeats. *Theoretical Applied Genetics*, 9: 998-1006.

57. Haas, H.U. 2000. The karyotype of grapevine (*Vitis vinifera* L.), in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 247-256. Acta Horticulturae 528, ISHS.

58. Hamada H., Petrino. M.G. and Kakunaga, T. 1982. A novel repeat element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 79: 6465-6469.

59. Hancock, J.M. 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms, in Goldstein, D.B. and Schlötterer, C. (Eds.). *Microsatellites: Evolution and Applications*. Oxford University Press.

60. Harker, N. 2001. Collection, Reporting and Storage of microsatellites genotype data, in Henry, R.J. (Ed.). *Plant Genotyping: The DNA Fingerprinting of Plants*. New York: CABI Publishing.

61. Hinrichsen, P., Narváez, C., Valenzuela, J., Muñoz, C., Bowers, J.E. and Meredith, C.P. 2000. Fingerprinting of grape cultivars grown in Chile: comparison of methods based on anonymous sequences and a set of microsatellite loci, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 161-168. Acta Horticulturae 528, ISHS.

62. Hodgkin, T., Roviglioni, R., De Vicente, M.C. and Dudnik, N. 2001. Molecular methods in the conservation and use of plant genetic resources, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 107-118. Acta Horticulturae 546, ISHS.

63. *ProAgri*. 2000. [World Wide Web] Available: http://www.proagri.co.za/E/E_41-02-04-18_1.htm.

64. Huang, W.-G., Cipriani, G., Morgante, M. and Testolin, R. 1998. Microsatellite DNA in *Actinidia chinensis*: isolation, characterisation and homology in related species. *Theoretical Applied Genetics*, 97: 1269-1278.

65. Ibañez, J. 2001. Mathematical analysis of RAPD data to establish reliability of varietal assignment in vegetatively propagated species, in Doré, C., Dosba, F. and Baril, C. (Eds). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 73-70. Acta Horticulturae 546, ISHS.

66. Jain, A., Apparanda, C. and Bhalla, P.L. 1999. Evaluation of genetic diversity and genome fingerprinting of *Pandorea* (Bignoniaceae) by RAPD and inter-SSR PCR. *Genome*, 42: 714-719.

67. Jeffreys, A.J. 2002. *Science Watch* [Online] Available: http://www.sciencewatch.com/interviews/sir_alec_jeffreys.htm.

68. Jeffreys, A.J., Wilson, V. and Thein, L. 1985. Hypervaribale 'minisatellite' regions in human DNA. *Nature*, 314 (6006): 67-73.

69. Kijas, J.M.H., Fowler, J.C.S. and Thomas, M.R. 1995. An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. *Genome* 38: 349-355.

70. Kim, C.S., Lee, C.H., Shin, J.S., Chung, Y.S. and Hyung, N.I. 1997. A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. *Nucleic Acids Research*, 25 (5): 1085-1087.

71. Kimura, M. and Crow, J.F. 1964. The number of alleles that can be maintained in a finite population. *Genetics*, 49: 725-738.

72. Kimura, M. and Ohta, T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the United States of America*, 75: 2868-2872.

73. Klug, W.S. and Cummings, M.R. *Concepts of Genetics*. 2nd Edition. Sixth Edition. New Jersey: Prentice Hall, Inc.

74. *KWV International*. N.d. [Online] Available: http://www.kwv-international.com.

75. Lamboy, W.F. and Alpha, C.G. 1998. Using Simple Sequence Repeats (SSRs) for DNA fingerprinting germplasm accessions of grape (*Vitis* L.) species. *Journal of the American Society of Horticultural Science*, 123 (2): 182-188.

76. Larkin, PJ. and Scowcroft, W.R. 1981. Somaclonal variation - a novel source of variability from cell cultures for plant improvement. *Theoretical Applied Genetics*, 60: 197-214.

77. Lefort, F., Anzidei, M., Roubelakis-Angelakis, K.A. and Vendramin, G.G. 2000. Paper presented at the 6th International Symposium on Grapevine Physiology and Biotechnology, Heraklion, Crete, Greece.

78. Leroy, X.J. and Leon, K. 2000. A Rapid Method for Detection of Plant Genomic Instability Using Unanchored-Microsatellite Primers. *Plant Molecular Biology Reporter*, 18 (3): 283.

79. Levinson, G. and Gutman, G.A. 1987a. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, 4 (3): 203-221.

80. Levinson, G. and Gutman, G.A. 1987b. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Research*, 15 (13): 5323-5338.

81. Lin, H. and Walker, A. 1998. Identifying Grape rootstocks with Simple Sequence Repeat (SSR) markers. *American Journal of Enology and Viticulture*, 49 (4): 403-407.

82. Litt, M. and Luty, J.A. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, 44 (3): 397-401.

83. Lodhi, M.A. and Reisch, B.I. 1995. Nuclear DNA content of *Vitis* species, cultivars and other genera of the Vitaceae. *Theoretical Applied Genetics*, 90: 11-16.

84. Lodhi, M.A., Daly, M.J., Ye, G.-N., Weeden, N.F. and Reisch, B.I. 1995. A molecular marker based linkage map of *Vitis*. *Genome*, 38: 786-794.

85. Lodhi, M.A., Ye, G., Weeden, N.F. and Reisch, B.I. 1994. A simple and efficient method for DNA extractions from grapevine cultivars and *Vitis* species. *Plant Molecular Biology Reporter*, 12: 6-13.

86. Lopes, M.S., Mendonça, D., da Câmara Machado, A., da Câmara Machado, M.L., Steinkellner, H and Sefc, K.M. 2000. Genotyping of *Vitis* used for the production of VQPRD and VLQPRD in the Azores by microsatellite markers, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine*

Genetics and Breeding. Montpellier, France: 189-192. Acta Horticulturae 528, ISHS.

87. Lopes, M.S., Sefc, K.M., Dias, E.E., Steinkellner, H., da Câmara Machado, M.L. and da Câmara Machado, A. 1999. The use of microsatellites for germplasm management in a Portuguese grapevine collection. *Theoretical Applied Genetics*, 99: 733-739.

88. Malossini, U., Grando, M.S., Roncador, I. and Mattivi, F. 2000. Parentage analysis and characterisation of some Italian *Vitis vinifera* crosses, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 139-146. Acta Horticulturae 528, ISHS.

89. Markert, J.A., Danley, P.D. and Arnegard, M.E. 2001. New markers for new species: microsatellite loci and the East African cichlids. *Trends in Ecology and Evolution*, 16 (2): 100-107.

90. Martínez-Zapater, J.M., Cabezas, J.A. and Cervera, M.T. 2000. AFLPs in genetic identification and genome analysis of grapevines, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 105-110. Acta Horticulturae 528, ISHS.

91. Merdinoglu, D., Butterlin, G., Baur, C. and Balthazard, J. 2000. Comparison of RAPD, AFLP and SSR (microsatellite) markers for genetic diversity analysis in *Vitis vinifera* L., in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 193-198. Acta Horticulturae 528, ISHS.

92. Meredith, C.P. 1998. *Viticulture Project Summary*. [Online]. Available: http://www.avf.org/reports/9697v114.html.

93. Messier, W., Li, S. and Stewart, C. 1996. The birth of microsatellites. *Nature*, 381: 483.

94. Michelmore, R.W. and Paran, I. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical Applied Genetics*, 85: 985-993.

95. Monteiro, F.F., Nunes, E., Magalhaes, R., Martins, A., Bowers, J.E. and Meredith, C.P. 2000. Fingerprinting of the main *Vitis vinifera* varieties

grown in the northern region of Portugal, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 121-128. Acta Horticulturae 528, ISHS.

96. Moreno, S., Gogorcena, Y. and Ortiz, J.M. 1995. The use of RAPD markers for identification of cultivate grapevine (*Vitis vinifera* L.). *Scientia Horticulturae*, 62: 237-243.

97. Moreno, S., J. P. Martin, and J. M. Ortiz. 1998. Inter-simple sequence repeats PCR for characterisation of closely related grapevine germplasm. *Euphytica*, 101: 117-125.

98. Morgante, M. and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*, 3: 175-182.

99. Mullis, K.B. and Faloona, F.A. 1987. Specific synthesis of DNA *in vitro* via a polymerase- catalysed chain reaction. *Methods in Enzymology*, 155: 335-350.

100. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M. and Kumlin, E. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, 235 (4796): 1616-1622.

101. Nei, M. 1973. Analysis of gene diversity in sub-divided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70: 3321-3323.

102. Olmo, H.P. 1976. Grapes, in Simmonds, N.W. (Ed.). *Evolution of Crop Plants*. London: Longman: 294-298.

103. Paetkau, D. and Strobeck, C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology*, 4: 519-520.

104. Passos, I.R.S., Appezzato-da-Gloria, B. and Vieira, M.L.C. 1999. Embryogenic responses of Vitis spp.: Effects of genotype and polyvinylpyrrolidone. *Vitis*, 38 (2): 47-50.

105. Patil, S.G. and Patil, V.P. 1992. Karyomorphology of *Vitis vinifera*, *V. rotundifolia* and their hybrid. *Cytologia* 57: 91-95.

106. Pemberton, J.M., Slate, J., Bancroft, D.R. and Barrett, A. 1995. Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, 4: 249-252.

107. Perold, A.I. 1927. *A Treatise on Viticulture*. London: Macmillan.

108. Pongrácz, D.P. 1983. *Rootstocks for Grapevines*. Cape Town: David Philip.

109. Porebski, S., Bailey, G. and Baum, B.R. 1998. Modifications of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 123 (2): 182-188.

110. Primikirios, N., Lefort, F., Burger, J.T. and K.A. Roubelakis-Angelakis. 2000. Classification of Vitis vinifera (L.) in spermatophytes molecular phylogeny based on chloroplast 16S and 23S rDNA sequences. Paper presented at 6th International Symposium on Grapevine Physiology and Biotechnology, Heraklion, Crete, Greece.

111. Regner, F., Sefc, K.M., Glössl, J. and Steinkellner, H. 2000a. Parentage analysis and pedigree reconstruction of vine cultivars using microsatellite markers, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 133-138. Acta Horticulturae 528, ISHS.

112. Regner, F., Stadlbauer, A. and Eisenheld, C. 2001. Molecular markers for genotyping grapevine and for identifying clones of traditional varieties, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 331-341. Acta Horticulturae 546, ISHS.

113. Regner, F., Stadlhuber, A., Eisenheld, C. and Kaserer, H. 2000b. Considerations about the evolution of grapevine and the role of Traminer, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 177-182. Acta Horticulturae 528, ISHS.

114. Reisch, B.I. 2000. Molecular markers: The foundation for grapevine genetic mapping, DNA fingerprinting and genomics, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 85-90. Acta Horticulturae 528, ISHS.

115. Reisch, B.I., Weeden, N.F., Lodhi, M.A., Ye, G. and Soylemezoglu, G. 1995. Linkage Map Construction in two hybrid grapevine (Vitis sp.) populations. Paper presented at Plant Genome IV Conference, Town and Country Conference Centre, San Diego, USA.

116. Rossetto, M. 2001. Sourcing of SSR markers from related plant species. In Henry, R.J. (Ed.). *Plant genotyping: The DNA Fingerprinting of Plants*. New York: CABI Publishing.

117. Ryan, F.J., Ledbetter, C.A., Ramming, D.W., Palmquist, D.E., Bell, D.E. and Peterson, S.J. 2001. Challenges in developing molecular markers for almond (*Prunus amygdalus)* and grape (*Vitis* species), in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 629-639. Acta Horticulturae 546, ISHS.

118. Sánchez-Escribano, E.M., Martin, J.P., Carreño, J. and Cenis, J.L. 1999. Use of sequence-tagged microsatellite sit markers for characterising table grape cultivars. *Genome*, 42: 87-93.

119. Schlötterer, C. and Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, 20 (2): 211-215.

120. Scott, K.D., Ablett, E.M., Lee, L.S. and Henry, R.J. 2000a. AFLP markers distinguish an early mutant of Flame Seedless grape. *Euphytica*, 113: 245-249.

121. Scott, K.D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E.M., Lee, L.S. and Henry, R.J. 2000b. *Theoretical Applied Genetics*, 100: 723-726.

122. Scott, K.D., Lee. L.S., Dow, T. and Henry, R.J. 2000c. Isolation and characteristics of new grape microsatellites, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 199-200. Acta Horticulturae 528, ISHS.

123. Sefc, K.M., Glössl, J. and Steinkellner, H. 2000a. Broad range genotyping using microsatellite markers identified in *Vitis riparia*, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 111-120. Acta Horticulturae 528, ISHS.

124. Sefc, K.M., Guggenburger, S., Regner, F., Lexer, C., Glössl, J. and Steinkellner, H. 1998a. Genetic analysis of grape berries and raisins using microsatellite markers. *Vitis*, 37 (3): 123-125.

125. Sefc, K.M., Lopes, M.S., Lefort, F., Botta, R., Roubelakis-Angelakis, K.A., Ibáñez, J., Pejić, I., Wagner, H.W., Glössl, J. and Steinkellner, H. 2000b. Microsatellite variability in grapevine cultivars from different European regions and evaluation of assignment testing to assess the geographic origin of cultivars. *Theoretical Applied Genetics*, 100: 498-505.

126. Sefc, K.M., Regner, F., Glössl, J. and Steinkellner, H. 1998b. Genotyping of grapevine and rootstock cultivars using microsatellite markers. *Vitis*, 37 (1): 15-20.

127. Sefc, K.M., Regner, F., Turetschek, E., Glössl, J. and Steinkellner, H. 1999. Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome*, 42: 367-373.

128. Sefc, K.M., Steinkellner, H., Glössl, J., Kampfer, S. and Regner, F. 1998c. Reconstruction of a grapevine pedigree by microsatellite analysis. *Theoretical Applied Genetics*, 97: 227-231.

129. Sefc, K.M., Steinkellner, H., Wagner, H. W., Glössl, J. and Regner, F. 1997. Application of microsatellite markers to parentage studies in grapevine. *Vitis*, 36 (4): 179-183.

130. Sensi, E., Vignani, R., Rohde, W. and Biricolti, S. 1996. Characterisation of genetic biodiversity with *Vitis vinifera* L. Sangiovese and Colorino genotypes by AFLP and ISTR DNA marker technology. *Vitis*, 35 (4): 183-188.

131. Silvestroni, O., Di Pietro, D., Intrieri, C., Vignani, R., Filipetti, I., Del Casino, C., Scali, M. and Cresti, M. 1997. Detection of genetic diversity

among clones of cv. Fortana (Vitis vinifera L.) by microsatellite DNA polymorphism analysis. *Vitis*, 36 (3): 147-150.

132. Smith, J.M. 1998. *Evolutionary Genetics*. 2nd edition. Oxford University Press.

133. *South African Wine Industry Information and Systems*. 2000. [Online] Available: http://www.sawis.co.za/eng/public/HomePage.asp.

134. Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98: 503-517.

135. Sturzeneker, R., Bevilacqua, R.A., Haddad, L.A., Simpson, A.J. and Pena S.D. 2000. Microsatellite instability in tumours as a model to study the process of microsatellite mutations. *Human Molecular Genetics*, 9 (3): 347-352.

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

137. Tautz, D. and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research*, 12: 4127-4138.

138. Testolin, R. and Cipriani, G. 1997. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in the genus Actinidia. *Theoretical Applied Genetics*, 94: 897–903.

139. *The Wine Room*. 2000. [Online] Available: http://www.thewineroom.ie/pinotnoir.html.

140. This, P., Lahogue, F., Adam-Blondon, A.F., Doligez, A. and Bouquet, A. 2000. Towards marker-assisted selection for seedlessness in grapevine, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 221-230. Acta Horticulturae 528, ISHS.

141. Thomas, M.R. and Scott, N.S. 1993. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theoretical Applied Genetics*, 86: 985-990.

142. Thomas, M.R., Cain, P and Scott, N.S. 1994. DNA typing of grapevines: A universal methodology and database for describing

cultivars and evaluating genetic relatedness. *Plant Molecular Biology*, 25: 939-949.

143. Thomas, M.R., Franks, T. and Iocco, P. 2000. Transgenic grapevines: Status and future, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 279-288. Acta Horticulturae 528, ISHS.

144. Thomas, M.R., Matsumoto, S., Cain, P. and Scott, N.S. 1993. Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. *Theoretical and Applied Genetics*, 86: 173-180.

145. Ulanowsky, S., Gogorcena, Y., Martínez-de-Toda, Borrego, J., Ibáñez, J. and Ortiz, J.M. 2001. Characterisation of grapevine accessions at germplasm banks with RAPD and microsatellite markers, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 271-300. Acta Horticulturae 546, ISHS.

146. Vidal, J.R., Coarer, M. and Defontaine, A. 1999. Genetic relationships among grapevine varieties grown if different French and Spanish regions based on RAPD markers. *Euphytica*, 109: 161-172.

147. Vidal, J.R., Delavault, P., Coarer, M. and Defontaine, A. 2000. Design of grapevine (*Vitis vinifera* L.) cultivar-specific SCAR primers for PCR fingerprinting. *Theoretical Applied Genetics*, 101: 1194-1201.

148. Vignani, R., Bowers, J.E. and Meredith, C.P. 1996. Microsatellite DNA polymorphisms analysis of clones of *Vitis vinifera* "Sangiovese". *Scienta Horticulturae*, 65: 163-169.

149. Vosman, B., Cooke, R., Ganal, M., Peeters, R., Isaac, P. and Bredemeijer, G. 2001. Standardisation and application of microsatellite markers for variety identification in Tomato and Wheat, in Doré, C., Dosba, F. and Baril, C. (Eds.). *Proceedings of the International Symposium on Molecular Markers for Characterising genotypes and Identifying cultivars in Horticulture*. Montpellier, France: 307-316. Acta Horticulturae 546, ISHS.

150. Wagner, H.W. and Sefc, K.M. 1999. IDENTITY 1.0. Centre for Applied Genetics, University of Agricultural Sciences, Vienna.

151. Walker, M.A. and Boursiquot, J.M. 1992. Ampelographic and isozyme data correcting the misnaming of the grape rootstock SO4 at the University of California, Davis. *American Journal of Enology and Viticulture*, 27: 68-73.

152. Wallace, D.C., Brown, M. and Lott, M.T. 1999. Mitochondrial DNA variations in human evolution and disease. *Gene*, 238: 211-230.

153. Wang, X., Wang, Z. and Zou, Y. 1996. An Improved Procedure for the Isolation of Nuclear DNA from Leaves of Wild Grapevine Dried with Silica Gel. *Plant Molecular Biology Reporter*, 14 (4): 369-373.

154. Ware, D., Jaiswall, P., Nil, J., Pan, X., Chang, K., Clark, K., Teytelman, L., Schmidt, S., Zhao, W., Cartinhour, S., McCouch, S and Stein, L. 2002. Gramene: a resource for comparative grass genomics. *Nucleic Acids Research*, 30 (1): 103-105.

155. Weber, J.L. and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics,* 44: 388-396.

156. Weihl, T. and Dettweiler, E. 2000. Differentiation and identification of 500 grapevine (*Vitis* L.) cultivars using notations and measured leaf parameters, in Bouquet, A and Boursiquot, J.-M. (Eds.). *Proceedings of the Seventh International Symposium on Grapevine Genetics and Breeding*. Montpellier, France: 37-44. Acta Horticulturae 528, ISHS.

157. Weising, K. and Gardner, R.C. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, 42(1): 9-19.

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

159. *Wines of South Africa*. 2000. [Online] Available: http://www.wosa.co.za.

160. Ye G., Söylemezeoglu, G., Weeden, N.F., Lamboy, W., Pool, R.M. and Reisch, B.I. 1998. Analysis of the relationship between grapevine cultivars, sports and clones via DNA fingerprinting. *Vitis*, 37: 33-38.

161. Zabeau, M. and Vos, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. *European Patent Application* EP0534858.

162. Zietkiewicz, E., Rafalski, A. and Labuda, D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, 20: 176-183.