TR 95-06

DETERMINATION OF DISTINCTNESS AMONG CITRUS CULTIVARS USING BIOCHEMICAL AND MOLECULAR MARKERS

THESIS

Submitted in fulfilment of the

requirements for the Degree of

MASTER OF SCIENCE of Rhodes University

by

KARIN CARSTENS

February 1994

AAN: MY OUERS

"Education makes a people easy to lead,
but difficult to drive;
easy to govern,
but impossible to enslave."

ABSTRACT

Citrus is among the most important fruit crops worldwide, and therefore the preservation and improvement of citrus germplasm is of the essence. Citrus breeders are often faced with the difficulty of distinguishing between new and existing cultivars because of the ambiguous nature of morphological traits due to environmental influences and error in human judgement. The protection of new varieties is very important to the breeder. New varieties cannot be patented in South Africa, but it can be protected by Plant Breeders' Rights, only if it is genetically distinguishable and significantly different economically from existing varieties.

Cultivars in four genera (*C. sinensis, C. paradisi, C. grandis* and *C. reticulata*) included in the Citrus Improvement Programme (CIP) or cultivars awaiting recognition of Plant Breeders' Rights by the International Union for the Protection of New Plant Varieties (UPOV) were analyzed with Isoenzymes, Restriction Fragment Length Polymorphism (RFLP) and Random Amplified Polymorphic DNA (RAPD).

Five enzyme systems (PGM, PGI, MDH, GOT and IDH) were analyzed and founded to be suitable for grouping together cultivars belonging to the same genera. It was not suited for routine discrimination of cultivars in a particular genus.

RFLP studies were conducted on five grapefruit cultivars, using cDNA clones from a genomic library of Rough Lemon. RFLP studies were valuable for the discrimination of closely related cultivars which probably originated from a common ancestor by bud mutations. This technique was, however, abandoned due to its biohazardous nature and

replaced by the PCR-based Random Amplified Polymorphic DNA.

RAPDs are easy to perform and gave promising results which were exploited to reveal polymorphisms between cultivars within the various groups. Although the interpretation of data produced by this method is often suspicious, it is the best method currently available for cultivar identification. It can play a complementary role in the protection of new varieties when classical morphological interpretation of differences is not capable of determining sufficient distinctness for the awarding of Plant Breeders' Rights.

CONTENTS

ABSTRACT ii

LIST OF TABLES AND FIGURES vii

ACKNOWLEDGEMENTS viii

CHAPTER 1. GENERAL INTRODUCTION AND LITERATURE REVIEW 1

1.1 INTRODUCTION 1

1.2 THE ORIGIN OF CITRUS AND ITS RELATIVES 1

1.3 CLASSIFICATION OF THE MAJOR CITRUS GROUPS 2

1.4 CITRUS BREEDING IN SOUTH AFRICA 4

1.5 THE STATUS OF GENETIC MARKERS IN CITRUS BREEDING 5

1.6 TECHNIQUES, OTHER THAN MORPHOLOGICAL IN THE ORGANIZATION OF GENETIC DIVERSITY CONCERNING CITRUS 10
1.6.1 Electrophoresis of Proteins and Isozymes 10
1.6.2 DNA Profiling 14
1.6.2.1 Restriction Fragment Length Polymorphism 15
1.6.2.2 Random Amplified Polymorphic DNA 18
1.6.2.3 Amplified Fragment Length Polymorphism 24

1.7 OBJECTIVE 24

CHAPTER 2. EVALUATION OF ISOZYME, RFLP AND RAPD TECHNIQUES FOR THE DISCRIMINATION BETWEEN CITRUS CULTIVARS 26

2.1 INTRODUCTION 26

2.2 MATERIALS AND METHODS 27
2.2.1 Isozyme Analysis 27
2.2.1.1 Extraction and Separation of Isozymes 27
2.2.2 Restriction Fragment Length Polymorphism 28
2.2.2.1 DNA Extraction 28

2.2.2.2 Restriction Enzyme Digests 29
2.2.2.3 Preparation of DIG- and ^{32}P-labelled probes 29
2.2.2.4 Southern blot, Hybridization and Detection of RFLPs 30
2.2.3 Random Amplified Polymorphic DNA 30
2.2.3.1 DNA Amplification 30
2.2.3.2 Detection of RAPD Markers 31

2.3 RESULTS AND DISCUSSION 31
2.3.1 Enzyme analysis 31
2.3.2 Restriction Fragment Length Polymorphism 34

2.4 DISCUSSION 35

CHAPTER 3. ANALYSIS OF THE GENETIC RELATIONSHIPS IN CITRUS FROM RAPD DATA, WITH PARTICULAR EMPHASIS ON CULTIVAR IDENTIFICATION 37

3.1. INTRODUCTION 37

3.2. AMPLIFICATION CONDITIONS AND OPTIMISATION 41

3.3. ANALYSIS OF RELATIONSHIPS BETWEEN CITRUS CULTIVARS 46
3.3.1. Sweet Oranges (*Citrus sinensis*) 48
3.3.2. Grapefruit (*C. paradisi*) 49
3.3.3. Pummelo/Shaddock (*C. grandis*) 52
3.3.4. Mandarins (*C. reticulata*) 54

3.4 PRACTICAL SOLUTIONS TO THE DIFFICULTIES IN THE PROTECTION OF NEW VARIETIES 59
3.5 DISCUSSION 61

CHAPTER 4. THE ROLE MOLECULAR TECHNIQUES CAN PLAY IN AWARDING PLANT BREEDER'S RIGHTS IN CITRUS 64

CHAPTER 5. GENERAL DISCUSSION AND CONCLUSIONS 69

REFERENCES 73

APPENDIX A. BUFFERS, STAINS AND SOLUTIONS

APPENDIX B. GLOSSARY OF TERMS

APPENDIX C. ABBREVIATIONS

APPENDIX D. DATA SETS

LIST OF TABLES AND FIGURES

TABLE 1. ISOZYME GENOTYPES 33

TABLE 2. CULTIVARS IN THE CIP 38

TABLE 3. APPLICATIONS FOR PLANT BREEDER'S RIGHTS 40

TABLE 4. SEQUENCES OF PRIMERS USED IN THE PCR 45

FIGURE 1. SYSTEMATICS OF CITRUS 4

FIGURE 2. RFLP PROFILES OF GRAPEFRUIT 34

FIGURE 3. EFFECT OF $MgCl_2$ CONCENTRATION ON THE PCR 43

FIGURE 4. PCR AMPLIFICATION PRODUCTS OF MANDARINS 45

FIGURE 5. MATHEMATICAL COMPARISON OF GENOTYPES 46

FIGURE 6. PCR-RAPD PRODUCTS OF NAVELS 49

FIGURE 7. PHYLOGENETIC TREE OF GRAPEFRUIT (NJOIN) 49

FIGURE 8. PHYLOGENETIC TREE OF GRAPEFRUIT (SAHN-UPGMA) 51

FIGURE 9. PHYLOGENETIC TREE OF GRAPEFRUIT (MIX) 51

FIGURE 10. PHYLOGENETIC TREE OF SHADDOCK (NJOIN) 53

FIGURE 11. PHYLOGENETIC TREE OF SHADDOCK (SAHN-UPGMA) 53

FIGURE 12. PHYLOGENETIC TREE OF SHADDOCK (MIX) 54

FIGURE 13. PHYLOGENETIC TREE OF MANDARIN (NJOIN) 56

FIGURE 14. PHYLOGENETIC TREE OF MANDARIN (SAHN-UPGMA) 57

FIGURE 15. PHYLOGENETIC TREE OF MANDARIN (MIX) 58

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Ralph Kirby for introducing me to RAPDs and for his support and interest in the work; also for assisting me in the final statistical analysis of the data.

I am grateful to the Agricultural Research Council for the financial support of this project, especially in the persons of Drs. J.H. Terblanche (former director of the Institute for Tropical and Subtropical Crops; currently President of the ARC) and J.H. Van Zyl, director of the ITSC.

My thanks to Dr. C.P. Du Plooy for his constant support and encouragement; also to my colleagues at Cultivar Development and Biotechnology, for sharing their expertise.

To my parents, Frikkie and Wiena Carstens and my sister Elsa, my deepest appreciation for all their love, encouragement and financial support during my studies.

Glory to my Creator, for physical strength and means to complete this assignment.

CHAPTER 1

GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

Citrus is among the most important economic fruit crops worldwide, and therefore the preservation and improvement of citrus germplasm is very important. The reason is that wild germplasm resources are scarce and/or inaccessible. At the same time, the introduction of wild resources into a breeding programme, is important, because of their specific characteristics like adaptability to different production areas, tolerance to parasites, cold, drought, salt, etc. Breeding is aimed at the introduction of new characteristics into existing cultivars, rather than the breeding of cultivars with totally different characteristics. The citrus genetic system is quite unusual. Citrus is almost entirely diploid, with 2n=18 chromosomes. A few polyploid (triploids and tetraploids) cultivars are known, but these are rare. Two major difficulties have delayed the progress of citrus genetics: the long juvenile period (5-10 years) of the progenies and the high frequency of apomixis by adventitious polyembryony. Interspecific hybridization is relatively easy, somatic hybridization now allows many new intergeneric hybrids to be created, and genetic engineering methods should overcome the long-generation time and allow breeders to alter specific traits.

1.2 THE ORIGIN OF CITRUS AND ITS RELATIVES

Citrus is considered to be native to South-East Asia, and especially to eastern India, but it shows phylogenetic relationships which extend through the East Indies, central China, Japan, Australia and Tropical Africa. The spread of the genus, *C. limon* (lemon), from the southern

slopes of the Himalaya in the direction of China where *Poncirus trifoliata* and *Fortunella japonica* were indigenous, was very slow (Hansen, 1983). By the close of the fifteenth century, the different species had reached almost all the tropical and subtropical sections of the Eastern hemisphere except Southern Africa. Here, exact evidence of the date of their introduction into South Africa and first fruiting was given in the daily journal of events kept by Van Riebeeck. The first sweet orange trees were brought from the island of St. Helena in 1654 - commonly used by the early voyagers as a stopping place - and evidently the orange, and perhaps other citrus trees, had been taken there from India as an intermediate point in their transfer to Europe (Webber, Reuther and Lawton, 1967).

1.3 CLASSIFICATION OF THE MAJOR CITRUS GROUPS

The status of species within the genus *Citrus* is in a state of contradiction. Lack of agreement reflects two basic problems: (1) what degrees of difference justify species status; and (2) whether supposed hybrids among naturally occurring forms should be assigned species rank.

Ferrari, Jonstonis and Volckamer paved the way in systematics of citrus varieties. Following Linaes's development of a classification system, Augustin P. de Candolle published the first comprehensive account of the orange subfamily in 1824.

The genus *Citrus* and its close relatives are members of the family Rutaceae. The Rutaceae family has seven subfamilies, with *Citrus* belonging to the subfamily Aurantioideae. The Aurantioideae is divided into two tribes, i.e. the Clauseneae and Citreae. Citreae has three

subtribes, the Triphasiinae, Citrinae and Balsamocitrinae. *Citrus* belongs to the subtribe, Citrinae. The subtribe Citrinae is divided into 13 genera (Hansen, 1983).

According to Swingle and Reece (1967), the true citrus fruit trees comprise of *Citrus* and five other genera, namely *Poncirus* (trifoliate orange), *Fortunella* (kumquat), *Eremocitrus* (xerophytic), *Microcitrus* and *Clymenia*, which constitute a group within the subtribe *Citrinae*. The commonly cultivated citrus fruits belong to three genera, *Citrus, Fortunella* and *Poncirus*, all closely related. The genus *Citrus* is divided into two very distinct subgenera, *Eucitrus* and *Papeda*. The subgenus, *Citrus*, includes all the commonly cultivated species of *Citrus*, all of which have pulp-vesicles filled with pleasantly tasting acid, subacid or sweet juice. On the contrary, none of the species of Citrus belonging to the subgenus *Papeda*, have edible fruits. Species can be divided into cultivars such as *Citrus sinensis* cult. *sanguinea* (blood orange). Apart from cultivars such as the Washington navel, Valencia and Eureka lemon, there are also various selections of cultivars. The well-known Palmer Washington navel, the Olinda Valencia and Frost nucellar Eureka lemon are selections commonly used in citrus circles. A wild species, *C. halimii* Stone was recently discovered in the highlands of Malaysia and Thailand (Roose, personal communication). The systematics of citrus according to Swingle, is given in Figure 1.

Many present-day citrus varieties have been cultivated since ancient times and their wild progenitors are usually not definitely known. The major groups of cultivated citrus (sweet oranges, mandarins, grapefruit and sour oranges) are believed to have derived from 3 species: *C. maxima* (pummelo), *C. medica* (citron) and *C. reticulata* (mandarin). A wide variety of clones are believed to have arisen by hybridization among these taxa.

DIVISION:	SPERMATOPHYTA
SUB-DIVISION:	ANGIOSPERMAE
CLASS:	DICOTYLEDONEAE
SUB-CLASS:	ARCHICHLAMYDEAE
ORDER:	GERIANIALES
SUB-ORDER:	GERANIINEAE
FAMILY:	RUTACEAE
SUB-FAMILY:	AURANTIOIDEAE
TRIBE:	CITREAE
SUB-TRIBE:	CITRINAE
GENERA:	*Fortunella*
	Poncirus
	Citrus

FIGURE 1. The classification of *Citrus* and its relatives according to Swingle (Swingle and Reece, 1967).

1.4 CITRUS BREEDING IN SOUTH AFRICA

Citrus has long been the object of intensive breeding programmes aimed at improvement of fruit quality, disease resistance and adaptability to varying climatic conditions (Torres, Soost and Mau-Lastovicka, 1982). South Africa possesses one of the largest gene banks in the world. Approximately 1 300 cultivars and selections are currently included and there is an ongoing search for superior cultivars. Although these selections are not all suited for commercial purposes, they are used as parents in the breeding programme. Citrus breeders, however are faced with two major problems. The first is the difficulty of distinguishing

nucellar (asexual) seedlings from zygotic seedlings. The second is that, when a new cultivar is derived, the trait whereby the selection was made, is often not phenotypic, e.g. yield, disease resistance, salt tolerance, etc. New cultivars arise through natural pollination, e.g. Ellendale tangor, or by means of controlled hand pollination, e.g. Minneola and Orlando tangelo (Ashari, Aspinall and Sedgley, 1989). New selections in a particular group, e.g. navels and mandarins are often so closely related that conventional morphological methods of identification cannot distinguish between new and existing selections.

1.5 THE STATUS OF GENETIC MARKERS IN CITRUS BREEDING

The identification of citrus cultivars was classically based on phenotypic assessment of morphological traits recorded in the field. Morphological features are still extremely useful, but alone can be ambiguous due to the influence of environmental conditions and cultural practices on identical genetic material grown on different locations. Furthermore, these methods involve a lengthy survey of plant growth that is costly, labour intensive and sensitive to the environment.

In promising selections, viruses (tristeza virus and exocortis viroid complexes) are removed through shoot tip grafting, whereafter it is pre-immunized with a mild strain of citrus tristeza virus. It is believed that the mild strain (which does not cause the disease), ensures some protection against the disease causing agent. These so-called superplants are multiplied at the Outspan Foundation Block situated at Uitenhage in the Eastern Cape, as a source of bud material for the citrus industry. The morphological characteristics of these trees are evaluated annually to verify trueness to type. In the current evaluation system, the

distribution of material is delayed at least three to four years since the trees must first reach productivity.

Seedlings derived from hybridization might exhibit variation which is often heritable and therefore results from genetic change. Several types of genetic changes occur, notably gene copy number, DNA mutations, transpositional changes, rearrangements and amplification of minor forms of the genome, like the chloroplast and mitochondria (Isabel, Tremblay, *et al.*, 1993). Because the genetic control of many of these characters is complex, their use in the detection of hybridization and pedigree determination is limited (Newbury and Ford-Lloyd, 1993). Karyological analysis cannot reveal alterations in specific genes or small chromosomal rearrangements.

Citrus reproduction is characterised by polyembryony, that is the formation of multiple apomictic (asexual embryos that originate from diploid nucellar cells) and usually only one zygotic embryo (Torres, Soost and Mau-Lastovicka, 1982). Among the multiple embryos, none, one or rarely more may be of sexual origin; the proportion of each type varies with genotype and environment. Most citrus cultivars produce seed with both nucellar and zygotic embryos, but some cultivars produce seed with only zygotic embryos. Nearly all commercial rootstock cultivars produce seeds with a high proportion of nucellar embryos, which favour genetic uniformity among the seedlings (Roose and Traugh, 1988). In the nursery, uniformity among plants is increased further by roguing (removal of phenotypically different plants) off-types from seedbeds. This procedure is important because citrus is generally very heterozygous and outcrossing does occur, resulting in sexual progeny that is often quite different from their maternal parent. If not eliminated, these zygotic seedlings may be

budded and produce undesirable trees when planted in the field. Upon germination one or several seedlings may emerge from a single seed, but generally only one will survive. The single zygotic embryo is usually suppressed by the numerous somatic embryos in the same ovule. Consequently, most seedlings produced from these cultivars are genetically identical to the pistillate (mother) parent (Preston, *et al.*, 1993). Most cultivated citrus types, including sweet orange, grapefruit, lemon and lime have nucellar embryony and are believed to have originated by natural hybridization. While nucellar seedlings obviously represent a method for propagating a certain genotype (e.g. rootstocks), the breeder is generally interested only in sexual recombinants for horticultural evaluation. In the past, this has restricted the citrus breeder to the few parent cultivars which are obligate monoembryonics, producing zygotic embryos only.

In most crossing combinations it would not be possible to determine the genetic origin of young seedlings from morphology because of the absence of a phenotypic marker such as trifoliate leaves (when *Poncirus trifoliata* - a citrus relative - is used as one of the parents). Yet, such a determination is essential in order to avoid the 5- to 10-year costs of growing out and maintaining until fruiting when unwanted nucellar seedlings that are genetically identical to already available seed parents, and which easily can be propagated by grafting or budding.

Developing reliable and discriminatory methods for identifying cultivars has become increasingly important to plant breeders and nurserymen who need sensitive tools to differentiate among and identify cultivars for plant breeder's rights. In the past, cultivar and race identification in plants was limited to horticultural, morphological and physiological descriptions. In most cases, the descriptions and measurements varied considerably due to

environmental fluctuation and human judgement (Torres *et al.*, 1993). The development of new cultivars that lack distinguishing morphological characteristics has furthered the need for more sensitive identification methods.

At the inception of a breeding programme, knowledge of the genetic relationships among genotypes can be used to complement phenotypic information in the development of breeding populations. A major issue in genetic resources has been the size of collections in relation to their effective management and use. In formal breeding programmes, effective management of the citrus collection requires solutions to several problems. One problem is distinguishing plants with unique genes from accessions already in the collection. Citrus cultivars are often given local names - the collection may thus contain many duplications; the same clone under different names. Another problem is in setting priorities for processing of foreign imports. Often, some of the imported material collected is simply a duplication of germplasm already in the collection (Roose, personal communication). Any technique which would allow screening for duplicate samples, would be an asset to genebank work. Little attention is usually paid to the measurement of genetic diversity because of the problems involved with its accurate measurement in large numbers of organisms. There is therefore a great demand for techniques which do truly measure genetic polymorphisms, and which could be applied on a large scale. A range of plants characteristics are available for distinguishing between closely related individuals. Classical phenotypic features are still extremely useful, but they require that a plant be grown to a suitable developmental stage before certain characters can be scored.

Until recently, virtually all progress in both breeding and model genetic systems has relied on a phenotypic assay of the genotype. Because the efficiency of a selection scheme or genetic analysis based on phenotype is a function of the heritability of the trait, factors like the environment, multigenic and quantitive inheritance, or partial and complete dominance often confound the expression of a genetic trait. As an alternative, a number of laboratory methods have been developed in the past two decades, such as isozyme analysis, storage protein electrophoresis and high performance liquid chromatography (HPLC) of various substances in the seeds and fruit (Hu and Quiros, 1991). Several research groups have attempted to resolve citrus taxonomy and identification by using biochemical markers, such as the analysis of different chemical compounds such as the essential oils, flavonoids, cuticular waxes (Gogorcena and Ortiz, 1988) and more recently electrophoretic separation of proteins and isoenzymes (Soost and Williams, 1980). All have had their limitations, ascribable in large part to a lack of understanding of the underlying genetic control of the compounds studied.

Many of the complications of a phenotype based assay can be mitigated by direct identification of the genotype with a DNA-based diagnostic assay. For this reason, DNA-based genetic markers are being integrated into several plant systems (Tingey and del Tufo, 1993). Recent advances in techniques for DNA analysis and subsequent data analysis have greatly increased our ability to understand the genetic relationships among organisms at molecular level. A range of plant characteristics are currently available for distinguishing between closely related individuals. Classical phenotypic features are still extremely useful, but can sometimes be influenced by environmental conditions. They require that a plant be grown to a suitable developmental stage before certain characters can be scored.

DNA-based markers allow the direct comparison of the genetic material of two individual plants avoiding environmental influences and gene expression. RFLP (restriction fragment length polymorphism) is widely used for gene mapping and studying diversity in plant populations. VNTR (variable number of tandem repeats) identifies repeated DNA regions of differing lengths resulting from variable numbers of serial repeats of a core DNA sequence. These core sequences are referred to as mini-satellites or micro-satellites. Another technique is that of PCR-sequencing. The technique reveals variation at a very high level of resolution, since differences are measured at the base level (Newbury and Ford-Lloyd, 1993). More recently, a rapid and sensitive technique using the polymerase chain reaction (PCR) has been introduced, namely RAPD (random amplified polymorphic DNA) which offered advantages in speed, technical simplicity and the frequency of identification of polymorphisms (Williams *et al.*, 1990; Welsh and McClelland, 1990).

1.6 TECHNIQUES, OTHER THAN MORPHOLOGICAL IN THE ORGANIZATION OF GENETIC DIVERSITY CONCERNING CITRUS

1.6.1 Electrophoresis of proteins and isozymes

The development of isozyme and other biochemical markers represented a significant improvement in the management of germplasm collections since they offer greater diversity (Weining and Langridge, 1991).

Isozymes (different molecular forms of an enzyme) have often been used as markers in genetic and taxonomic studies of citrus (Wenpin, Shanwen and Gengfeng, 1988). Isozymes

are found in all plants. Isozyme profiles have been used to distinguish between cultivars of avocado's, olives, apples and citrus (Gogorcena, Zubrzycki & Ortiz, 1990).

Citrus cultivars are very heterozygotic. The genetics of perceptible characteristics are complex since one characteristic is often controlled by more than one gene (Gogorcena, et al., 1990).

Leaves, bark, fruit and pollen serve as good sources for enzymes (Gogorcena, et al., 1990). Protopapadakis (1987) found that pollen is an excellent source for isozymes as results are reproducible and the physiological condition of pollen is constant. However, pollen is not always readily available therefore leaves are preferred. The isozyme composition of leaves is influenced by the age of the plant as well as the interaction between scion and rootstock (Protopadakis, 1987).

Isozymes can be separated by electrophoresis on starch, polyacrylamide, or agarose gels, or by isoelectric focusing. Starch gels are preferred to polyacrylamide gels because of their convenience for analysis of large numbers of individuals, and their resolution is adequate for most purposes.

A variety of isozyme systems can be used for the identification, however, according to results obtained by Wenpin *et al.* (1988), peroxidase gives better results than polyphenol oxidase and cytochrome oxidase. However, for identification purposes peroxidase is not recommended as it yields undefined bands, probably due to the uncertainty of this enzyme's

exact substrate. The expression of peroxidase has been associated with greening disease of citrus (Dr. L. van Lelyveld, personal communication).

Protopapadakis (1987) examined four enzyme systems, namely esterase, acid phosphatase (slow and fast types) and peroxidase. Esterase produced the best results for cultivar identification. The electrophoretic separation of this enzymes from leaves and peel resulted in an unique banding pattern for each cultivar which was not influenced by the rootstock. This enzyme is more suitable as a genetic marker than the others as the gene is collinear (there is a direct relationship between the sequence of the mutation sites within the gene and the altered amino acid sequence of the enzyme for which it codes). The allele that specifies the enzyme is co-dominant, therefore there is no suppression of the gene's expression.

The following isozyme systems have been analyzed in *Citrus* and *Poncirus*: Aspartate aminotransferase, Phosphoglucose isomerase (PGI), Phosphoglucomutase (PGM), Malate dehydrogenase (MDH), Leucine aminopeptidase (LAP), Hexokinase, Isocitrate dehydrogenase (IDH), Malic enzyme (ME), Superoxide dismutase, Peroxidase, Polyphenoloxidase, Cytochrome oxidase and Esterase.

On a zymogram, symbols are assigned to the subunits of the enzyme. F specifies fast, S slow and I intermediatory migrating subunits. The number of bands on a zymogram are an indication of the number of alleles of the gene which code for that specific enzyme. Additional bands on the zymograms are the result of complexes with cofactors, ions or in-complete transcripts or translations of polypeptides (Torres et al., 1978). Double bands are

often found in other organisms, however, in citrus, it is not clear whether the unusual banding patterns are caused by non-standard electrophoretic methods (Torres et al., 1982). Ashari et al. (1989) while studying Ellendale trees (tangerine cultivar) discovered that one of the trees differed from the other trees at two loci. This variation can be ascribed to a number of factors such as mutations, multiple origin of material, incorrect tree identification and propagation through mono-embryonic seed. To minimize these influences on variation, plants of the same age, grown under the same conditions, should be selected for isozyme studies. Samples should also be taken at the same time.

The relative migration velocity (rMV) of an enzyme is described as the coefficient of the distance the enzyme migrated in the gel divided by the distance of the bromophenol blue from the cathode. The gene, *MeO2*, specifies the fast moving malate isozyme in Citrus and Poncirus, *Sod1* the slower migrating group of the superoxide dismutase isozymes.

As an adjunct to morphological and physiological methods, identification tests based on isozyme and protein patterns have been introduced to fingerprint cultivars of various species. Codominant markers such as isozymes are particularly useful for distinguishing nucellar from zygotic seedlings because most citrus genotypes are quite heterozygous. An additional advantage of using isozymes over other biochemical markers is that enzymes are nearly all direct gene products and products of a series of biosynthetic reactions such as those leading to the production of pigments, oils and various other classes of compounds (Torres *et al.*, 1978). The use of individual proteins as molecular markers offers advantages since proteins are the direct products of individual genes. Protein markers are normally compared by monitoring their migration in gels during electrophoresis; they are detected by using either

a general protein stain or a stain to detect a specific enzyme (iso-enzyme analysis) (Newbury and Ford-Lloyd, 1993). Again, environmental effects can influence results, but this is not normally regarded as a problem.

1.6.2 DNA profiling

Recent advances in techniques for DNA analysis and subsequent data analysis have greatly increased our ability to understand the genetic relationships among organisms at the molecular level.

With the advent of molecular techniques, DNA-based procedures have been proposed for cultivar identification. This include DNA profiling or fingerprinting. A DNA profile (or "DNA fingerprint") is a visual product derived from an analysis of some parts of the DNA molecule. It allows the analysis of any area of the genome that displays polymorphism, which is most often present in non-coding portions of the DNA molecules. DNA profiles can be likened to a 'bar-code' or a human fingerprint (Jeffreys, Wilson and Thein, 1985). It is the identifying of the characteristics of the product or individual, but it bears no relationship to the appearance or performance of that product or individual. DNA fingerprinting is more precise in determining the degree of relatedness among individuals than classic protein polymorphisms and it provides information that cannot be obtained from field observation and morphogenetic studies.

DNA-based markers allow the direct comparison of the genetic material of two individual plants avoiding any environmental influences on gene expression (Newbury and Ford-Lloyd, 1993). The structure of DNA (on which these techniques are based) will not be influenced

by external factors. This is in contrast to the quantity and nature of all other molecules which are more or less influenced by climate, environment, latitude or developmental stage of the plant. There are currently two major DNA profiling techniques of relevance to varietal characterization: Restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) analysis.

1.6.2.1 Restriction Fragment Length Polymorphism

In 1968, molecular biologists discovered that DNA could be cut into pieces with various types of restriction enzymes. Every DNA molecule has a unique restriction map for a particular restriction endonuclease enzyme. One enzyme would cut the DNA only at one specific sequence of base pairs, and others at other sequences. The fragments of DNA could be subjected to electrophoresis (migration of the molecules in an electric field) and separated according to the molecular weight of the fragments. If DNA from a particular strain was cut by a particular restriction enzyme and separated by electrophoresis repeatedly, the fragments always migrated in the same order. However, if the same enzyme was used to cut the DNA of another strain, a few differences would appear when compared with the banding pattern of the first. This was found to be due to small but appreciable differences in the size of certain fragments. If a mutation occurred in the target sequence for a restriction endonuclease and the DNA was exposed to the enzyme, one cut normally made in the wild-type DNA would not be made, and the pattern of bands formed after gel electrophoresis of the digest would differ: thus two bands present in the digest of wild type DNA would be missing and a new band would be present whose molecular weight equals the sum of the molecular weights of the missing bands (Irvine and Moore, 1991).

Southern (1975) devised a technique that transferred DNA from the electrophoresis gel to a nitrocellulose membrane where differences between individuals could be repeatedly detected by various radioactive nucleic acid probes. These were named restriction fragment length polymorphisms (RFLPs). RFLPs depend on the use of probes to identify single or low copy sequences in DNA. RFLPs were quickly seized as potential markers and a milestone paper (Botstein *et al.*,1980) appeared with the radical proposal that RFLPs could be used to map human genes.

RFLP has been used widely for gene mapping and studying diversity in plant populations. It was proposed as an additional approach to detecting genetic polymorphisms at the molecular level in plant species. Burr *et al.* (1983) suggested the use of RFLP in estimating diversity and mapping genes that control quantitively inherited traits. RFLP markers are useful for sampling various regions of the genome and are potentially unlimited in number. RFLP are also helpful in assessing citrus organization: cultivar identification, heterozygosity and phylogeny. RFLP have a greater resolving power than isozyme studies. It can therefore be applied for exploring a greater part of the genome.

RFLP alleles are typically codominants, thus offering the possibility to distinguish both homozygous and heterozygous structures. Most of the variations do not modify the phenotype directly. RFLPs are very convenient markers for establishing phylogenetic relationships between taxa. The expression of these markers does not depend upon environmental conditions and this is essential for cultivar identification. Many of these markers are not subject to selective pressure or undergo convergent evolution. This is an important advantage for phylogenetic studies. RFLP analysis, while providing significant

information about the organism's genotype, however, they are time-consuming, costly and they require large amounts of plant tissue.

RFLPs occur because the DNA of different individuals differs in the presence or absence of a specific 4-8 base pair sequence which is recognized and cleaved by a restriction endonuclease. Alternatively, restriction endonucleases can detect polymorphisms which occur due to single nucleotide changes. DNA of the genotypes may differ by insertions, deletions or other rearrangements which alter the distance between a pair of restriction sites. Recognition sites for restriction enzymes are thus destroyed, or created.

RFLPs are detected by isolating DNA from each genotype, digesting it with a restriction enzyme, separating the fragments according to size using electrophoresis, transferring the fragments to a membrane which binds the DNA (Southern blot), and hybridizing the DNA bound to the membrane with a specific cloned DNA sequence which is homologous to some of the DNA on the membrane. This "probe" DNA binds to only those fragments with which it is homologous, and these fragments are visualized by autoradiography if the probe is radioactively labelled, or visually if using chemically modified probes.

Although these variable loci occur as arbitrary sites they are not necessarily associated with specific genes. Those which are, e.g. *Hpa* I (Sickle Cell Anaemia in humans), are valuable as polymorphic markers. It is used in the construction of genetic linkage maps which are used to study Mendelian inheritance of genetic traits.

RFLP alleles are typically codominants, thus offering the possibility to distinguish both homozygous and heterozygous structures. Most of the variations do not modify the phenotype directly. RFLPs are very convenient markers for establishing phylogenetic relationships between taxa. The detection of these markers does not depend upon environmental conditions and this is an important advantage for cultivar identification.

1.6.2.2 Random Amplified Polymorphic DNA

Over two years ago, a new genetic assay was developed independently by two different research groups (Welsh and McClelland, 1990; Williams *et al.*, 1990.). This technique, called "random amplified polymorphic DNA" (RAPD), uses short synthetic deoxyribonucleotides of arbitrary sequence to detect polymorphisms between genotypes in a DNA amplification-based-assay. RAPDs have been proposed as a new species of genetic marker, one that overcomes many of the technical limitations of RFLP analysis (Williams *et al.* 1990; Welsh and McClelland 1990; Rafalski *et al.* 1991). RAPD would find applications in the identification of cultivars and selections by genomic fingerprinting. The basis of RAPD methodology is the "polymerase chain reaction" (PCR).

PCR is an ingenious new tool for molecular biology that has had an effect on research similar to that of the discovery of restriction enzymes and the Southern blot. PCR has provided an alternative approach to many procedures in molecular biology and is replacing many standard techniques. It can be used to amplify specific target sequences for subsequent cloning, and it provides an extremely sensitive method for the detection of specific DNA and RNA sequences. This has led to the use of PCR to reveal variability of simple sequences in eukaryotic genomes (Weining and Langridge, 1991). PCR can amplify double- or single-

stranded DNA, and with the reverse transcription of RNA into a cDNA copy, RNA can also serve as a target (Erlich Gelfand and Sninsky, 1991). PCR is so sensitive that a single DNA molecule has been amplified, and single copy genes are routinely extracted out of complex mixtures of genomic sequences and visualized as distinct bands on agarose gels (Innes, *et al.*, 1990). The use of thermostable DNA polymerases and automation have fostered the development of numerous and diverse PCR applications.

PCR is an *in vitro* method of nucleic acid synthesis by which particular segments of DNA can be specifically replicated. The principles involved in the PCR are not novel, but, like so many previous advances in DNA handling, it results from the availability of an enzyme with useful properties. Organisms that live in hot springs are a rich source of thermostable enzymes, and the development of the PCR has been carried out with the DNA copying enzyme Taq polymerase (from *Thermus aquaticus*) which works optimally at 72°C. This means that this enzyme can be used to make copy strands of DNA under the elevated temperature conditions employed.

The PCR is based on the annealing and extension of one or more oligonucleotide primers that flank the DNA fragment to be amplified. After denaturation of the target DNA, the primer binds to one of the two separated strands such that the extension from each 3' hydroxyl end is directed toward the other (Erlich *et al.*, 1991). The enzyme can only start to copy from a double-stranded piece of DNA and this is provided by the primer/genomic DNA duplex. DNA sequences have polarity, and the enzyme can only extend the primer in one direction. Replication starts from those points of hybridization, producing varying lengths of DNA, depending on how close together the primers were on opposing DNA strands. The lengths

of the products generated during the PCR is equal to the sum of the lengths of the primers and the distance in the target DNA between the primers. When the temperature is raised to above 90°C, the strands of genomic DNA separate. Lowering the temperature to 37°C, allows the primer to anneal to specific regions of the strands. At temperatures between 35 and 50 °C, primers can be expected to anneal to many sequences with a variety of mismatches. Some of these will be within a few hundred base pairs of each other and on opposite strands. Sequences between these positions will be PCR amplifiable. The extent to which sequences will amplify, depend on the efficiency of priming at each pair of primer annealing sites and the efficiency of extension (Welsh and McClelland, 1990). Raising the temperature to 72°C removes those primers that are not tightly hybridised. DNA strand elongation takes place at this temperature. These three steps (denaturation, primer binding, and DNA synthesis) represent a single PCR cycle (Erlich *et al.*, 1991).

A key feature of the process is that re-copying the copy strand results in a DNA sequence of defined length. Since the extension products are also complementary to and capable of binding primers, successive cycles of amplification essentially doubles the amount of target DNA synthesized in the previous cycle. The result is an exponential accumulation of the specific target fragment, approximately 2^n, where n is the number of cycles of amplification performed. A discrete DNA product is produced through thermocyclic amplification. This process is allowed to continue through a number of replications so that the fragments created are in sufficient quantities to be visualised on an agarose or polyacrylamide gel. When one subjects the reaction mixture to electrophoresis, then it is this amplified sequence that one sees as a DNA band. The number, reproducibility and intensity of bands in a fingerprint is dependent on the concentration of salts, annealing temperature, template concentration,

primer length and primer sequence (Welsh and McClelland, 1990). Each primer of similar length but of different sequence gives a different set of patterns, as expected since the template/primer interactions are different. The reason for the enormous DNA duplication rate is that once a DNA sequence is copied, the copy can be copied so that an exponential rate of increase in the number of copies of that sequence occurs.

While the PCR products are produced from random regions of the genome, they are specific and reproducible since they are primed from specific DNA sequences within the genome. The polymerase chain reaction has led to the development of several novel genetic assays based on selective DNA amplification. One of the advantages of these assays is that they are more amenable to automation than conventional techniques (Tingey and del Tufo, 1993). The utility of DNA-based diagnostic markers is determined by the technology that is used to reveal polymorphisms.

The presence of each amplification product identifies complete or partial nucleotide sequence homology between the genomic DNA and the primer at each end of the amplified region. Each primer will direct the amplification of several discrete loci in the genome, making the assay an efficient way to screen for nucleotide sequence polymorphism between individuals. The number of potential 10-base primers is unlimited so that numerous polymorphisms between even closely related organisms can be obtained (Stiles, *et al.*, 1993). The major advantage of this assay is that no prior information of the DNA sequence is required.

Random Amplified Polymorphic DNA (RAPD) is based on the *in vitro* expression of some polymorphic regions of the DNA molecule. By using single 10 base primers of arbitrary

sequences to amplify DNA with PCR, fingerprint-like bands are generated. RAPD markers are capable of detecting polymorphisms in single copy and repetitive DNA, resulting from insertions, deletions, rearrangements, or single base changes (Echt, Erdahl and McCoy, 1991). Because RAPD polymorphisms are the result of either a nucleotide base change that alters the primer binding site, or an insertion or deletion within the amplified region (Williams *et al.*, 1990) polymorphisms are usually noted by the presence or absence of an amplification product from a single locus. The bands detected by RAPD amplification are usually treated as dominant markers. Individuals containing two copies of the same allele (homozygotes) are not readily distinguishable from those containing two different alleles (heterozygotes), although it is possible by analysis of band segregation in crosses to identify co-dominant bands (Tingey and del Tufo, 1993). RAPDs are useful for DNA fingerprinting where there is a need to identify varieties of a crop species or to determine parentage in breeding material. The utility of RAPDs as genetic markers has been demonstrated in soybean (Williams *et al.*, 1990). In addition to varietal identification, phylogenetic relationships, particulary at intra-specific level, can be examined using RAPDs. Since these markers can be shown to segregate in a Mendelian fashion (Torres *et al.*, 1993) it can result in the early selection of progeny in a breeding programme, saving considerable time in the production of new varieties. Random Amplified Polymorphic DNA (RAPD) offers advantages in speed, technical simplicity and the frequency of identification of polymorphisms and need for very small amounts of genomic DNA which allows the analysis of young seedlings (Hu and Quiros, 1991).

A key feature of RAPD is that the primers used, passes a base sequence that is arbitrarily defined. Although the primer sequence is known, the investigator has no idea to which, if

any, gene or repeated sequence in the plant genome the primer may have homology. Any bands observed in a gel can be used for the comparison of genotypes.

Fingerprints are generated from DNA segments that amplify in DNA preparations from one parent but not the other. While the PCR products are produced from random regions of the genome, they are specific and reproducible since they are primed from specific DNA sequences within the genome, providing that all of the reaction conditions are consistent. In order that the randomly-defined primers result in the amplification of some sequences, short primers are usually employed. The approach of Williams *et al.* (1990) uses a 10-base synthetic oligonucleotides with a GC content of 50-60%. Purely on average, a 10-mer will hybridise to a strand of DNA about once every million bases. Since the higher plant genome is very large, several amplified fragments are normally observed when one 10-mer is used. The products are separated by standard agarose gel electrophoresis and visualised with ultraviolet light by staining with ethidium bromide. Polyacrylamide gels are also used and combined with silver staining of DNA to increase the resolution and detection of less amplified fragments.

The essential feature of RAPDs, is the identification of polymorphism by the detection of differences in DNA occurring between plants. Detected DNA polymorphisms are inherited in a Mendelian fashion and can be produced from any species without any other DNA sequence information.

RAPD markers have been used to create DNA fingerprints for the study of individual identity and taxonomic relationships in both eukaryotic and prokaryotic organisms (Caetano-Anollés

et al., 1991; Hu and Quiros, 1991). Several groups have reported on the utility of RAPD markers as a source of phylogenetic information. Hu and Quiros (1991) showed that the amplification products from only four random primers were sufficient to discriminate between 14 different broccoli and 12 different cauliflower cultivars (*Brassica oleracea* L.).

1.6.2.3 Amplified Fragment Length Polymorphism

AFLP is a PCR based technique in which small restriction fragments, obtained by cleaving genomic DNA with restriction enzymes, are amplified to produce simple DNA fragment patterns. The technique is based on the use of specifically designed PCR primers which selectively amplify small fragments out of a complex mixture. The method yields DNA fingerprints which can be programmed by the choice of restriction enzymes and primers. These AFLP fingerprints display 50 to 100 bands. According to Zabeau (1992) the fingerprints are repeatable both in terms of banding patterns and the relative intensity of each band. The frequency of AFLP polymorphisms closely match that found using conventional RFLP analysis. Several pilot studies have been conducted in which AFLP fingerprinting was used to measure genetic distance between commercial varieties of hybrid corn, tomato, lettuce, cucumber and oilseed rape.

1.7 OBJECTIVE

The objective of this study was the development of laboratory techniques to differentiate between citrus cultivars in aid of the following purposes:

1. To differentiate between nucellar and zygotic seedlings in Phase I of the breeding programmes where only the hybrid is of importance to the breeder.

2. To verify the trueness to type of a particular commercial or semi-commercial cultivar or selection before the budwood is released to the industry.
3. To assist UPOV (Union for the Protection of New Varieties) in the awarding of plant breeder's rights when morphological traits are not sufficient to ensure significant distance between new and existing varieties.
4. To screen for uniformity in rootstock seedlings, where only the nucellar seedlings are important.
5. To implement this technique as a diagnostic service to the citrus industry.

CHAPTER 2

EVALUATION OF ISOZYME, RFLP AND RAPD TECHNIQUES FOR THE DISCRIMINATION BETWEEN CITRUS CULTIVARS

2.1 INTRODUCTION

The aim in adopting a technique for processing large amounts of plant samples on a routine basis, was to minimize the labour and cost involved. The methods described in the literature are time-consuming, not cost effective, sometimes biohazardous and often designed for specific laboratory use and not applicable to the industry where the demands are mostly more technical than research orientated. This investigation was designed to evaluate the potential and applicability of molecular and biochemical markers for assessing the genetic integrity of citrus cultivars which are important to both the breeder of new cultivars, the nurseryman who has to propagate material of ultimate quality and to the commercial citrus farmers.

Citrus was chosen as the crop of interest because of the problems that exist in the awarding of plant breeder's rights. Cultivars and selections used in this study are either included in the Citrus Improvement Programme of the Institute for Tropical and Subtropical Crops or awaiting plant breeder's rights awarded by the International Union for the Protection of New Varieties of Plants (UPOV).

2.2 MATERIALS AND METHODS

2.2.1 Isozyme analysis

2.2.1.1 Extraction and separation of isozymes.

The following enzyme-systems were analyzed: Phosphogluco-mutase (PGM), Phosphogluco-isomerase (PGI), Glutamate oxaloacetate transaminase (GOT), Isocitrate dehydrogenase (IDH) and Malate dehydrogenase (MDH).

Approximately 1 cm^2 leaf discs were grounded with a glass rod, in 150 μl extraction buffer (0.025 M, Potassium-phosphate buffer, pH 7.5, 12% PVP-40, 2% ß-mercaptoethanol). Filter paper wicks (Whatmann 3 MM) were saturated with the extract and loaded with a pair of forceps onto a prepared 10.5% starch gel, containing 3.5% sucrose. IDH and MDH were run in a pH 7.0, 0.3 M Tris-citrate electrode buffer; the gel buffer was 0.03 M Tris-citrate, pH 7.0. Gels were run overnight at 10 mA. GOT, PGM and PGI were run in 0.37 M Sodium-borate electrode buffer, pH 8.7; the gel buffer was 0.019 M Tris-citrate, pH 7.7. This gel was run at 35 mA for approximately 4 hours. Gels were stained specifically for individual enzymes[1]. Gels were sealed in plastic bags after the addition of the stain. Colour development took place at 37 °C in the dark. Results were scored as soon as isozyme bands could be detected.

[1]Composition of stains are given in Appendix A.

2.2.2 Restriction Fragment Length Polymorphism

2.2.2.1 DNA Extraction.

DNA was extracted from fresh, young leaves according to a method by Doyle and Doyle (1991) with slight modifications. The extraction buffer (3% (w/v) CTAB, 1.4 M NaCl, 0.2% (v/v) ß-mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl, pH8.0) was preheated to 60 °C. Approximately 0,8 grams of fresh tissue were ground to a fine powder in liquid nitrogen. The powder was scraped into 15 ml polypropylene centrifuge tubes containing 5 ml extraction buffer. The tubes were gently swirled to mix and incubated at 60 °C for 30 to 40 minutes. The mixture was extracted once with chloroform:isoamyl-alcohol (24:1). Cellular debris was removed by centrifugation at 5 000 rpm for 10 minutes at room temperature. The aqueous phase was transferred with a pipet to a clean tube and 4 ml cold isopropanol was added and left for one hour at room temperature to precipitate the DNA. DNA was recovered by centrifugation at 10 000 rpm for 10 minutes at 4 °C. The pellet was then washed in wash buffer (76% ethanol, 10mM ammonium acetate) for at least 20 minutes (this was a convenient stop for the night). The loose pellet was centrifuged at 5 000 rpm for 5 minutes. The supernatant was poured of gently, not to loose the pellet. Tubes were inverted on paper towel to allow the pellet to dry slightly. The DNA pellet was resuspended in 1 ml TE buffer (10 mM Tris-HCl pH7.4, 1mM EDTA). RNAse was added to a final concentration of 10 μg/ml and incubated for 30 minutes at 37 °C. Usually the DNA is suitable for the polymerase chain reaction at this stage. However, when restriction enzyme digests were being done, the DNA had to be highly pure and free from contaminants such as tannins. A phenol:chloroform extraction was usually necessary to remove proteineous contamination. This step was followed by the addition of 7.5M ammonium acetate to a final concentration of 2.5M and 2.5 volumes of cold ethanol to precipitate the DNA. For RFLP,

DNA was resuspended in TE buffer to a final concentration of 500 μg/ml. Five microlitres of the prepared DNA was run on a 1% agarose gel to check the quality and estimate the quantity thereof.

2.2.2.2 Restriction enzyme digests.

Approximately 50 μg plant DNA was digested with 50 units *Eco* RI, *Eco* RV, *Bam* HI, *Hind* III and *Pst* I in a 100 μl reaction. The reaction was allowed to continue 12 - 18 hours at 37 °C. Resultant fragments were separated on a 0.8% agarose gel in TBE (Tris-borate-EDTA) buffer for 14 - 16 hours at a constant 50 V.

2.2.2.3 Preparation of DIG- and ^{32}P-labelled probes.

A cDNA library of Rough Lemon was obtained from Prof. Mikael L. Roose (University of California, Riverside, USA). cDNA was cloned into the *Pst* I restriction site of the pUC9 plasmid. The size of the inserts varied between 340 and 1 600 base pairs. Nineteen cDNA clones from this library were screened as potential probes for the detection of RFLPs among citrus cultivars.

Plasmids were extracted from the host bacteria (*Escherichia coli* JM83) by an alkali-method (Ish-Horowicz and Burke, 1981). The inserts were recovered from *Pst* I digests and purified with Gene Clean^R. Inserts were labelled with radio-active dCTP or with non-radio-active digoxigenin-11-dUTP (a nucleotide analog) (Boehringer Mannheim, W-Germany) according to the manufacturer's instructions.

2.2.2.4 Southern blot, Hybridization and detection of RFLPs.

The gels were depurinated in 0.25 M HCl until the bromophenol blue dye changed from blue to yellow (10 - 15 minutes). DNA was denatured by soaking the gel in 0.5 M NaOH for 30 minutes. After it was rinsed in distilled water, it was soaked for 60 minutes in 1 M NH_4OAc and 0.02 M NaOH. The gel was alkali blotted to a positively charged Hybond N^R nylon membrane by capillary force according to the protocol of Maniatis *et al.* (1987). The blotting reaction was allowed to proceed for 18 - 24 hours. The membrane was rinsed in 0.5 M Tris-Cl, pH7.2 and 1 M NaCl for 15 minutes and air dried for 30 minutes. The membrane was prehybridized for 4 to 6 hours in a commercial hybridization fluid (Boehringer Mannheim). Heat denatured probes were added to fresh hybridization fluid and hybridization was allowed for 12 to 18 hours at 65 °C. Unbound probe was removed by a low-stringency wash buffer (2X SSC, 0.5% SDS), followed by a high stringency wash (0.1X SSC, 0.1% SDS) at 68 °C. When a radio-active probe was used or when AMPPD (energy is emitted as light pulses) was used as the substrate for a DIG-labelled probe, RFLPs were detected by exposing the membrane to X-Ray film for the required time. A colour substrate was also used with a DIG-labelled probe colour substrate.

2.2.3 Random Amplified Polymorhic DNA

DNA was prepared in the same way as was described for RFLPs.

2.2.3.1 DNA amplification.

The polymerase chain reaction (PCR) was conducted in a JDI 8012 automated thermocycler (JD Instruments, Noordhoek, Cape Town) by combining 1 μl undiluted DNA, 5 μl 10X Taq Polymerase Buffer, 10 picamoles primer, 400 μmoles deoxynucleotides (adenine, guanine,

thymine and cytocine), 1.5 mM $MgCl_2$ and 0,75 U Taq Polymerase (Promega, Madison, Wisconsin) in a 50 μl reaction volume. The reaction mix was overlaid with 50 μl paraffin oil. The thermocycler was programmed for 35 cycles at 90 °C for 45 seconds, 37 °C for 30 seconds and 72 °C for 120 seconds for denaturing, annealing and primer extension respectively. An initial denaturing step at 90 °C for 120 seconds and a final extension step at 72 °C for 600 seconds were included. Running time for this programme was approximately 4 hours.

2.2.3.2 Detection of RAPD markers.

After all the cycles were completed, 20 μl of each of the amplified products were loaded on 1,4% agarose gels in 1X TBE buffer (Tris-Borate-EDTA, pH 8.0) and run at ca. 10 V per cm gel for approximately 3 hours. Gels were stained with 0.5 μg/ml ethidium bromide for 20 minutes and destained for 10 minutes in tap water. Gels were photographed under UV light with Polaroid film 667. In order to confirm cultivar specific markers, the amplification was repeated at least twice.

2.3 RESULTS

2.3.1 Enzyme analysis

Although they are believed to be quite heterozygous, all the sweet oranges (*C. sinensis*) have the same isozyme phenotype for PGI, IDH and MDH (Table 3). PGM was the only enzyme which exhibited polymorphisms among the navels. No polymorphism was detected in the four Valencias studied.

Henderson was the only grapefruit cultivar which deviated from the isozyme phenotypes observed in all the other cultivars of this group (*C. paradisi*). Henderson had a FS genotype for PGI-1 and not SS like the other grapefruit cultivars. Its isozyme pattern for PGM were also different from the other grapefruit. Henderson was selected from a bud mutation in Ruby Red. Could this be a case where the expression of the isozyme phenotype was indeed changed by a bud mutation? Minor variation was detected in GOT, but it was impossible to score the genotypes of the cultivars, because of the difficulty of interpretation of GOT patterns.

IDH was the only enzyme which was polymorphic among the shaddocks (*C. grandis*). Tahiti pummelo and Chandler were homozygous - II and MM respectively - for IDH. Tahiti pummelo was FS for PGM, whilst Melogold, Chandler and Oroblanco were SS for PGM. The second band observed in Tahiti pummelo's PGM zymogram could have been a ghost band. It was not possible to confirm the FS phenotype of this cultivar because of the unavailability of a control cultivar which was definitely FS for PGM.

TABLE 3. Isozyme phenotypes of some citrus cultivars for five enzyme systems.

GENUS	CULTIVAR	PGI	PGM	GOT	IDH	MDH
C. sinensis (Navel)	Californian Lane Late	FS	FM	-	MI	FS/FF
	Fisher	FS	GM	-	MI	FS/FF
	Gillemberg	FS	GM	-	MI	FS/FF
	Lane Late	FS	FM	-	MI	FS/FF
	Leng	FS	FM	-	MI	FS/FF
	Painter Early	FS	GM	-	MI	FS/FF
	Patensie	FS	FM	-	MI	FS/FF
	Robyn	FS	GM	-	MI	FS/FF
	Royal Late	FS	GM	-	MI	FS/FF
	Washington	FS	FM	-	MI	FS/FF
C. sinensis (Valencia)	Broedershoek	FS	GM	-	MI	FS/FF
	Delport	FS	GM	-	MI	FS/FF
	Margaret	FS	GM	-	MI	FS/FF
	Turkey	FS	GM	-	MI	FS/FF
C. paradisi (Grapefruit)	Flame	SS	SS	-	II	FS/FF
	Henderson	FS	FI	-	II	FS/FF
	Nelruby	SS	SS	-	II	FS/FF
	Oran Red	SS	SS	-	II	FS/FF
	Ray Ruby	SS	SS	-	II	FS/FF
	Rio Red	SS	SS	-	II	FS/FF
	Ruben Seedless	SS	SS	-	II	FS/FF
C. grandis (Shaddock)	Chandler	SS	SS	-	MM	FS/FF
	Melogold	SS	SS	-	MI	FS/FF
	Oroblanco	SS	SS	-	MI	FS/FF
	Tahiti Pummelo	SS	FS	-	II	FS/FF

2.3.2 Restriction fragment length polymorphism

Five grapefruit cultivars (Ray Ruby, Nelruby, Oran Red, Ruben seedless and Rio Red) were digested with three restriction enzymes (*Bam* HI, *Eco*RI and *Hind* III) and hybridized to a DIG-labelled 1.23 kilobase insert of the cDNA clone, pRL94. By combining three restriction enzymes and only one probe, it was possible to distinguish between these cultivars (Figure 2). Oran Red and Rio Red have similar banding patterns when digested with *Bam* HI and hybridized to probe pRL94. However, when the banding patterns produced by *Eco* RI and *Hind* III were analyzed, they were significantly different.

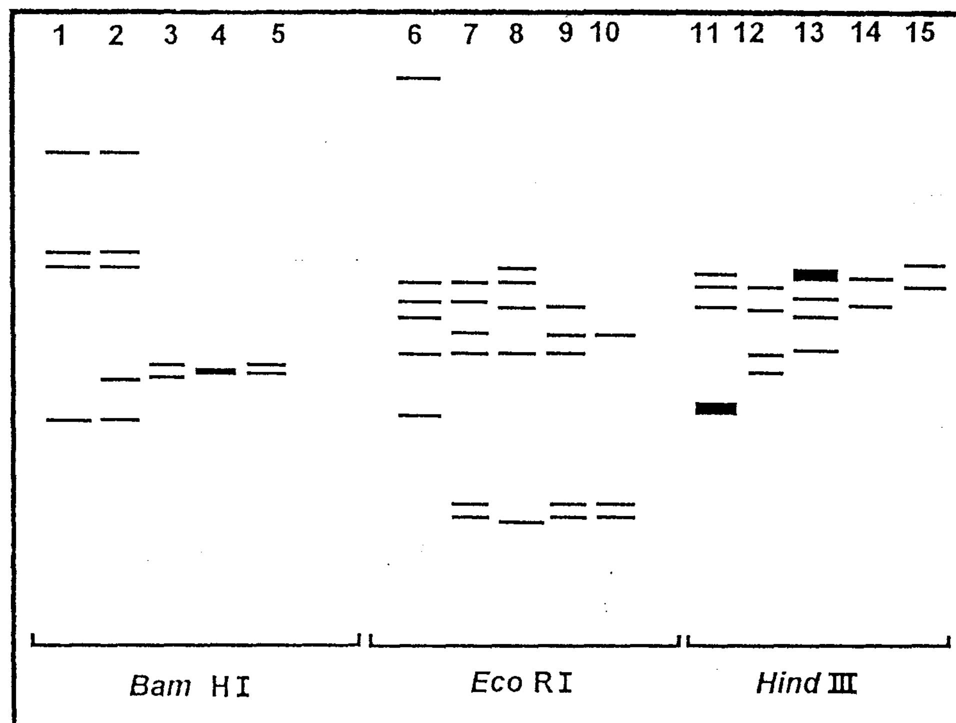


FIGURE 2. RFLP banding patterns of 5 grapefruit hybridised to a DIG-labelled probe (pRL94). DNA in lanes 1-5 was digested with *Bam* HI, lanes 6-10 with *Eco* RI and lanes 11-15 with *Hind* III. Lanes 1, 6, 11 - Ray Ruby; 2, 7, 12 - Nelruby; 3, 8, 13 - Oran Red; 4, 9, 14 - Ruben Seedless and 5, 10, 15 - Rio Red.

2.4 DISCUSSION

Most cultivated citrus types, including sweet orange, grapefruit, lemon and lime, are believed to have originated by natural hybridization (Roose, personal communication), and are therefore expected to be quite heterozygous. It was possible to group all the cultivars into the three species, *C. sinensis, C. paradisi,* and *C. grandis*, with the aid of isozyme analysis. It was not possible to distinguish between Valencia and Navels (both groups belong to the sweet oranges, but are morphological quite distinct).

Isozyme phenotypes of cultivars within a particular group, were the same in most instances. Differences observed in some cultivars were too insignificant to be used for routine cultivar typing. The homozygous nature of the isozyme phenotypes in *C. paradisi* (grapefruit) and *C. grandis* (shaddocks) supports the believe the most of the commercial cultivars in the South African citrus industry originated as from mutated buds. Morphological differences might therefore be the result of somatic mutations and not the introduction of foreign genetic material via natural or artificial hybridization. Isozymes are secondary metabolic products coded for by only one gene and is therefore the phenotypic expression of a very small region of the plant's genome. It is therefore not surprising that the variation within this groups is not always detectable in the isozyme patterns. Bud mutation which causes phenotypic change, do not necessarily alter the expression of the isozymes.

RFLPs were suited for routine cultivar identification in grapefruit. It was possible to identify a cultivar that originated from a natural mutation. Nelruby is believed to be a nucellar seedling of Ray Ruby. There was thus no introduction of foreign genetic material.

However, results show different banding patterns for the probe/enzyme combinations tested. The differences observed between these cultivars, are therefore probably the result of somatic variation caused by chromosomal rearrangements or spontaneous mutations in the DNA sequence.

Melogold and Oroblanco are becoming very important shaddock cultivars in the citrus industry. It was not possible to differentiate between them with isozymes, but RAPDs revealed significant DNA polymorphisms.

CHAPTER 3

ANALYSIS OF THE GENETIC RELATIONSHIPS IN CITRUS FROM RANDOM AMPLIFIED POLYMORPHIC DNA DATA, WITH PARTICULAR EMPHASIS ON CULTIVAR IDENTIFICATION

3.1 INTRODUCTION

Random Amplified Polymorphic DNA (RAPD) is a useful method for DNA fingerprinting to identify varieties of crops species. In addition to varietal identification, phylogenetic relationships can be examined by RAPDs. RAPD polymorphisms are the result of nucleotide base changes that alter the primer binding site, or an insertion or deletion within the amplified region. Polymorphisms are noted as the absence or presence of a band. RAPD markers are capable of detecting polymorphisms in single copy and repetitive DNA.

Since the higher plant genome is very large, several amplified fragments are normally observed. The DNA fragments amplified in the Polymerase Chain Reaction, is separated by agarose or polyacrylamide gel electrophoresis.

RAPD markers are inherited in Mendelian fashion and can be produced from any species without prior DNA sequence information.

TABLE 2. Cultivars included in the Citrus Improvement Programme (CIP)

GENUS	GROUP	SELECTION
C. sinensis	Navel	Amanzi
		Californian Lane Late
		Cara Cara
		Fisher
		Gillemberg
		Lane Late
		Leng
		Navelina
		Painter Early
		Patensie
		Prins Navel
		Robyn
		Royal Late
		Tuligold
		Washington
	Valencia	Broedershoek
		Delport
		Margaret
		Ruben Seedless

TABLE 2. (cont.)

GENUS	GROUP	SELECTION
C. paradisi	Grapefruit	Flame
		Henderson
		Nelruby
		Oran Red
		Ray Ruby
		Rio Red
		Ruben Seedless
C. grandis	Shaddock	Chandler
		Melogold
		Oroblanco
		Pomelit
		Tahiti pummelo
C. reticulata	Mandarin	Daisy
		Fortuna
		Kiyomi Satsuma
		Lomati Tangelo
		Minneola
		Robin
		Sweet Spring
		Temple Thoro

TABLE 3. Cultivars for which applications for plant breeders' rights (PBR) were filed on 10 March 1993.

GENUS	GROUP	CULTIVAR	A/S
C. paradisi	Grapefruit	Marsh	S
		Nelruby	A
		Ray Ruby	S
		Redblush	S
		Star Ruby	S
C. reticulata	Mandarin	Edelgard	A
		Ellendale	S
		Elno	A
		Lanique	A
		Lee	S
		Nova	S
		Novel	A
		Novelty	S
		Orlando	S
		Ortanique	S
		Page	S
		Robin	A
		Robinson	S
		Roma	
		Ronel	A

TABLE 3. (cont.)

GENUS	GROUP	CULTIVAR	A/S
C. sinensis	Navel	Gillemberg	A
		Mias	A
		Nieuwoudt	A*
		Palmer	S
		Rautenbach	A
		Robyn	A*
		Royal Late	A
		Washington	S
C. grandis	Shaddock	Chandler	S
		Pomardine	A
		Pomelia	A
		Pomelit	A

* Approved by UPOV for plant breeders' rights
A Application received for PBR
S Cultivars used as standards

3.2 AMPLIFICATION CONDITIONS AND OPTIMISATION

Weeden *et al.* (1992), emphasised the importance of high quality DNA for the generation and resolution of amplified products. It was found that crude DNA, containing moderate levels of impurities, were sufficient to obtain the required results when RAPDs were analyzed. However, for RFLPs, ultra-pure DNA was necessary for the fractionation of the DNA with the restriction enzymes. The best RFLP blots were obtained when the plant DNA was purified through CsCl/Ethidium bromide gradient ultra-centrifugation, but this was too expensive for the routine screening of multiple samples.

According to Devos and Gale (1992), the template DNA only gave reproducible results in the concentration range between 10 and 30 ng when the primer concentration was kept constant at 200nM: too little template DNA resulted in non-specific amplification. However, it was not necessary to quantify the DNA with a fluorometer for every sample, as any amount between 3 and 300 ng gave acceptable amplification products during the PCR reaction. A constant primer concentration of 7 to 10 picamoles per 50 μl reaction was used throughout the experiments. It was found that a Taq polymerase concentration of 0.75 U per 50 μl reaction gave consistent satisfying results.

Although the annealing temperature did not seem to have any significant influence in the products produced during the PCR, it was found that when the annealing time was increased from 10 seconds to 30 seconds per cycle, the products obtained were more consistent. The lower the denaturing temperature, the better the average results obtained. When a denaturing temperature of 91 °C instead of 94 °C was implemented, the amplification of all the DNA analyzed, increased. However, Devos and Gale (1992) found that results obtained with different thermocylers, were not consistent. These variations suggest that the cycling conditions for every individual machine should be optimized. Amplification products from different thermocyclers should therefore not be compared.

The effect of Mg^{2+} on the PCR was investigated by varying the concentration of $MgCl_2$ in the standard PCR between 1 and 4.5 mM. Valencia Turkey DNA was used as tem-plate DNA and the experiment was repeated. Specific and reproducible results were obtained at 1.5 mM Mg^{2+}. Devos and Gale (1992) found that low template concentra-tions and

high Mg^{2+} resulted in the generation of some novel bands and the disappearance of others (Figure 3). According to Welsh and McClelland (1990) a high Mg^{2+} (> 4 mM) concentration enhanced the stability of primer/template interactions. This suggests that the stringency of the annealing process is influenced by the concentration of magnesium ions in the reaction vial.

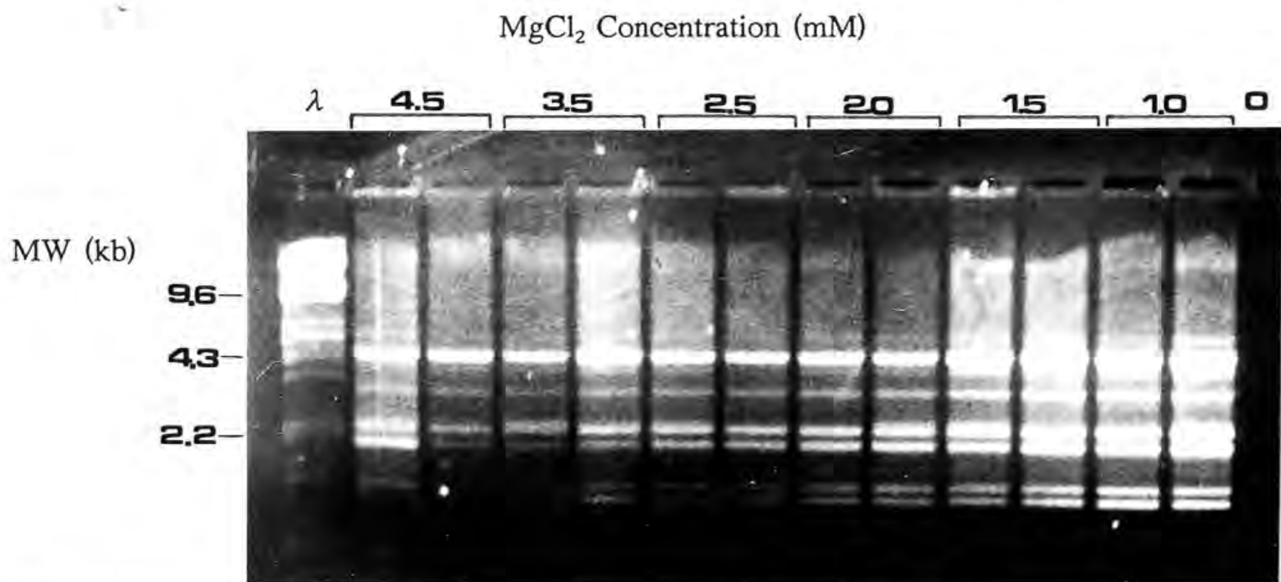


FIGURE 3. The effect of $MgCl_2$ concentration on the amplification of Valencia Turkey DNA during the Polymerase Chain Reaction. DNA fragments were separated on a 1.4% agarose gel at a constant 80 volts for 2 hours. Samples were loaded in duplicate.

Taq DNA polymerase has no 3' to 5' exonuclease ("proofreading") activity, but has a 5' to 3' exonuclease activity during polymerization. The initial estimates of the misincorporation rate by Taq DNA polymerase during PCR is about 10^{-4} nucleotides per cycle. According to Erlich *et al.* (1991), lower dNTP and $MgCl_2$ concentrations, higher annealing temperatures and shorter extension times will reduce the misincorporation rate to less than 10^{-5} nucleotides per cycle. The probability that primer extension will not be

completed, is dependent on the distance between two primers, quality of the template DNA, enzyme limitation and extension time (Erlich *et al.*, 1991).

The primers used for generation of fingerprints for individual cultivars and selections, are listed in table 4. Fourteen primers were screened of which 7 yielded data that revealed polymorphisms. Not all the primers performed equally well. In general, the size of amplified DNA fragments generated by most of the primers used, ranged from 300 to 2 500 base pairs. The number of bands in the profiles varied between three and twenty, depending on the primer and cultivar tested. Some primers did not produce any bands and none of the primers consistently yielded small or large bands. Primer #3 initiated the generation of bands in almost all the cultivars, but when the PCR products were analyzed on a 1.4% agarose/TBE gel, it appeared smear-like. Of all the primers tested, primer #2 and primer OPA-04 gave the best average results in most of the cultivars tested. Primer OPA-04 was able to produce individual fingerprints for all the mandarins awaiting plant breeders' rights as well as for the cultivars used as standards (Figure 4).

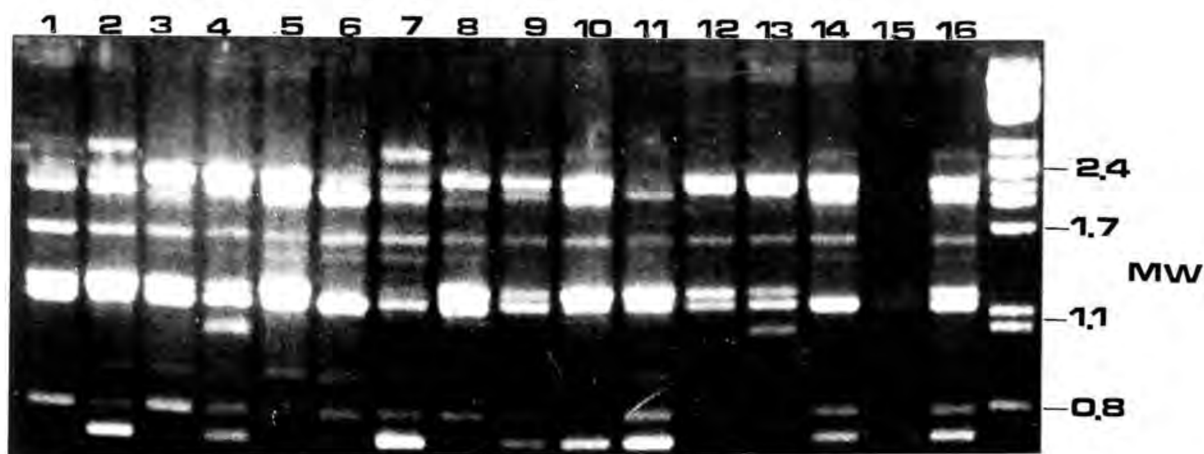


FIGURE 4. Ethidium bromide-stained electrophoretic pattern of mandarin DNA amplified with primer OPA-04. Lanes 1-16 contain, in the regular order, Ellendale Beauty, Lanique, Ortanique I, Page, Novel, Robin, Ronel, Elno, Nova, Roma, Edelgard, Ortanique II, Novelty, Lee, Orlando and Robinson.

TABLE 4. Primers used to initiate DNA synthesis from various loci in the citrus genome.

PRIMER	SEQUENCE	AVERAGE NUMBER OF BANDS PRODUCED
#288	5'GCAAGTAGCT3'	10 - 12
#2	5'GCAAGTAGCT3'	8-9
#3	5'CGGCCCCTGT3'	6-10
#8	5'TCTCGATGCA3'	5-8
#10	5'TGGTCACTGA3'	5-6
OPA-01	5'CAGGCCCTT3'	2-4
OPA-02	5'TGCCGAGCTG3'	9
OPA-03	5'AGTCAGCCAC3'	8 (?)
OPA-04	5'AATCGGGCTG3'	6
OPA-05	5'AGGGGTCTTG3'	6
OPA-06	5'GGTCCCTGAC3'	-
OPA-07	5'GAAACGGGTG3'	7-10
OPA-08	5'GTGACGTAGG3'	8-12
OPA-09	5'GGGTAACGCC3'	2-5

3.2 ANALYSIS OF RELATIONSHIPS BETWEEN CITRUS CULTIVARS

Genetic distance as detected by RAPD marker analysis is the result of a series of comparisons, each of which can have two possible outcomes - similarity or difference. Genetic distance can be defined as the ratio of differences to the total number of comparisons. The comparison between genotypes were based on the presence ("1") or absence ("0") of a RAPD marker for each genotype under investigation. If each outcome is given the value of 1 for difference and 0 for similarity, then genetic distance is equal to the numerical mean of this set of observations. The genetic distance for each genotype pair is thus redefined as the mean of the observations of similarity or difference over all marker loci tested. Genetic distance between two genotypes can be calculated from RAPD data by the application of the following simple formula:

	Genotype		
	A	B	|A - B|
1	1	0	1
2	0	1	1
3	1	1	0
4	0	0	0
5	0	0	0
6	1	0	1
7	1	1	0
8	1	1	0
9	0	1	1
10	1	0	1

$$\text{Genetic distance (A,B)} = \frac{\Sigma_i |A_i - B_i|}{10} = 0.50$$

FIGURE 5. Comparison of two hypothetical genotypes for 10 RAPD markers (Skroch *et al.*, 1992).

The variance of genetic distance, can be calculated using the formula for sample variance. For a genetic distance, d, based on n RAPD bands, the variance and standard error are given by the following two formulas:

$$\text{variance (d)} = nd(1-d)/(n-1)$$

$$\text{estimated standard error} = [var(d)/n]^{1/2}$$

A data matrix was constructed for every set of data generated by a particular primer. Combined data matrices (containing data from various primers) were constructed for each of the Grapefruits, Shaddocks and Mandarins (Appendix D). Data were analysed by three clustering programs, NJOIN and SAHN, of NTSYS-pc (Numerical Taxonomy and Multivariate Analysis System) (Exeter Software, New York) and MIX (which produces unrooted consensus trees).

NJOIN (Neighbour-Joining method) is based on the idea of parsimony and generates estimated phylogenetic trees, whilst SAHN-clustering (Sequential, Agglomerative, Hierarchical and Nested clustering methods) uses UPGMA (Unweighted pair-group method, arithmetic average). The neighbour-joining method was developed as a method for estimating phylogenetic trees. It attempts to find a tree close to the true phylogenetic tree, rather than attempting to find the shortest possible tree for a set of data. The algorithm used to find NJOIN trees is similar to that of the distance Wagner procedure. In the NJOIN method, trees are constructed by linking together the two original taxonomic units (OTU) that are the closest mutual "neighbours".

PENNY is another program that will find all of the most parsimonious trees implied by the data by using the more sophisticated "branch and bound" algorithm. The search strategy used by PENNY starts by making a tree consisting of the first two species (the first three if the tree is to be unrooted). Then it tries to add the next species in all the possible places. This process is continued for all the data points. It will thus generate all the possible trees, of which there are a large number, even when the number of species is moderate. Although this programme was evaluated for analysis of clustering data, too many different trees were found. This program cannot handle the comparison of many samples. It was therefore decided not to pursue the use of this program any further.

MIX is a general parsimony program which carries out the Wagner an Camin-Sokal parsimony methods in mixture, where each character can have its method specified separately. The program defaults to carrying out Wagner parsimony. The criterion is to find the tree which requires the minimum number of changes.

3.2.1 Sweet oranges (*Citrus sinensis* [L.] Osbeck)

The most distinctive feature of the navel oranges, in which they differ from all others, is anatomical in nature and consists of the presence of the navel - a small secondary fruit embedded in the apex of the primary fruit. Seedlessness is also a characteristic of navel oranges, resulting from the absence of functional pollen and viable ovules. Navel orange clones in general and notably the Washington variety, is unstable and give rise to somatic mutations. Nearly all known navel cultivars have originated from Washington, as bud mutations. It is almost impossible to differentiate between navels on a morphological or physiological level. The difficulty of determining variation among navel cultivars is

surprising. The small amount of variation detectable so far may also be ascribed to the small amount of primers screened. Most navel cultivars were selected on the maturity date of the fruit. Although this characteristic indicates differences within the group, it does not necessarily have a multigenetic and broad genetic basis. It was impossible to distinguish between the selections with RADPs (Fig. 6).

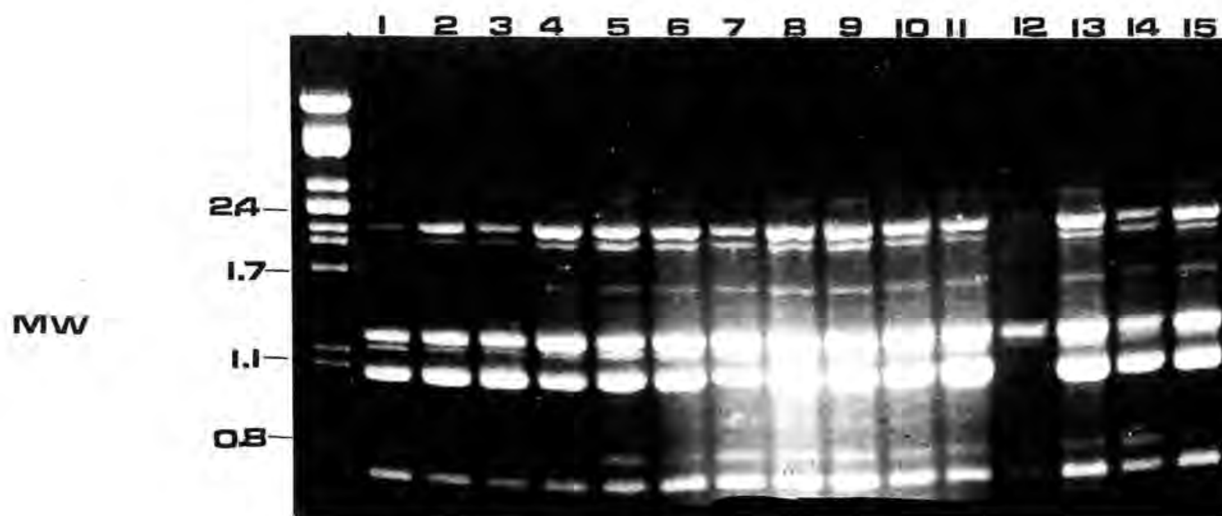


FIGURE 6. RAPD products produced by amplification of 15 Navel orange cultivars with primer OPA-04. Fragments were separated on 1.4% agarose gels and stained with ethidium bromide. No differences were observed between the RAPD profiles of the Navels. Lanes 1 to 15 contain Amanzi, Californian Lane Late, Cara Cara, Fisher, Gillemberg, Lane Late, Leng, Navelina, Painter Early, Patensie, Prins Navel, Robyn, Royal Late, Tuligold and Washington.

3.2.2 Grapefruit (*C. paradisi)*

There is no absolute certainty about the origin of the grapefruit. Morphological characteristics support the theory that grapefruit is the result of a cross between a sweet orange (*C. sinensis*) and a shaddock (*C. grandis*) or between *C. grandis* and *C. reticulata*. Another theory suggests that the grapefruit could possibly have resulted from a bud mutation

of the shaddock. Most modern grapefruit cultivars and selections originated through bud mutations or from superior nucellar seedlings - implicating somatic variation. RAPDs were sensitive enough to detect these small, but significant changes in the genome of the grapefruit.

It was possible to identify every grapefruit cultivar from trees drawn from data generated by single primers and from the combined data, analysed with NJOIN (Fig. 7). This data compared favourably with clustering results obtained from SAHN-UPGMA (Fig. 8) and MIX (Fig. 9). However, SAHN-UPGMA analysis was unable to separate out all the cultivars, even when the data from separate matrices were pooled to create a combined data tree.

Nelruby (a red-fleshed cultivar with a red-pigmented skin) clustered with Marsh (a thick-skinned, yellow-fleshed cultivar) in all the data trees. Nelruby originated as a nucellar seedling from Ray Ruby (also a red cultivar) - Marsh could therefore not have been a direct ancestor of Nelruby. Specific evolutionary and breeding data on Marsh is not available. An explanation for this somewhat peculiar phenomenon is therefore not possible. Nelruby and Ray Ruby do separate at close distance (at a relative genetic distance of 0.25 on the UPGMA scale).

Although it was not possible to distinguish between Ray Ruby, Rio Red, Henderson and Flame with UPGMA, the cultivars separated out as individuals on the NJOIN tree, but in one cluster, with Rio Red branching at ca. 4 on this scale.

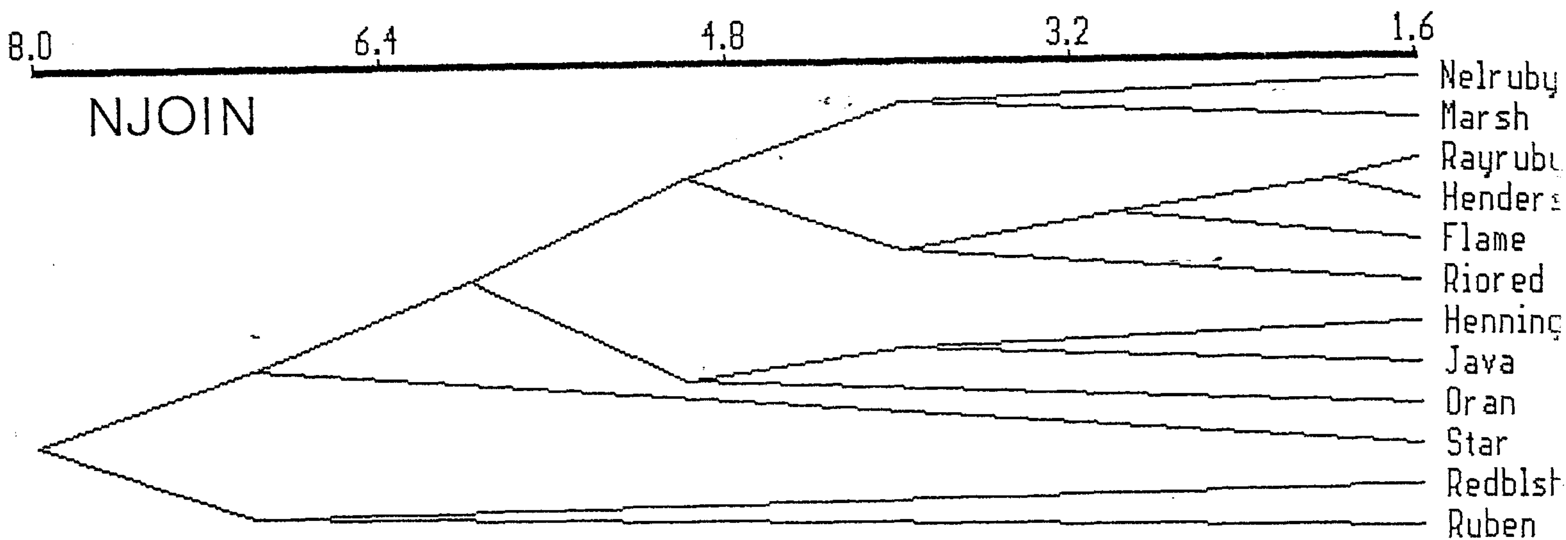


FIGURE 7. A phylogenetic tree of 12 grapefruit (*C.paradisi*) cultivars produced from Neighbour-Joining analysis (NJOIN). Combined datasets were generated by three 10-mer primers (#288, #10 and OPA-04).

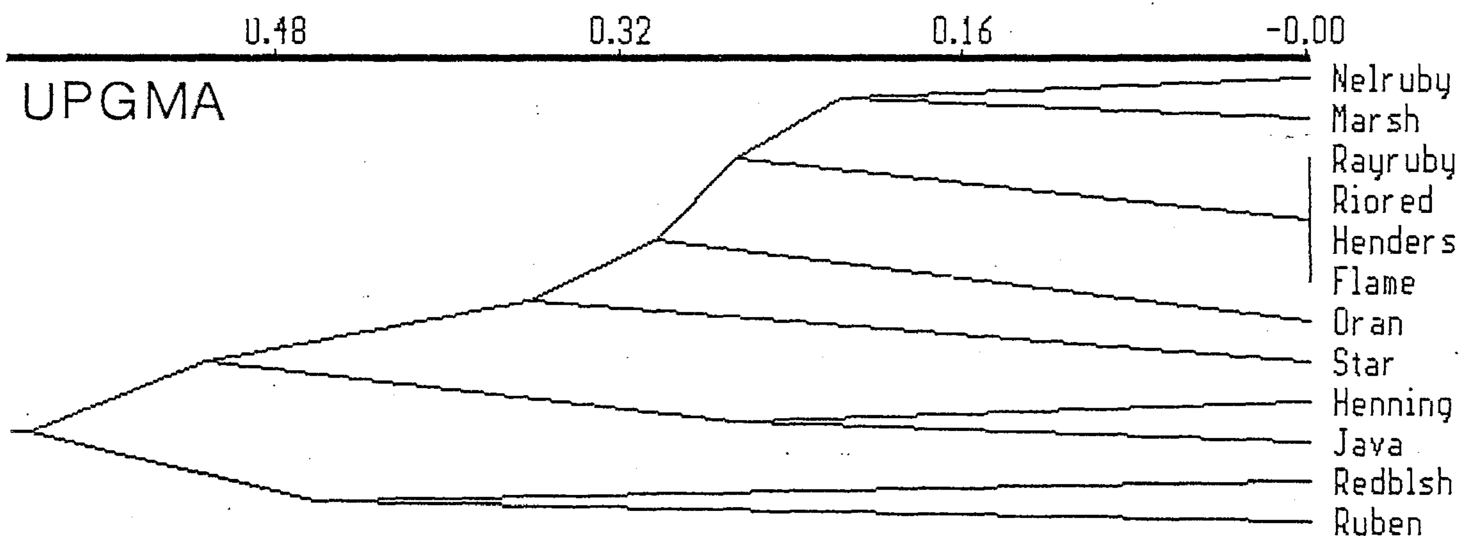


FIGURE 8. A phylogenetic tree of 12 grapefruit (*C.paradisi*) cultivars produced from SAHN-UPGMA cluster analysis. Combined datasets were generated by three 10-mer primers (#288, #10 and OPA-04).

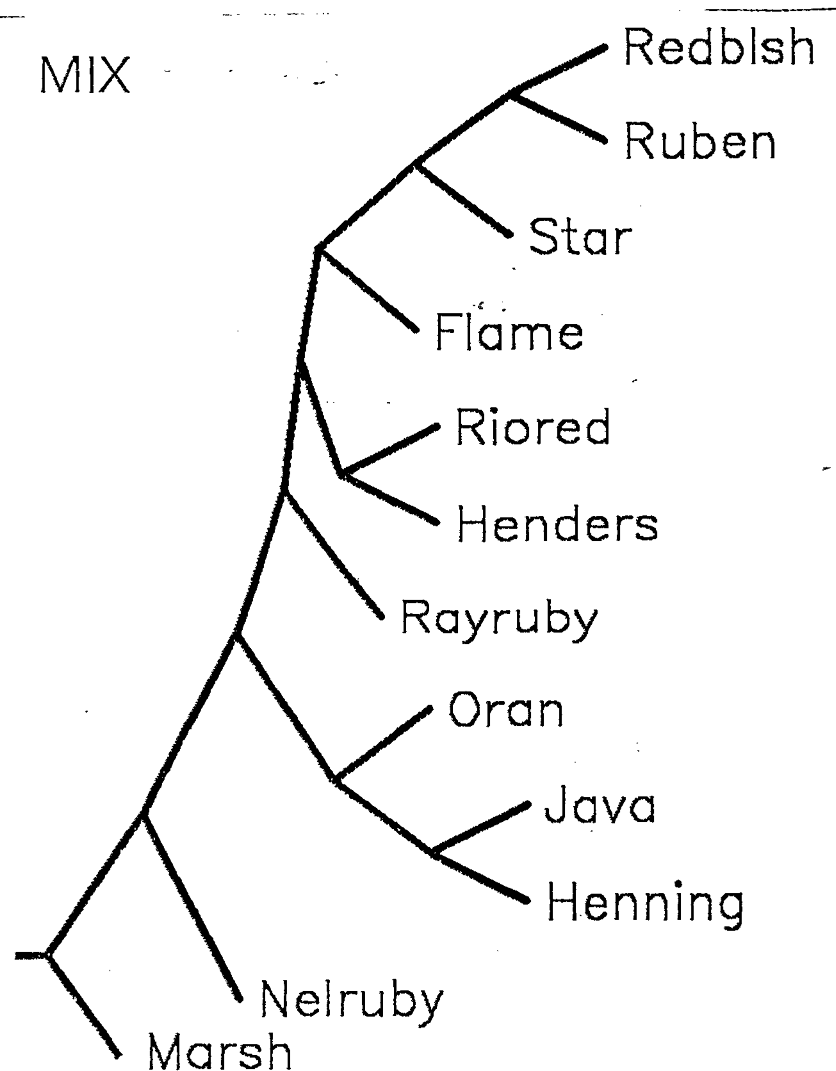


FIGURE 9. A phylogenetic tree of 12 grapefruit (*C.paradisi*) cultivars produced from MIX cluster analysis. Combined datasets were generated by three 10-mer primers (#288, #10 and OPA-04).

3.2.3 Shaddock (*C. grandis*)

Shaddocks, also known as pummelos, exhibit a range of variation in characteristics. The range of variation in fruit size, form and rind thickness is notable. Another reason for these important differences reside principally in the fact that the pummelos are monoembryonic and hence each seedling constitutes a different genotype. It is therefore expected that less cultivars in this genus originated from somatic mutation than in the related *C. paradisi* (grapefruit). This fact is depicted by the distinctness observed between the cultivars analyzed from data generated by single primers and the combined data set (Fig. 10).

Although NJOIN was the preferred clustering programme to separate the grapefruit and mandarins, it seemed that SAHN-UPGMA was more suited for separating shaddock cultivars (Fig. 11). Pomelit and Chandler clustered together in the trees generated by both programs. Genetic distance between these cultivars ranged from 0 to 0.75 in individual SAHN-UPGMA trees and averaged 0.54 in the combined tree. Again, breeding or evolutionary data is limited and can therefore not explain this observation.

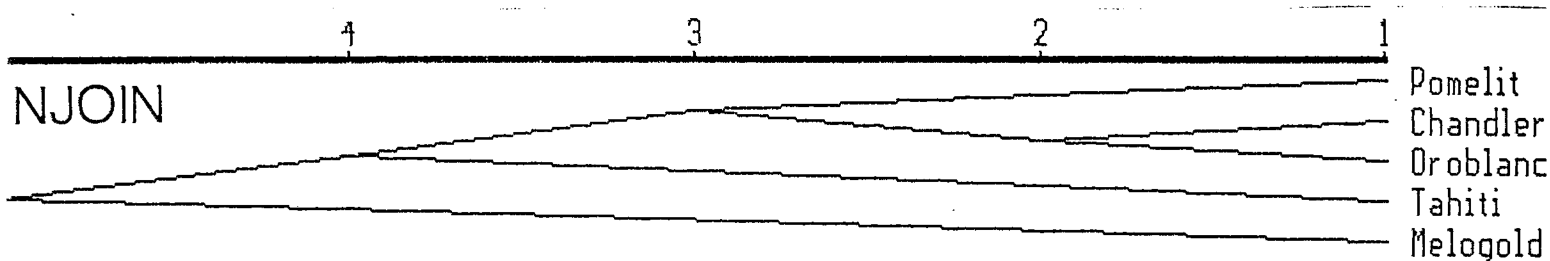


FIGURE 10. A phylogenetic tree of 4 shaddock (*C. grandis*) cultivars produced from NJOIN cluster analysis. Combined datasets were generated by four 10-mer primers (#288, #3, OPA-04 and OPA-07).

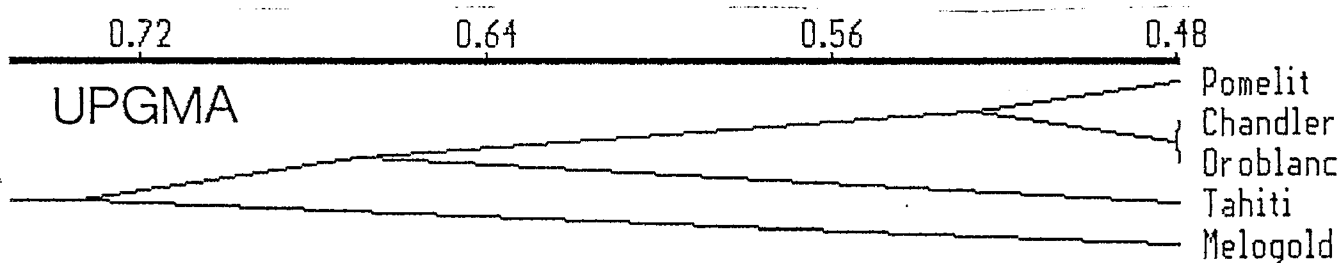


FIGURE 11. A phylogenetic tree of 4 shaddock (*C. grandis*) cultivars produced from SAHN-UPGMA cluster analysis. Combined datasets were generated by four 10-mer primers (#288, #3, OPA-04 and OPA-07).

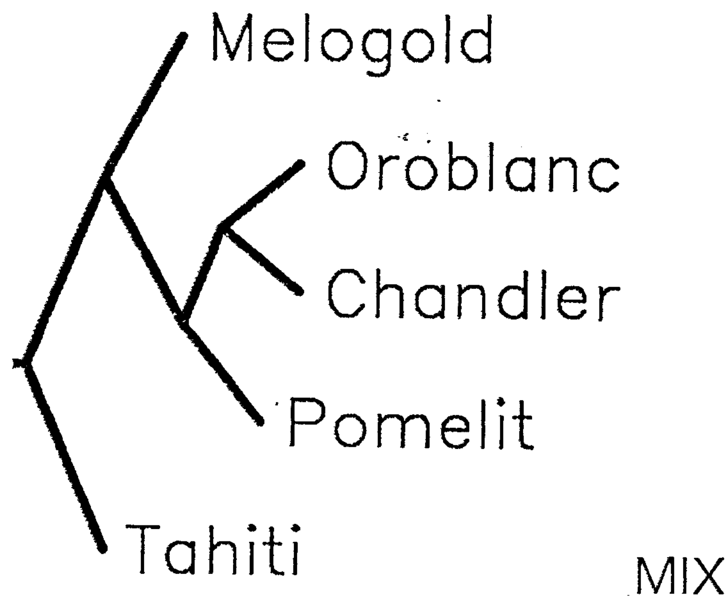


FIGURE 12. A phylogenetic tree of 4 shaddock (*C. grandis*) cultivars produced from MIX cluster analysis. Combined datasets were generated by four 10-mer primers (#288, #3, OPA-04 and OPA-07).

3.2.4 Mandarin (*C. reticulata* Blanco)

C. reticulata Blanco is usually referred to as the common mandarins. Mandarins consists of many hybrids which make it a very heterogenous group. Most mandarins originate from natural hybridization, controlled pollination or, in some instances, mutations. The heterogenous nature of this genus is therefore expected. As new cultivars are getting more closely related because of the selection of superior seedlings, rather than the introduction of foreign genetic material, the process of distinguishing between them, gets more difficult. The range of variation in characters exhibited by the mandarin group, is much greater than in the navel oranges, pummelos and grapefruit. Although obviously closely related, the mandarins are separable into several natural groups like Tangelo, Satsuma, Clementine and Tangor. Results confirm this fact in that most of the cultivars had unique DNA profiles with

single primers. When analyzed using the tree generating programmes on both the single primers and the combined data, good separating of cultivars was obtained (Fig. 13).

Lanique and Ronel separated into one cluster on the NJOIN tree (Fig. 13) at a distance of 1.8, and into two sub-clusters, still very close, on the UPGMA scale (Fig. 14). Lanique and Ronel are both awaiting the awarding of PBR.

UPGMA did not distinguish between Ellendale, Roma and Robinson, and between Elno and Ortanique-L13. All of these cultivars separated out in the NJOIN and MIX tree (Fig. 15), with Ellendale, Roma and Robinson in one cluster and Elno and Ortanique in another cluster. Two selections of Ortanique, L3 and L13 were included in this study to determine their genetic relationship. These two selections could be distinguished with all the primers used. It is therefore suspected that these two selections evolved in two separate directions, probably by the introduction of foreign pollen via open pollinated flowers.

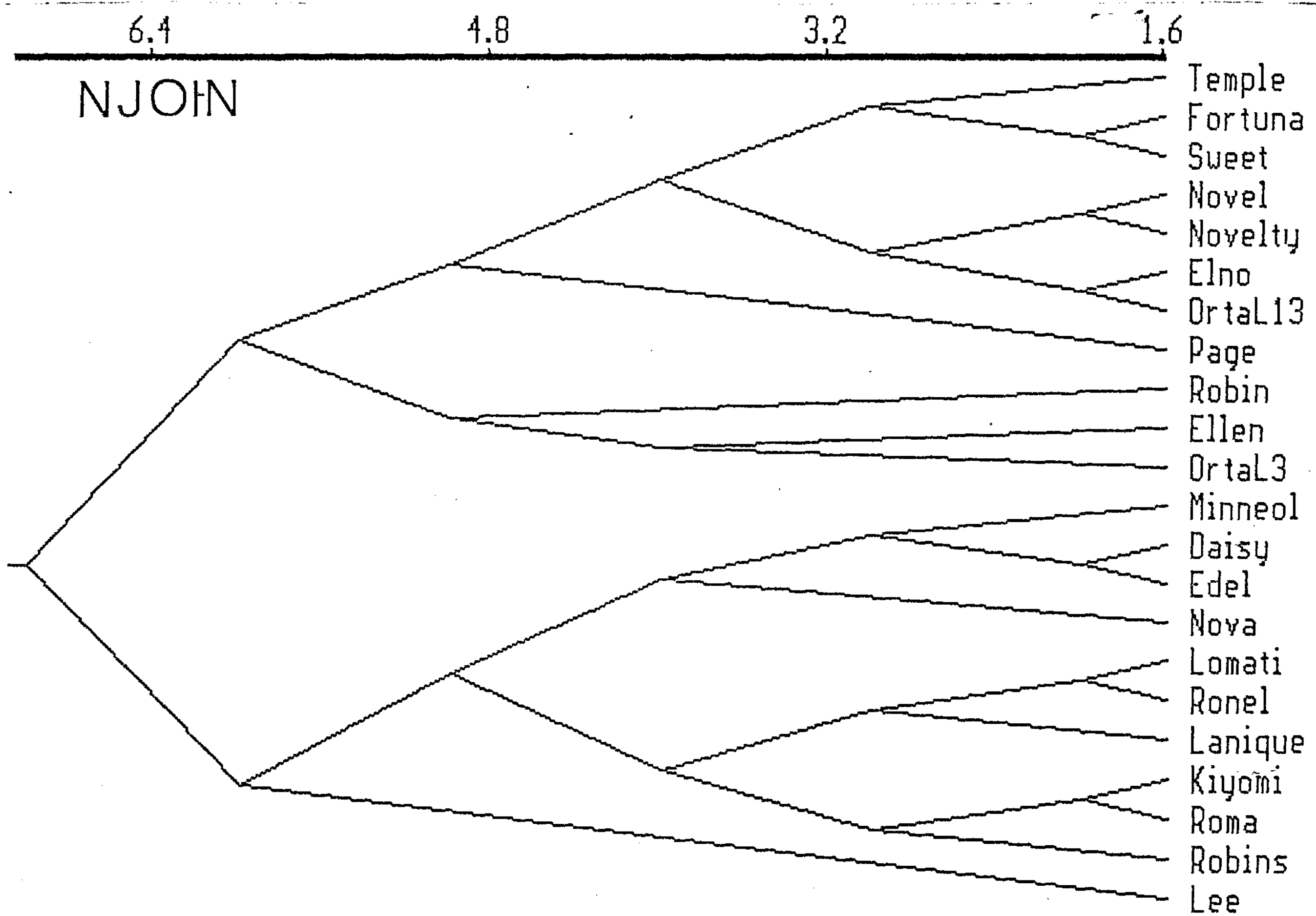


FIGURE 13. A phylogenetic tree of 22 mandarin (*C. reticulata*) cultivars produced from NJOIN cluster analysis. Combined datasets were generated by three 10-mer primers (OPA-04, OPA-07 and OPA-09).

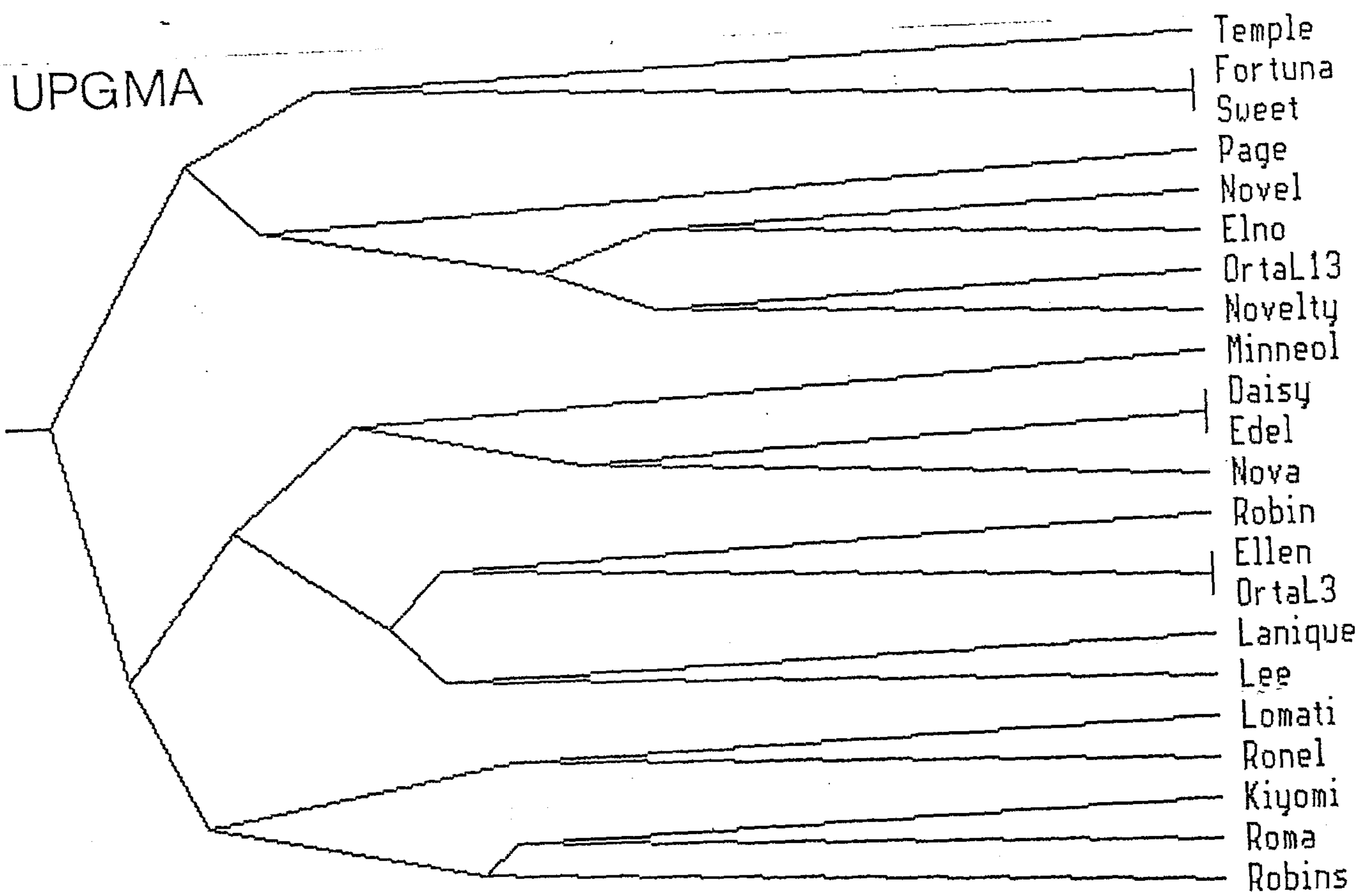


FIGURE 14. A phylogenetic tree of 22 mandarin (*C. reticulata*) cultivars produced from SAHN-UPGMA cluster analysis. Combined datasets were generated by three 10-mer primers (OPA-04, OPA-07 and OPA-09).

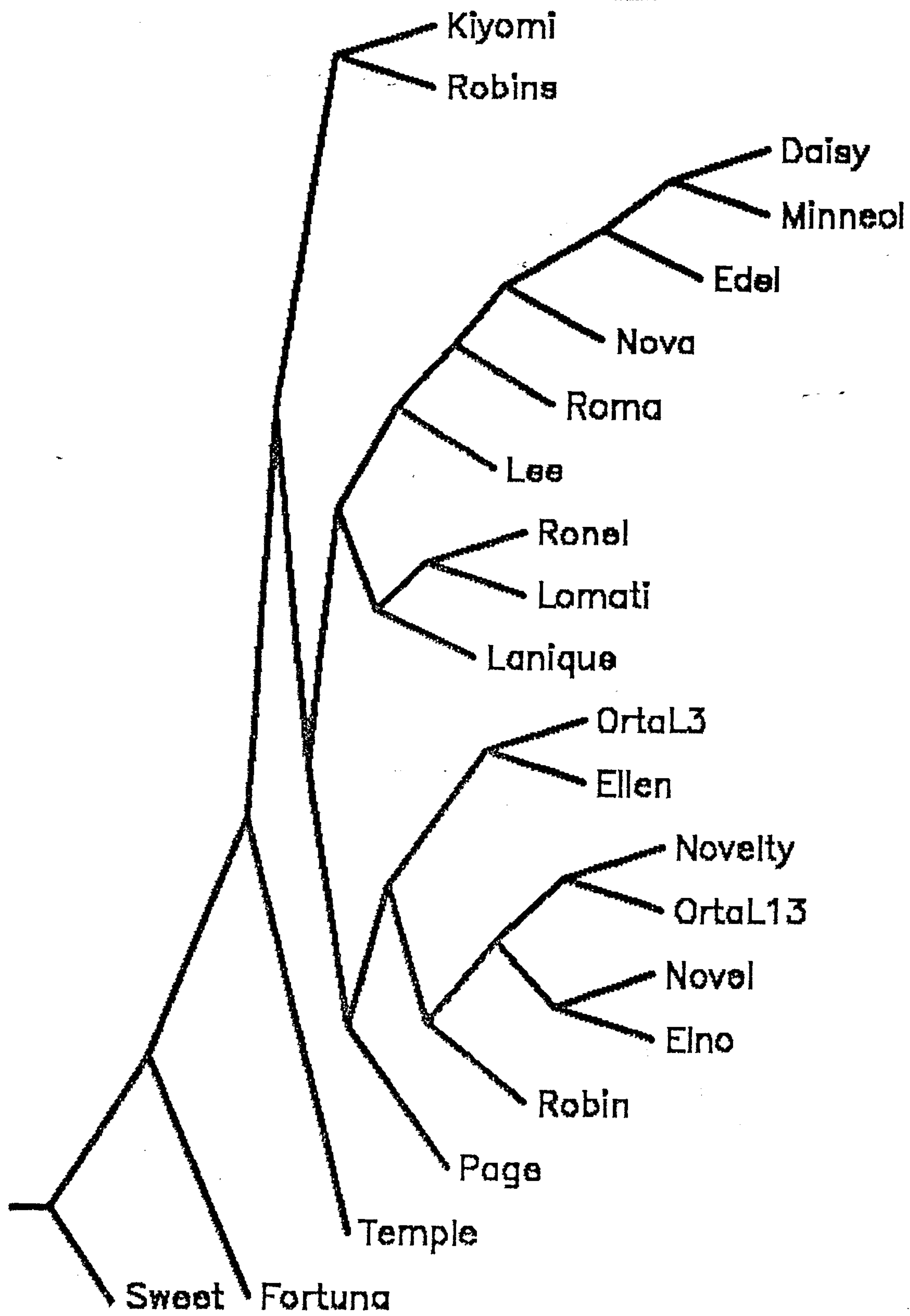


FIGURE 15. A phylogenetic tree of 22 mandarin (*C. reticulata*) cultivars produced from MIX cluster analysis. Combined datasets were generated by three 10-mer primers (OPA-04, OPA-07 and OPA-09).

3.3 PRACTICAL SOLUTIONS TO THE DIFFICULTIES IN THE PROTECTION OF NEW VARIETIES.

New candidate varieties are compared with known varieties of the same species, and as far as possible, the comparison are made at a single locality under uniform conditions. Although Plant Breeders' Rights may be granted after a single season of testing, the description of a variety is usually based on at least two season's data. Initially, very few characteristics were required to establish distinctness, as varietal differences could be observed easily. Observations were mainly based on visual assessments and were therefore very subjective. These unsophisticated methods are now insufficient when trying to distinguish between varieties which closely resemble each other. The difficulties experienced in identifying similar varieties often result in the testing period being extended for another year (Keetch, 1992).

Nelruby is a red-fleshed, seedless grapefruit, selected form a nucellar seedling of Ray Ruby by the ITSC and is currently being evaluated for the awarding of Plant Breeders' Rights. When an application is filed for Plant Breeders' Rights, the material is budded onto a rootstock and planted with reference cultivars. It usually takes more than three years for the tree to start producing fruit and then another two to three years for evaluation. The time delay from filing the application to the decision is thus more than five years. Nelruby is compared to Ray Ruby, Star Ruby (a red grapefruit, imported), Redblush (a pink grapefruit) and Marsh (white grapefruit) in order to establish significant distinctness, uniformity and stability. Results obtained from RAPDs, showed that Nelruby is genetically different from the reference cultivars when screened with only one primer, OPA-04.

The same difficulties are experienced with the mandarins as with the other selections. Again, it was possible to differentiate between all these selections with only one primer. The mandarins were mostly selected for specific characteristics; however, it is not UPOV's task to set guidelines for fruit quality and this characteristic is therefore not taken into account when the application for PBR is considered. The breeder selected the cultivar for a very specific trait and it is therefore assumed that it is different from any other known cultivar.

The application of DNA techniques to determine distinctness among varieties will therefore shorten the time needed to award PBR.

In navels, the various selections, differ mainly in their maturity date. It is important to select or breed cultivars which can extend the harvesting season. Variability in morphological and physiological characteristics are limited in navels. Six navel selections are currently being evaluated for PBR. None of these can be differentiated from the reference cultivars. It was not yet possible to distinguish these cultivars with RAPDs. Although there were a few insignificant differences in banding patterns in some of the selections, it could not be reproduced in repeated PCR's. It is therefore necessary to implement DNA techniques to differentiate between the selections.

3.4 DISCUSSION

According to Williams *et al.* (1992), the presence of a RAPD marker in both genotypes indicates a high level of sequence homology at this particular site. This is not necessarily the case; it is possible that similar sized fragments can be created from different priming sites, thus having different nucleotide sequences. DNA fragments (RAPD markers), are separated according to differences in molecular weight (=size of the fragment), in an electric field. It is thus possible that bands with different nucleotide sequences can co-migrate in an electric field, letting the interpreter to believe that the bands are the same. In the case where one genotype has the marker while the other lacks it, there is certainty of sequence difference. To determinate distinctness between genotypes, it is therefore necessary to concentrate on the absence rather than the presence of the bands. Increasing the number of primers to determine distinctness, will overcome this problem.

Most of the detected RAPD markers were common within a particular group investigated. This indicates that most of the cultivars and selections in a species are very closely related. Although species mainly originated from hybridizations between species, most of the commercial cultivars were selected from bud mutations or exceptional tree performance, like increased yields or specific desired fruit characteristics and harvest time. Cultivar status were thus not necessarily awarded on the bases of pure genetic diversity.

In a number of cases, especially in the sweet oranges, variation between cultivars could not be detected. Due to the random effect of the primers used, certain regions of the genome will inevitably be overamplified and underamplified; in other words, one primer may reveal

no or very few polymorphisms between the genomes, because it might amplify a fairly conserved region of the genome, while another primer might amplify a hypervariable region. Because of the random selection of the primers, the number of primers and the nucleotide sequence thereof obviously play an important role in the outcome of the results. The more primers used to amplify the genome, the more uniform the distribution of markers across the genome.

The more primers used to construct a matrix of similarity of the cultivars under investigation, the larger the part of the genome that can be analysed and the more appreciable the phylogenetic tree in the end. This will increase the resolution of RAPD analysis. A relatively accurate assessment of the genetic difference between genotypes can therefore be achieved.

RAPD phenotypes for many plant species typically display more than one amplified fragment. In order to observe the discriminative value of this technique, a large number of different cultivars were analyzed. The profiles of the amplified products from each cultivar were compared to each other for identification of cultivar specific markers. The number and sizes of the bands varied between cultivars, but variation was stable across plants within cultivars for replicate DNA samples obtained from the same plant or from vegetatively propagated material from different locations. All bands from multiple amplifications had to be reproducible before it was taken into account when scoring similarity between cultivars or individuals. It was found that under the described conditions, individual RAPD phenotypes were reproducible for individual samples. Differences in the intensity of some corresponding bands were observed. It is not uncommon for some RAPD products to be

intense, well-resolved bands that are very reliable genetic markers. On the other hand are faint products that are difficult to score, but which may also have a genetic component. Weeden *et al.* (1992) suggest that only the "clear" polymorphisms be scored. However, the definition of "clarity" is somewhat arbitrary. In a small number of cases, the intenser bands might be due to the co-migration of non-homologous bands which are the same size. Apart form DNA hybridization or DNA sequence analysis, it is very difficult to assess whether the amplified fragments in individuals of undefined relationship are homologous. It was decided not to take the intensity of the bands into account when scoring the profiles for polymorphisms, until more information on the scientific value of intensity becomes available.

Although too few primers were screened to measure true phylogenetic relationships, it was possible to separate all the cultivars with NJOIN and Mix. MIX produced much the same data as NJOIN and UPGMA. Although all the cultivars could be distinguished by MIX, too many weak branching points between clusters make the trees obtained by this program somewhat suspected. UPGMA gave an indication of the relatedness of the cultivars and the results corresponded well with that of NJOIN, but it was not suitable to separate out all the individuals. These results prove valuable for consideration as an alternative for the identification of new genotypes when morphological differences in characteristics are too limited to determine sufficient distinctness of new genotypes. NJOIN would therefore be the programme of choice for cultivar identification.

CHAPTER 4

THE ROLE OF MOLECULAR MARKERS IN THE PROTECTION OF NEW VARIETIES OF PLANTS

New, improved varieties are constantly produced to satisfy man's needs or wishes. This is known as "plant breeding" and is always directed towards a practical aim. The variety the breeder strives for must have one or more characteristics which cannot be found in existing varieties of that species of plants - it must be new. This notion of novelty needs to be defined if it is to be used as a basis for legal protection. The protection of new varieties is very important to the breeder. New varieties can be protected by a set of regulations commonly referred to as plant breeder's rights (PBR). It can be compared to patenting of new inventions. Plant breeding can be encouraged if the breeder is granted such legal protection. Although plants cannot be patented in South Africa, new varieties will be considered for the protection of plant breeder's rights, provided that they are genetically distinguishable and significantly different economically from existing varieties and can be maintained in that form.

DNA analysis for forensic and paternity testing, is achieving widespread use, and has become a very effective form of evidence in court. To date, however, there have been no court cases decided where plant DNA fingerprint analysis has been used as evidence. Over the last ten years, there have been rapid advancements in the application of biotechnology to the identification and fingerprinting of plants.

Inspite of the importance if this issue, the only guide-lines currently employed to differentiate between new and existing varieties, in order to decide whether or not the "new" variety qualify for protection, is strictly based on morphological traits. This may be sufficient for ornamentals, which is mainly cultivated for its aesthetic value. However, in food crops, hidden traits such as tolerance to diseases and increased yield potential are more important to the breeder. Yet, these traits are currently not taken into consideration when an application for plant breeders' rights is filed.

The criteria laid down for varietal protection, must exhibit certain minimum requirements. It must have a high power of discrimination. It should also exhibit minimal interaction with the environment. Morphological and agronomical performance data are constrained in this respect. The tests must be objective and not influenced by human judgement errors. The data must be repeatable and accurate and sufficient for the making of valid comparisons among varieties. Data and methodologies used in one laboratory must be recorded into an information database and made available to other laboratories. DNA profile data can easily be stored as chromatographs (peak position and area), fragment size tables or gel pictures. Electronically stored data can be made available by electronic media for comparisons of varietal profiles in the database. These criteria must be able to show relative distances among genotypes that correlate with those from other multigenically based descriptors (Smith and Chin, 1992).

Biochemical techniques such as the analysis of isoenzymes, restriction fragment length polymorphism and random amplified polymorphic DNA can be used to reveal additional

genetic markers. Even with these techniques, the overall documented genetic diversity between cultivars may still be insufficient for specific varietal identification.

Beckmann and Bar-Joseph (1986) suggested that recombinant DNA technology be use to insert tagging identifiers or "signatures" into the genomes of novel varieties. In this procedure, the specific tagging sequence introduced into the genome is a synthetic sequence connected in such a way that it can be recognized by simple dot blot analysis. For the tags to be useful, they need to be both stably inherited and phenotypically silent or neutral. However, it is clear that tagging methodologies are restricted to species for which transgenic plants can be recovered. Very few, if any, new varieties are "created" in the laboratory. Although we have to recognize this possibility, it is still impractical and can thus not be considered in the protection of plant breeder's rights.

The ability to patent plants in certain countries like the United States of America, has led to a greater emphasis on developing methods of detection of the similarities and differences between plants (Jondle, 1992). The most important advantage of the introduction of biochemical and molecular markers for plant breeder's rights testing, is that the results are independent from environmental factors. A further advantage of these techniques is that the parents of hybrids can be determined, thereby saving time and costs in that plants can be tested at an early stage. In contrast to DNA based characteristics, secondary metabolites (phenolics, pigments, lipids, etc.) which are products of a complex series of biochemical reactions, are normally severely affected by the environment, nutritional status, latitude and stage of plant development. They therefore provide no advantage for varietal characterization.

An important prerequisite of molecular techniques, is the repeatability there-of. RAPDs meet this essential prerequisite of revealing repeatable differences among citrus cultivars, at least in our laboratory. Isozyme analysis requires the application of a range of different enzyme systems to develop a "fingerprint" profile. Enzymes will vary for different tissues and species. By contrast, there is only one standard RAPD analytical procedure for all tissues and species.

DNA-techniques do not differentiate between the coding and non-coding regions of the genome, unless a specifically designed probe or primer is used to reveal polymorphisms in a specific gene. If small differences in the DNA profile are acceptable as a basis for the protection of new varieties, it might lead to the exploitation of the system. The aim of UPOV is to give a good solid identity and to provide reliable protection to a new variety (Buitendag, 1993).

There is currently no requirement that DNA fingerprinting be submitted in support of an application for the protection of a variety, but it will definitely play an increasingly important role in the acquiring and enforcing of plant breeder's rights. This phenomenon can be explained by considering the following example: Assume that a sunflower plant that is protected by breeder's rights, has an oleic acid content of 80%. A new sunflower plant, having an oleic acid content of 85% may not be novel and therefore not liable to be protected by law. If it is possible, however, to show by DNA fingerprinting that the 85% oleic acid content is the result of a different genetic pathway or a difference in DNA composition, the new sunflower plant may be recognized as unique. Because it will be necessary to claim these differences in order for the new sunflower to be protected, its DNA fingerprint may

be a part of the application if the difference between the new and existing variety cannot adequately be written in the application (Jondle, 1992).

Plant fingerprint analysis has many applications in both commercial and legal environments. An important use of fingerprint analysis in breeding programmes is to determine relatedness of genotypes. While RAPD analysis can play an important role as an objective means of establishing varietal distinctness, it provides no descriptive information. Visual (morphological and physical) characteristics are important in marketing and to minimize infringement. It is therefore still uncertain if DNA profiling techniques will ever completely replace or only have a complementary role to conventional variety characteristics.
I believe that plant breeder's rights should not be awarded solely on the results obtained by DNA profiling techniques, but that descriptive characters, performance data, naturally acquired hidden traits and RAPD profiles should complement each other when the granting of plant breeders' rights is considered. I summarise my argument in the words of Smith and Chin (1992): "In varietal protection, the need for large databases and wide acceptance of the technology and its standards of practice will place very high demands on any new technologies before the can conceivably replace those that are currently employed."

CHAPTER 5

DISCUSSION AND CONCLUSION

In contrast to isozyme analysis and DNA fingerprinting, the analysis of secondary metabolites (phenolics, pigments, lipids, etc.), which are products of a complex series of biochemical reactions, are normally severely affected by the environment, nutritional status, latitude and stage of plant development. They, therefore provide no advantage for varietal characterization.

Isozyme analysis requires the application of a range of different enzyme systems to develop a "fingerprint" profile. Enzymes vary for different tissues and species. Zygotics, which originated by hybridization can be distinguished from nucellars or selfs if the male parent has alleles not found in the female. Cultivars which are known to have developed by mutation, such as the true grapefruit cultivars, cannot be distinguished from one another by isozyme analysis. Genetics of isozymes in citrus leaves were widely studied by Torres *et al.* (1978; 1982) and they detected that well recognized groups of citrus cultivars are fairly homogenous. There are several species in which most cultivars cannot be distinguished by isozyme analysis. These include sweet oranges, sour oranges and trifoliate orange (Roose, 1988). Isozyme analysis is most suitable for routine screening of large numbers of progeny in order to distinguish between nucellars and zygotics, and between selfed progeny and hybrids. However, some genetic differences among cultivars may not be visualized by isozyme profiles, since the number of loci that can be resolved is limited and because of the

inability of isozymes to reveal polymorphisms that originated through single base pair changes.

DNA-based diagnostics are now well established as a means to assay diversity at whole genome levels. As the technology has advanced, DNA sequence-based assays have become easier to use and more efficient at screening for nucleotide sequence-based polymorphisms. Compared to isozyme techniques, analyzing DNA has many advantages: it is independent of environmental conditions; DNA sequence is identical throughout the genome, no matter what tissue or developmental stage is analyzed, and the number of scorable loci is unlimited. RFLP is more demanding in its requirement for technical expertise since it requires DNA extraction, restriction enzyme digestion, gel electrophoresis, blotting onto a membrane, the preparation of a labelled probe (usually radio-active), hybridization, filter washing and auto-radiography. By contrast, there is only one standard RAPD analytical procedure for all tissues and species, which consists of DNA extraction, PCR-based DNA amplification and electrophoretic separation of amplified fragments. RAPDs are useful for DNA fingerprinting where there is a need to identify varieties of a crop species or to determine parentage in a breeding population. In addition to varietal identification, phylogenetic relationships, particularly at the intra-specific level, can be examined using RAPDs.

For RAPD analysis, the DNA sequences used as primers, are short oligonucleotides which can be quickly synthesized or obtained from commercial companies. In the case of RFLP, a cDNA or genomic library is normally prepared and individual inserts are then used as probes.

The bands on autoradiographs that are the result of RFLP analysis are known to have closely related sequences; they would not hybridise to the labelled probe if this was not the case. Identical sized restriction fragments from different genotypes are interpreted as representing genetic similarities, where-as different-sized fragments are interpreted as genetic differences (Thornmann and Osborn, 1992). There is a chance that two highlighted fragments from different cultivars which have the same size as monitored by gel electrophoresis may not actually have the same sequence. In the analysis of RAPD data, it is usually assumed that co-migrating DNA fragments are identical but in a small number of cases this will not be true. Further work is, however, necessary to demonstrate that RAPD fragments might be highly conserved in certain taxa (Weeden *et al.*, 1992). At present, this problem is so serious that the value of RAPD in phylogenetic research data must be regarded as highly suspect.

RFLP results allow the discrimination of two alleles at a locus if the two alleles have different sizes. RAPD markers are considered dominant; polymorphism will be detected because a DNA sequence will be amplified from one individual, but may not be amplified from another. A size difference between alleles as monitored by RAPD is more likely to result in the presence or absence of a band. RAPDs do therefore not allow the discrimination between plants that are homozygous or heterozygous for a particular amplifiable sequence; in both cases a band of the same size is observed.

Since RAPD markers can be shown to segregate in a Mendelian fashion, it can result in the early detection of hybrid progeny in a breeding programme, saving considerable time in the production of new cultivars.

The mean frequency with which polymorphism is detected, is higher per primer when RAPDs are used, than it is per probe using the RFLP method. This is partly due to the fact that more scorable bands are typically produced using a single primer.

Advantages of RAPDs over RFLP analysis include the following: the equipment and supplies necessary are inexpensive relative to those needed for RFLP analysis; the speed of analysis is less than two days, since Southern blotting and labelled probes are not necessary.

Because of the enormous amplification of very low initial quantities of DNA sequence, all PCR based techniques are stated to be prone to artifacts caused by contamination of the reaction mixture by foreign DNA. This competition factor may not be large, but should be borne in mind when data is analyzed.

Finally, RAPD analysis provides a very powerful and exciting set of DNA markers that potentially can be applied in many of the same ways as isozyme analysis and RFLPs. Although DNA profiling techniques have limitations, the results of this study show that these techniques are to be used complementary to classical techniques for varietal identification in citrus. The ease of detecting RAPD markers make them an attractive choice for routine screening of large germplasm collections and the determination of genetic relationships. The ultimate genetic assay would, however, be based on the determining of the complete DNA sequence at any locus.

REFERENCES

ASHARI, S., ASPINALL, D. & SEDGLEY, M. 1989. Identification and Investigation of Relationships of Mandarin Types using Isozyme Analysis. *Scientia Horticulturae*, **40**: 305-315.

BECKMAN, J.S. & BAR-JOSEPH, M. 1986. The use of Synthetic DNA Probes in Breeders' Rights Protection: A Proposal to Superimpose an Alpha-numerical Code on the DNA. *Trends in Biotechnology*, **4**(9): 230-323.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, M. & DAVIS, R.W. 1980. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am. J. Human Genet.*, **32**: 314.

BUITENDAG, E. 1993. Travel Report. First Meeting of the BMT Working Group, organised by UPOV. Geneva, 19-20 April. (Unpublished).

BURR, E., EVOLA, S.V., BURR, F.A. & BECKMAN, J.S. 1988. The Application of Restriction Fragment Length Polymorphism to Plant Breeding. (*In*: SETLOW, J.K. & HOLLAENDER, A. Genetic Engineering: Principles and Methods, volume 5. Planum Publishing Corporation.).

CAETANO-ANOLLES, G., BASSAM, B.J. & GRESSHOFF, P.M. 1992. DNA Amplification Fingerprinting: A Strategy for Genome Analysis. *Plant Mol. Biol. Rep.*, **9**: 294-307.

CITRUS AND SUBTROPICAL FRUIT RESEARCH INSTITUTE. 1988. Botanical Classification of Citrus. 1 p. (CSFRI. Leaflet Citrus A.2.)

DEVOS, K.M. & GALE, M.D. 1992. The Use of Random Amplified Polymorphic DNA Markers in Wheat. *Theor. Appl. Genet.*, **84**: 567-572.

DOYLE, J.J. & DOYLE, J.L. 1990. Isolation of Plant DNA from Fresh Tissue. *Focus*, **12**: 13-15.

ECHT, C.S., ERDAHL, L.A. & McCOY, T.J. 1991. Genetic Segregation of Random Amplified Polymorphic DNA in Diploid Cultivated Alfalfa. *Genome*, **35**: 84-87.

ERLICH, H.A., GELFAND, D. & SNINSKY, J.J. 1991. Recent Advances in the Polymerase Chain Reaction. *Science*, **252**: 1643-1650.

GOGORCENA, Y. & ORTIZ, J.M. 1988. Morphometric and Biochemical Characterization of Spanish Sour Orange. *Proc. 6th Int. Citrus Congress*. p.27-36.

GOGORCENA, Y. & ORTIZ, J.M. 1993. Use of Multivariate Analysis in the Taxonomy of *C. aurantium* L. and Relatives. *Scientia Hortic.*, **53**: 301-310.

GOGORCENA, Y., ZUBRZYCKI, H. & ORTIZ, J.M. 1990. Identification of Mandarin Hybrids with the aid of Isozymes from Different Organs. *Scientia Horticulturae*, **41**: 285-291.

HANSEN, C. 1983. De Citrusteelt in Zuid-Afrika. (Thesis: Industriele Ingenieur, Industriele Hogeschool het Rijk C.T.L. Gent).

HU, J. & QUIROS, C.F. 1991. Identification of Broccoli and Cauliflower Cultivars with RAPD Markers. *Plant Cell Reports*, **10**: 505-511.

INNES M.A., GELFAND D.H., SNINSKY J.J. & WHITE T.J., *eds*.. PCR Protocols: A Guide to Methods and Applications. San Diego, California: Academic Press. 482p.

IRVINE, J.E. & MOORE, P.H. 1991. Biotechnology and its Application in Agriculture. *Proceedings of the South African Sugar Technologists' Association*. p.19-22.

ISABEL, N., TREMBLAY, M., MICHAUD, M., TREMBLAY, F.M. & BOUSQUET, J. 1993. RAPDs as an aid to Evaluate the Genetic Integrity of Somatic Embryogenesis-Derived Populations of *Picea mariana* (Mill.) B.S.P. *Theor. Appl. Genet.*, **86**: 81-87.

ISH-HOROWICZ, D. & BURKE, J.F. 1981. Rapid and Efficient Cosmid Cloning. *Nucleic Acids Research*, **9**(13): 2989-2998.

JEFFREYS, A.J., WILSON, V. & THEIN, S.L. 1985. Individual-specific "Fingerprints" of Human DNA. *Nature*, **316**: 76-79.

JONDLE, R.J. 1992. Legal Aspects of Varietal Protection using Molecular Markers. (*In*: Applications of RAPD Technology to Plant Breeding: Proceedings of the joint symposium of the Crop Science Society of America; the American Society for Horticultural Science and the American Genetic Association. p. 50-52.)

KEETCH, D.P. 1992. Plant Breeders' Rights in South Africa. *Plant Varieties and Seeds*, **5**: 163-165.

LOWRIE, P. & WELLS, S. 1991. Genetic Fingerprinting. *New Scientist*, **52**: 1-4.

MANIATIS, T., FRITSCH, E.F. & SAMBROOK, J. 1987. Molecular cloning: a laboratory manual. Cold Spring Harbour Press, Cold Spring Harbour, N.Y.

NEWBURY, H.J. & FORD-LLOYD, B.V. 1993. The Use of RAPD for Assessing Variation in Plants. *Plant Growth Regulation*, **12:** 43-51.

PRESTON, L.R., APPELS, R., LEE, L.S., BULLER, C.D.S & MORELL, M.K. 1993. Differentiation of Mandarin Varieties by DNA Profiling Techniques. (Lecture given at the first meeting of the working group on biochemical and molecular techniques and DNA profiling in particular of the International Union for the Protection of New Varieties of Plants (UPOV), 19-20 April. Geneva. (Unpublished.)

PROTOPAPADAKIS, E.E. 1987. Identification by Isozymes of Five Cultivars of *Citrus medica* Grafted on Four Rootstocks. *Journal of Hort. Science*, **62**(3): 413-419.

RAFALSKI, J.A., TINGEY, S.V. & WILLIAMS, J.G.K. 1991. RAPD Markers - A New Technology for Genetic Mapping and Plant Breeding. *AgBiotech. News Inf.*, **3**: 645-648.

ROOSE, M.L. 1988. Isozymes and DNA Restriction Fragment Length Polymorphisms in Citrus Breeding and Systematics. *Proc. 6th Int. Citrus Congress*. p.155-165.

ROOSE, M.L. & TRAUGH, S.N. 1988. Identification and Performance of Citrus Trees on Nucellar and Zygotic Rootstocks. *J. Amer. Soc. Hort. Sci.*, **113**(1): 100-105.

SCOTT, M.P., HAYMES, K.M. & SCOTT. M.W. 1992. Parentage Analysis using RAPD PCR. *Nuc. Acids Res.*, **20**(20): 5493.

SKROCH, P., TIVANG, J. & NIENHUIS, J. 1992. Analysis of Genetic Relationships using RAPD Marker Data. (*In*: Applications of RAPD Technology to Plant Breeding: Proceedings of the joint symposium of the Crop Science Society of America; the American Society for Horticultural Science and the American Genetic Association. p.26-30.)

SMITH, S. & CHIN, E. 1992. The Utility of Random Primer-mediated Profiles, RFLPs and Other Technologies to Provide Useful Data for Varietal Protection. (*In*: Applications of RAPD Technology to Plant Breeding: Proceedings of the joint symposium of the Crop Science Society of America; the American Society for Horticultural Science and the American Genetic Association. p. 46-49.)

SOOST, R.K. & WILLIAMS, T.E. 1980. Identification of Nucellar and Zygotic Seedlings of *Citrus* with Leaf Isozymes. *HortScience*, **15**(6): 728-729.

SOUTHERN, E.M. 1975. Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis. *J. Mol. Biol.*, **98**: 503.

STILES, J.I., LEMME, C., SONDUR, S., MORSHIDI, M.B. & MANSHARDT, R. 1993. Using Randomly Amplified Polymorphic DNA for Evaluating Genetic Relationships among Papaya Cultivars. *Theor. Appl. Genet.*, **85**: 697-701.

SWINGLE, W.T. & REECE, P.C. 1967. The Botany of Citrus and its Wild Relatives. *The Citrus Industry*, **1**: 190-422.

THORMANN, C.E. & OSBORN, T.C. 1992. Use of RAPD and RFLP Markers for Germplasm Evaluation. (*In*: Applications of RAPD Technology to Plant Breeding: Proceedings of the joint symposium of the Crop Science Society of America; the American Society for Horticultural Science and the American Genetic Association. p. 9-11.)

TINGEY S.V. & del TUFO J.P. 1993. Genetic Analysis with Random Amplified Polymorphic DNA Markers. *Plant Physiol.*, **101**: 349-352.

TORRES, A.M., MILLáN, T. & CUBERO, J.I. 1993. Identifying Rose Cultivars using Random Amplified Polymorphic DNA Markers. *HortScience*, **28**(4): 333-334.

TORRES, A.M., SOOST, R.K. & DIEDENHOFEN, U. 1978. Leaf Isozymes as Genetic Markers in Citrus. *Amer. J. Bot.*, 65 (8): 869-881.

TORRES, A.M., SOOST, R.K. & MAU-LASTOVICKA, T. 1982. *Citrus* Isozymes - Genetics and Distinguishing Nucellar from Zygotic Seedlings. *J. Hered.*, **73**: 335-339.

UPOV. 1992. Report of the Technical Committee: Determination of Distinctness, Uniformity and Stability of Varieties using DNA Profiling Techniques. Geneva. (TC/28/4.) (Chairman: H.L. Lloyd.)

WEEDEN, N.F., TIMMERMAN, G.M., HEMMAT, M. KNEEN, B.E. & LODHI, M.A. 1992. Inheritance and Reliability of RAPD markers. (*In*: Applications of RAPD Technology to Plant Breeding: Proceedings of the joint symposium of the Crop Science Society of America; the American Society for Horticultural Science and the American Genetic Association. p. 46-49.)

WEINING, S. & LANGRIDGE, P. 1991. Identification and Mapping of Polymorphisms in Cereals Based in the Polymerase Chain Reaction. *Theor. Appl. Genet.*, **82**: 209-216.

WELSH, J. & McCLELLAND, M. 1990. Fingerprinting Genomes using PCR with Arbitrary Primers. *Nuc. Acids Res.*, **18**(24): 7213-7218.

WENDEL, J.F. & WEEDEN, N.F. 1989. Visualization and Interpretation of Plant Isozymes. (*In*: SOLTIS, D.E. & SOLTIS, P.S., *eds.* Isozymes in plant biology. Portland, Oregon: Dioscorides Press. p.5-45.)

WENPIN, L., SHANWEN, H. & GENGFENG, L. 1988. A Study of Citrus Natural Resources in Hunan by Analysis of Leaf Isozymes. *Proc. Int. Hort. Germplasm Symposium*. p1-7.

WILLIAMS, J.G.K., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A. & TINGEY, S.V. 1990. DNA Polymorphisms Amplified by Arbitrary Primers are Useful as Genetic Markers. Nucleic Acids Research, **18**(22): 6531-6535.

WILLIAMS, J.G.K., RAFALSKI, J.A. & TINGEY, S.V. 1992. Genetic Analysis using RAPD Markers. Method. Enzymol.

ZABEAU, M. 1992. AFLP: A Practical Solution to Measurement of Genetic Distance and Dependency Issues. (Lecture given at the twenty-eighth meeting of UPOV's Technical Committee, 21-23 October. Geneva. (Unpublished).

APPENDIX A

COMPOSITION OF STAINS FOR ENZYMES

Glutamate oxaloacetate transaminase (GOT)

α-keto Glutaric acid	50 mg
Aspartic acid	100 mg
0.1 M Tris buffer (do not pH)	50 ml
dissolve to pH 8.0 (diluted HCl)	
Just before pouring, add:	
Fast Blue BB salt	50 mg
Pyridoxal 5-phosphate	5 mg

Phosphogluco-mutase (PGM)

0.1 M Tris-Cl, pH 8.5	35 ml
0.1 M $MgCl_2$	5 ml
Glucose-1-phosphate	85 mg
Glucose-6-phosphate dehydrogenase	40 units
NADP (10 mg/ml)	0.75 ml
MTT (10 mg/ml)	1.5 ml
PMS (10 mg/ml)	0.3 ml

Phosphogluco-isomerase (PGI)

0.2 M Tris-Cl, pH 8.0	35 ml
0.1 M $MgCl_2$	5 ml
Fructose-6-phosphate	20 mg
Glucose-6-phosphate dehydrogenase,	
(sodium-salt)	20 units
NADP (10 mg/ml)	1 ml
MTT (10 mg/ml)	1.5 ml
PMS (10 mg/ml)	0.3 ml

Malate dehydrogenase (MDH)

0.1 M Tris-Cl, pH 8.0	45 ml
2 M Malate, pH 7.0	5 ml
Just before pouring, add:	
NAD	1 ml
MTT	1 ml
PMS	0.25 ml

Isocitrate dehydrogenase (IDH)

0.1 M Tris-Cl, pH 8.0	6 ml
Isocitric acid (Disodium-salt)	10 mg
0.1 M $MgCl_2$	2 ml
Just before pouring, add:	
NADP	0.4 ml
MTT	0.4 ml
PMS	0.1 ml
1% agarose (80 °C)	7 ml

APPENDIX B
GLOSSERY OF TERMS

Alleles	Different forms of a gene
Autoradiography	The visualisation of radioactivity by exposure to an X-ray film
Coding	Those areas of the genome which are transcribes into RNA - leading to a protein product
DNA	Deoxyribonucleic Acid. The carrier of the genetic information in cells; Composed of 2 complementary chains of nucleotides wound in a double helix; capable of self replicating as well as coding for RNA synthesis
Electrophoresis	The separation by charge of nucleic acid or protein in a gel
Genome	The complete set of chromosomes (DNA), with their genes
Hybridisation	Binding of fragments of nucleic acids to compatible regions of the genome
Markers	Short fragments of DNA which bind to the genome at specific locations determined by their sequence
Non-coding	Portions of the genome which do not encode for RNA or protein products
Nucleotides	The basic unit of nucleic acids. There are 5 types: Guanine, Adenosine, Cytosine, Thymine and Uracil. Thymine is only found in DNA, and is substituted by Uracil in RNA
Oligonucleotide	Lengths of nucleic acids
Polymorphism	The presence in a population of 2 or more phenotypically distinct forms of a trait
Primers	Short fragments of nucleic acids which bind to the genome at specific locations determined by their sequence; act as starting points for nucleic acid replication
Probes	Fragments of nucleic acids incorporating radioactively, enzymatically or fluorescently labelled nucleotides which bind to the genome at specific locations determined by their sequence allowing the visualisation of these points of hybridisation

Restriction Enzyme	Enzymes that cleave the double helix at specific nucleotide sequences
Sequence	The pattern of nucleic acids in the DNA molecules
Southern Hybridisation	The process of hybridising DNA probes with DNA bound to a membrane support

APPENDIX C

ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
CIP	Citrus Improvement Programme
CTAB	N-Cetyl-N,N,N-Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid disodium
GOT	Glutamate oxaloacetate transaminase
IDH	Isocitrate dehydrogenase
MDH	Malate dehydrogenase
MTT	3-[4,5-Dimethylthiazol-2-yl]2,5-diphenyltetrazolium bromide
NAD	Nicotinamide adenine nucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NTP	Nucleotide Triphosphate
PBR	Plant Breeder's Rights
PCR	Polymerase chain reaction
PGI	Phosphogluco-isomerase
PGM	Phosphogluco-mutase
PMS	Phenazine methosulfate
PVP	Polyvinyl-Pyrrolidone
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SDS	Sodium Dodecylsulphate
UPOV	International Union for the Protection of New Varieties of Plants

APPENDIX D

DATA SETS AND DENDROGRAMS

Grapefruit Primer 1

Nelruby	1111011
Ray	1111011
Star	1101110
Redblush	1111010
Marsh	1111111

Grapefruit Primer 2

Marsh	11101
Nelruby	11101
Star	11111
Redbl	11111
Hennin	00011
Java	10111
Ray	11111
Riored	11111
Hender	11111
Ruben	11111
Oran	11111
Flame	11111

Grapefruit Primer 3

Marsh	1111111111
Nelruby	1111111111
Star	1111111110
Redblsh	1111111000
Henning	0111111111
Java	0111111111
Rayruby	1111111111
Riored	1111111111
Henders	1111111111
Ruben	0101011000
Oran	0111111111
Flame	1111111111

Grapefruit Combined Data Set

Nelruby	11110111111111111101111
Rayruby	11110111111111111111111
Star	1101111011111111111110
Redblsh	1111011011111111110010
Marsh	1111111111111111101111
Henning	9999990101110101111111
Java	999999090111011111111
Riored	9999991911111111111111
Henders	9999991911111111111111
Ruben	9999991901011011110010
Oran	999999190111111111111
Flame	9999991911111111111111

FIGURE D1. Datasets showing differences between Grapefruit (*C. paradisi*) cultivars with different primers.

Shaddock Primer 1

Pomelit 10011110011110
Chandler 000011110111101
Tahiti 111011111111111
Oroblan 000100000111101

Shaddock Primer 2

Pomelit 0010001000
Chandler 1111001100
Melogold 1111001011
Tahiti 1101101100
Oroblan 0010011001

Shaddock Primer 3

Pomelit 1001011101
Chandler 1001011001
Melogold 0111111101
Tahiti 1111011111

Shaddock Primer 4

Pomelit 0011101111
Chandler 0011101111
Tahiti 1111111111
Oroblan 1001101111

Shaddock Combined Dat Set

Pomelit 10100010010011001111010011111110
Chandler 100100100010110001010100111110111
Melogold 019119119911190119090911919999999
Tahiti 11111101101011101111110011111111
Oroblanc 990009100100990900010111111110111

FIGURE D2. Datasets showing differences between Shaddock (*C. grandis*) cultivars with different primers.

Mardarin Primer 1

```
Ellendale 110111110
Lanique   110111111
OrtaniquA 010111110
Page      101111110
Novel     000111110
Robin     000101110
Ronel     110101111
Elno      000111010
Nova      110111010
Roma      110111110
Edelgard  110111000
OrtaniquB 000111010
Novelty   001111010
Lee       110101110
Robinson  110111110
```

Mandarin Primer 2

```
Robin     101010100101
Daisy     101010100011
Sweet     000001101001
Kiyomi    100001111011
Lomati    100110001001
Minela    111010100011
Temple    100111111101
```

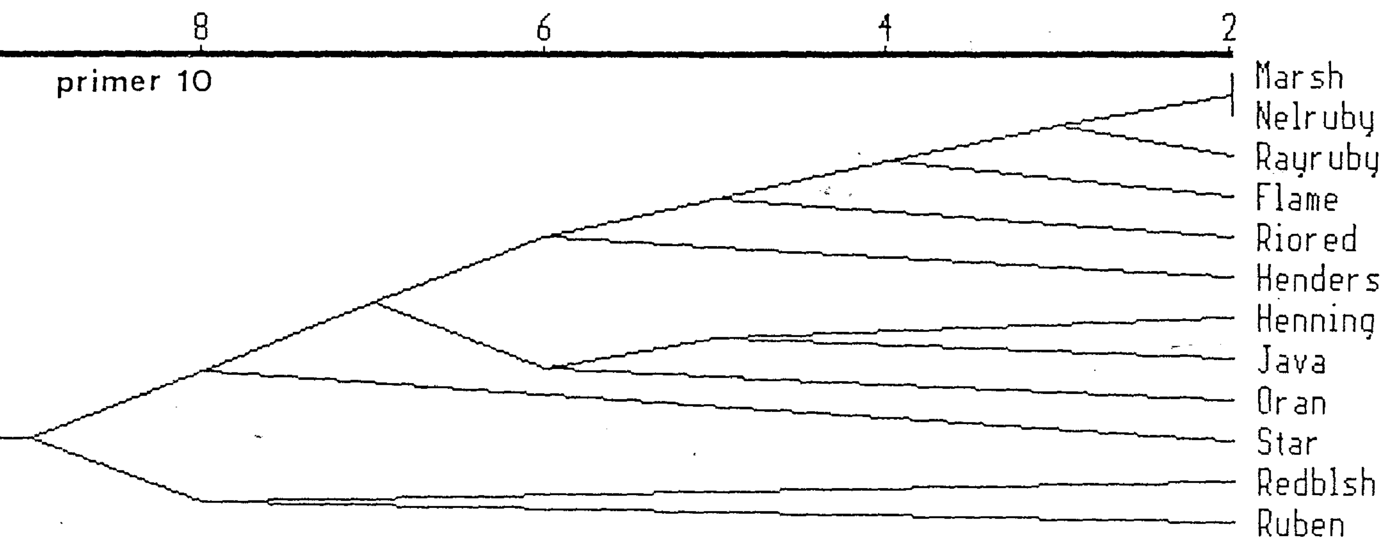
Mandarin Primer 3

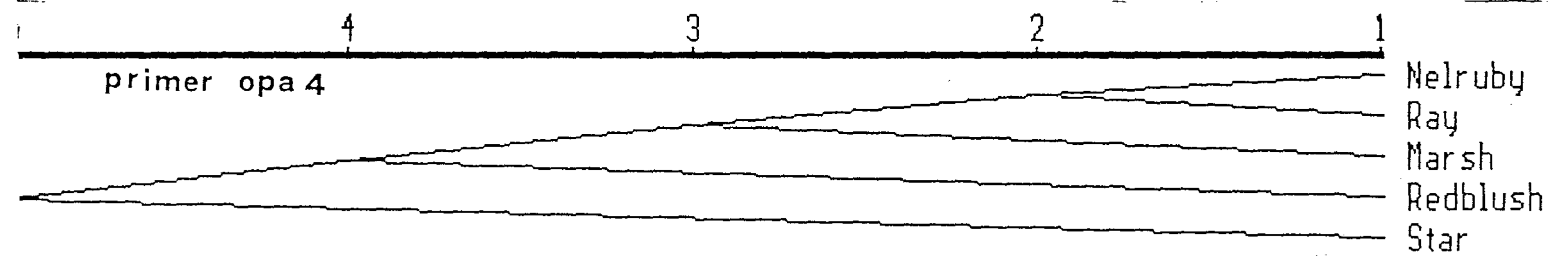
```
Temple    00001010110
Minneo    10000000010
Lomati    11001011110
Kiyomi    00101011111
Fortuna   00001110000
Sweet     00111011111
Daisy     00001111010
```

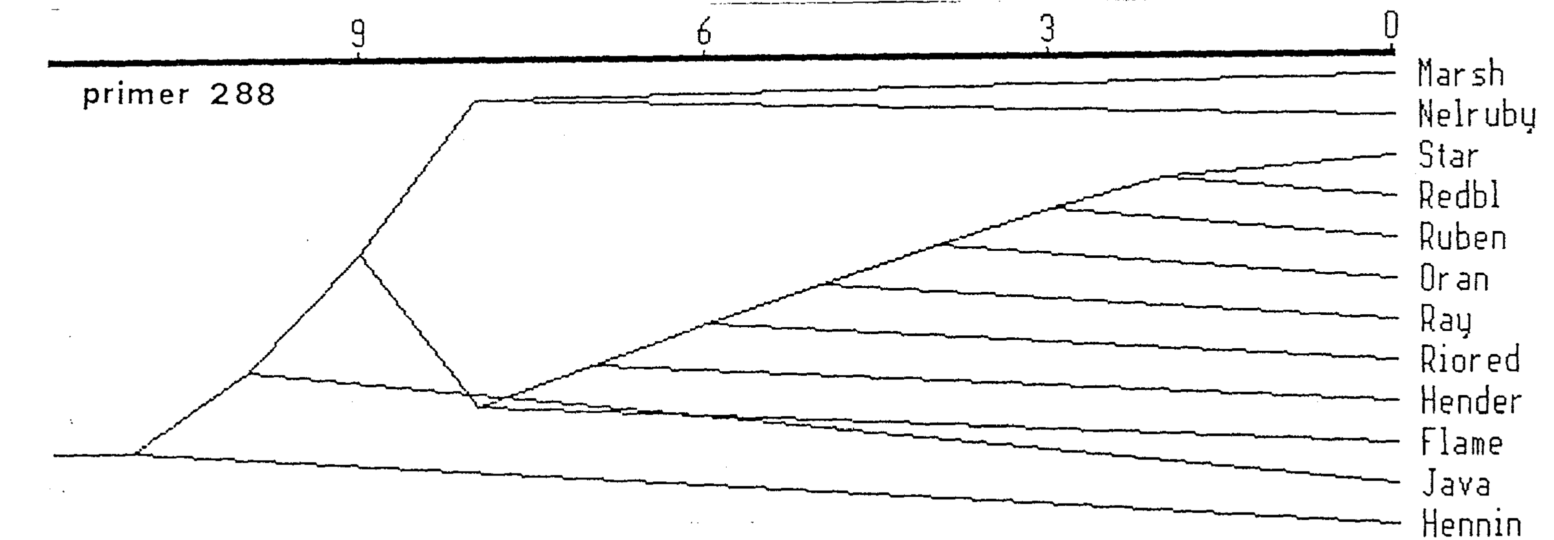
Mandarin Combined Data Set

```
Temple    001009019111191019011010
Minneol   101009109100190009000110
Lomati    111009019110090119101110
Kiyomi    001109009011191109101111
Fortuna   009009999919999119090090
Sweet     000009009011190119000010
Daisy     001009109110190119100110
Robin     901101101190100010919009
Ellen     909191191999919910999109
Lanique   919191191999919910999119
OrtaL3    909191191999919910999109
Page      909191191999919911999019
Novel     909190191999919911999009
Ronel     919191091999919910999119
Elno      909190191999919910999009
Nova      909190191999919910999119
Roma      909190091999919911999119
Edel      909090191999919910999119
OrtaL13   909190091999919910999009
Novelty   909190091999919911999009
Lee       909191191999909910999119
Robins    909191091999919910999119
```

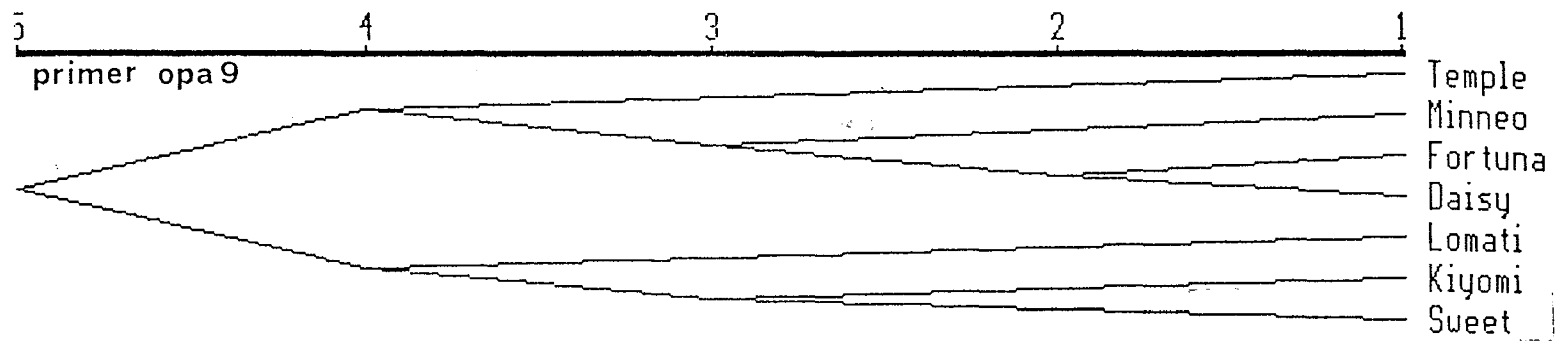
FIGURE D3. Datasets showing differences between Mandarin (*C. reticulata*) cultivars with different primers.

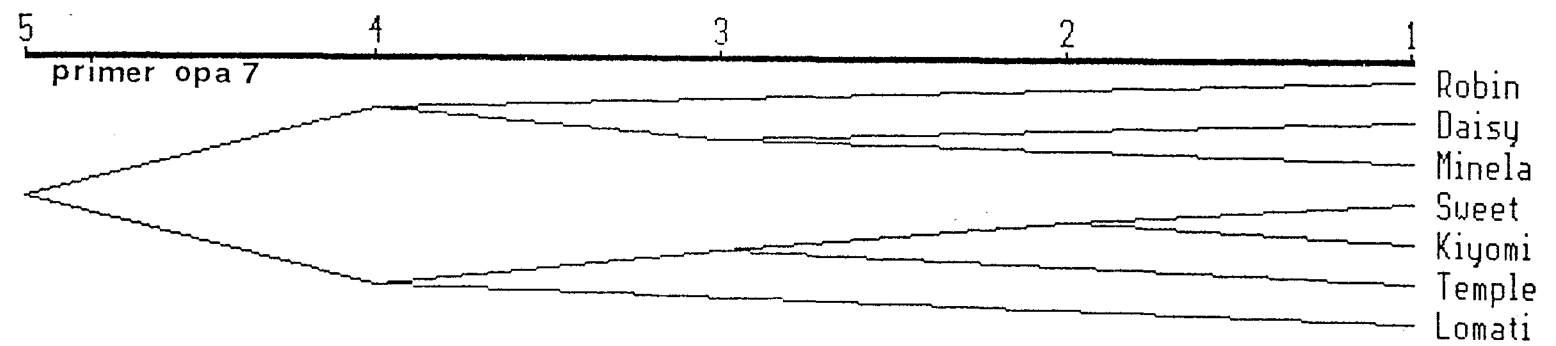


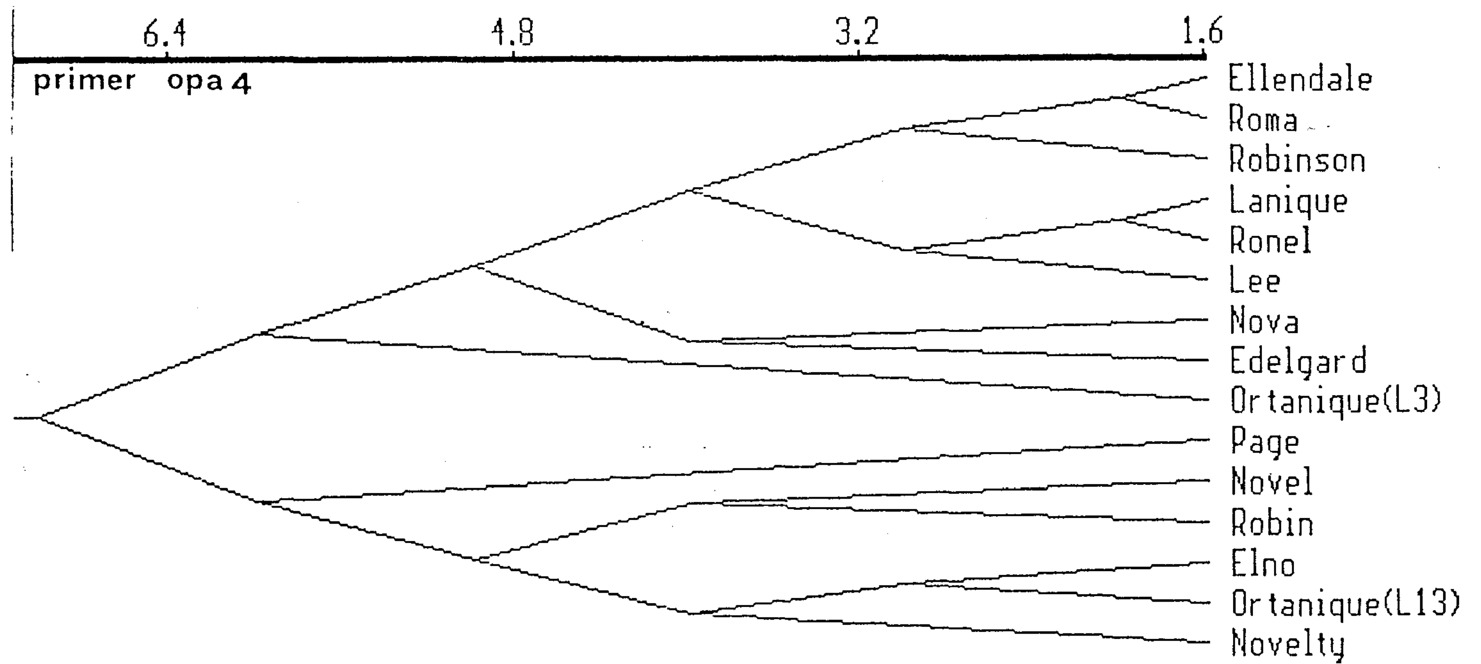




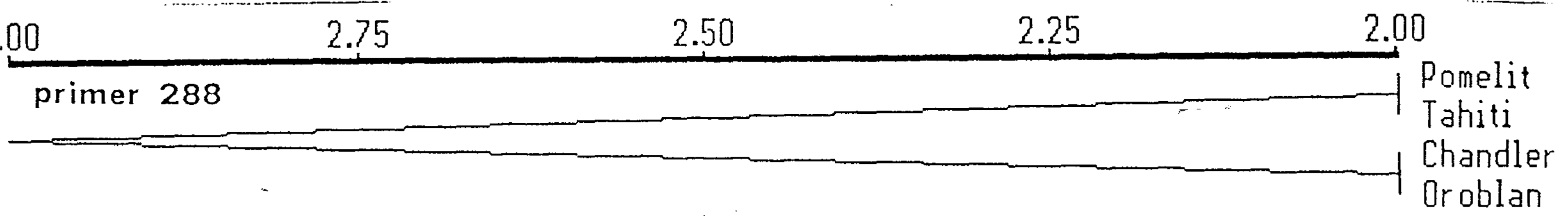
NJOIN phylogenetic trees of Grapefruit with three primers.

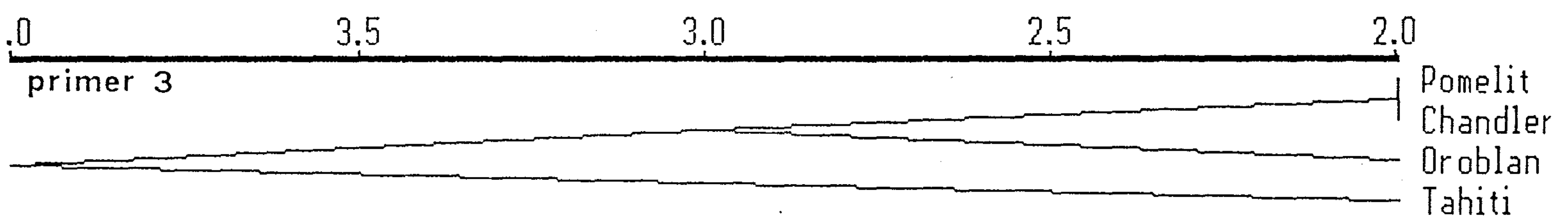


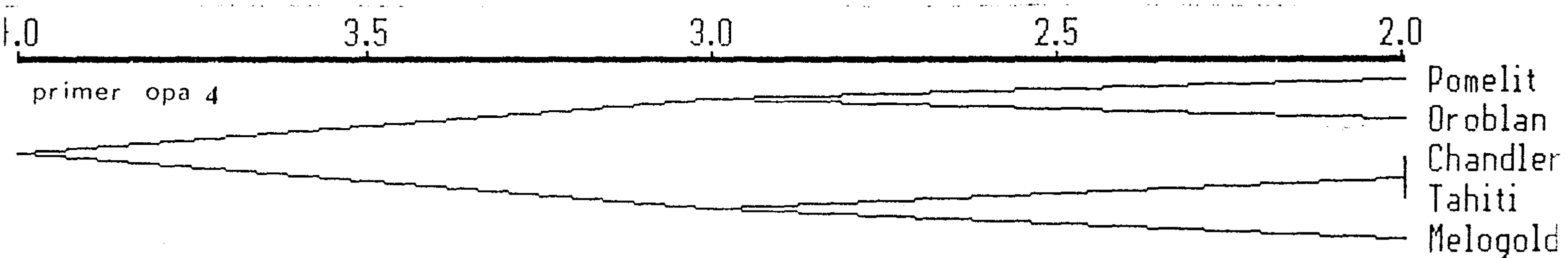


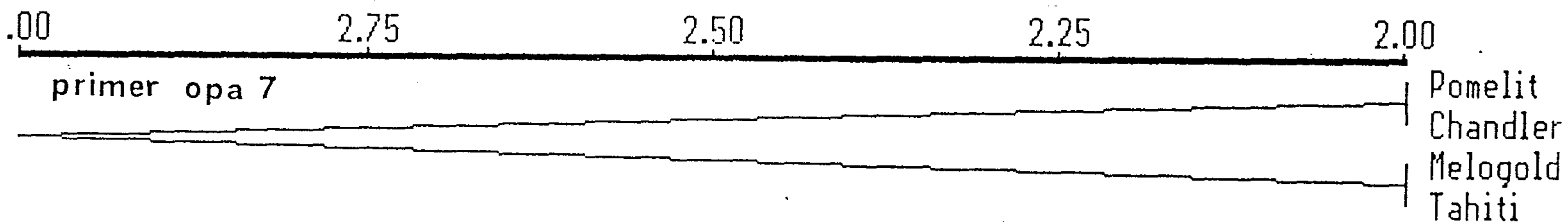


NJOIN phylogenetic trees of Mandarins with three primers.

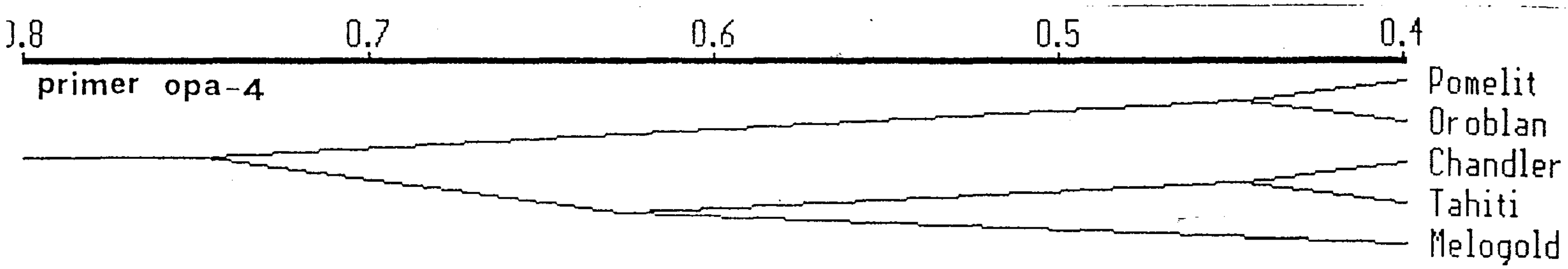


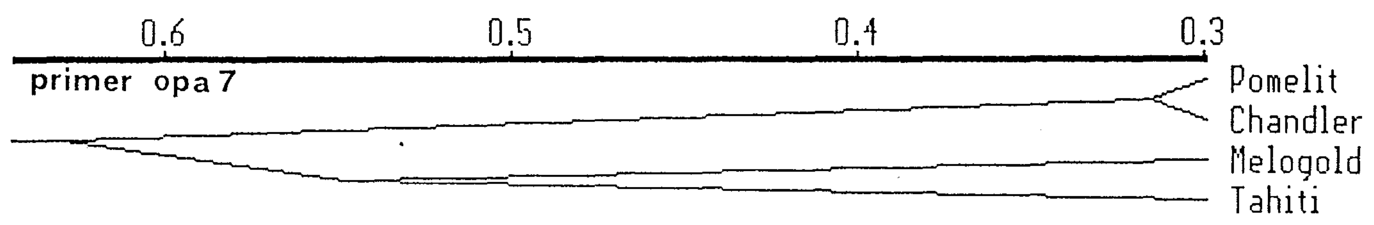


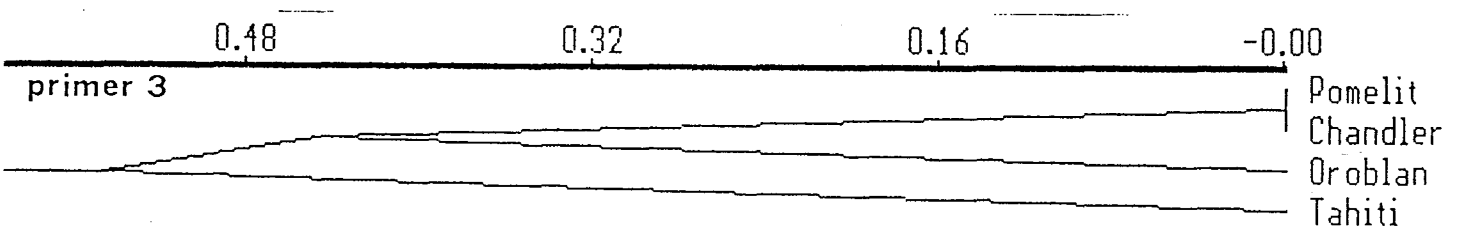


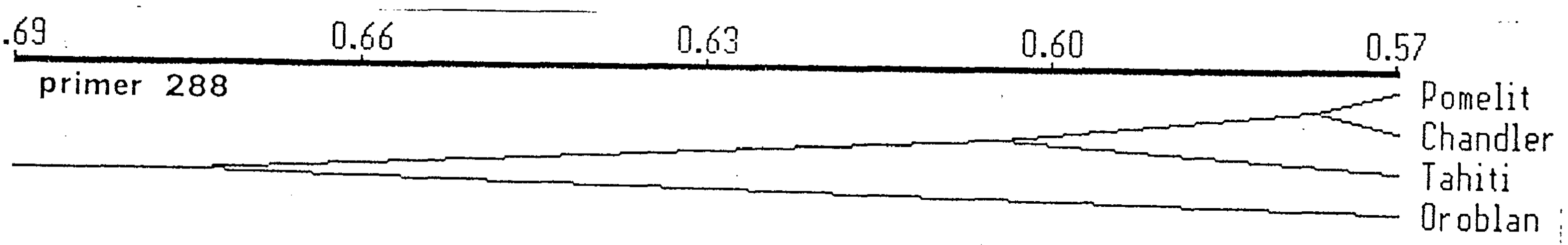


NJOIN phylogenetic trees of Shaddocks with four primers.

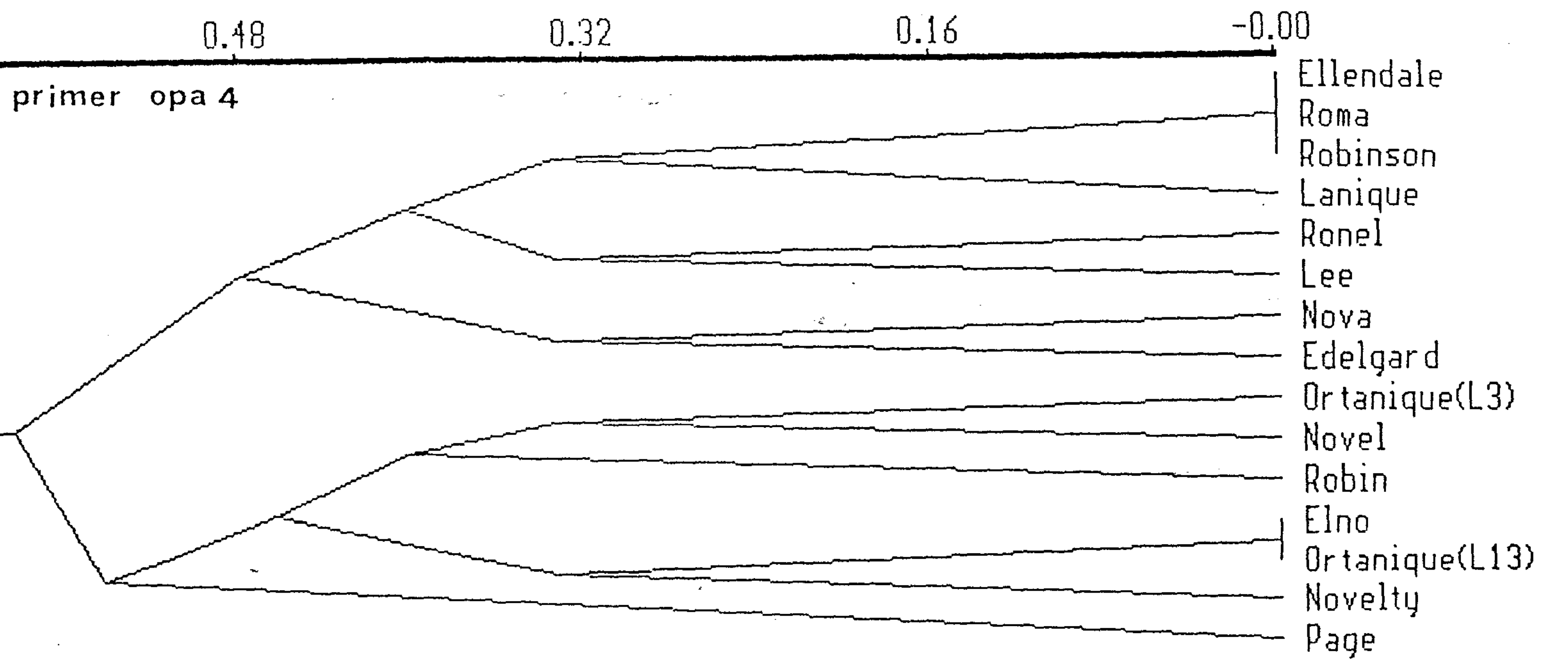


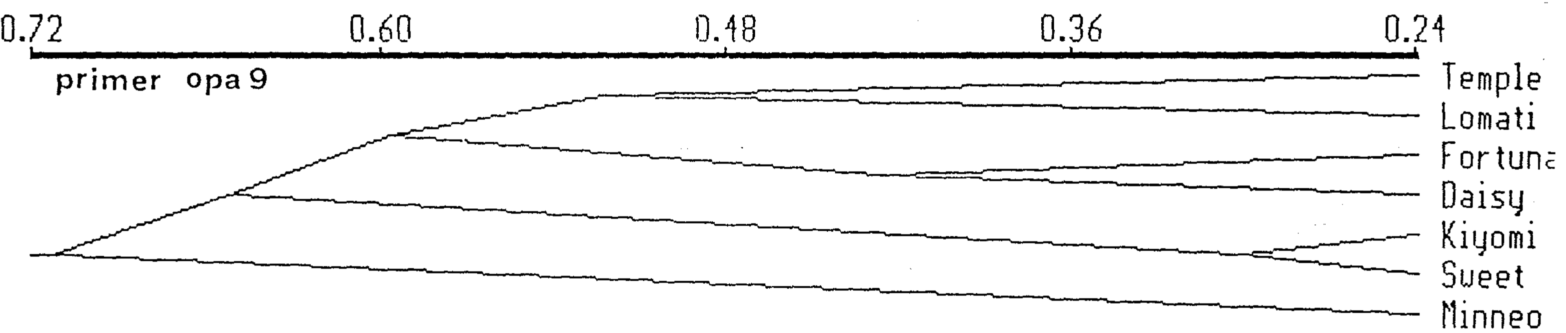


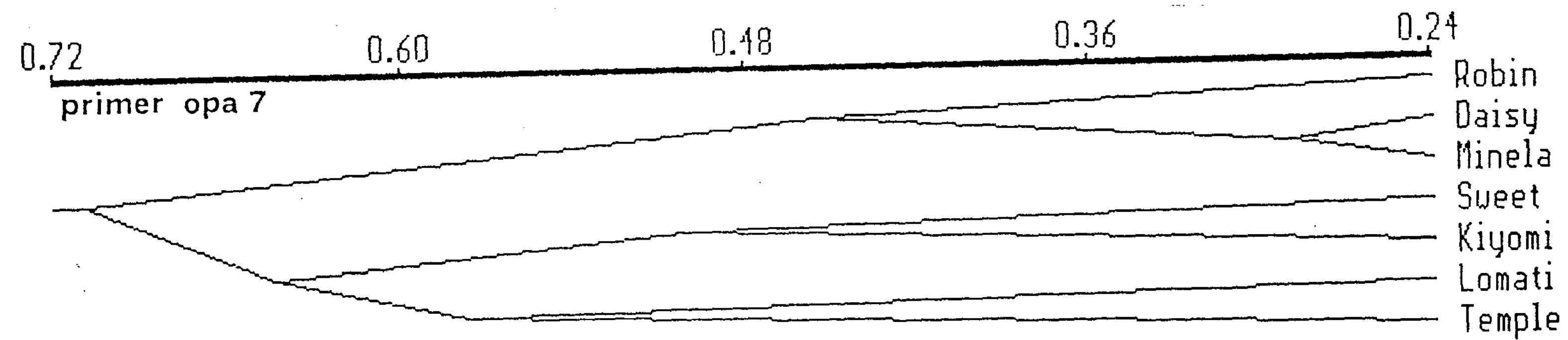




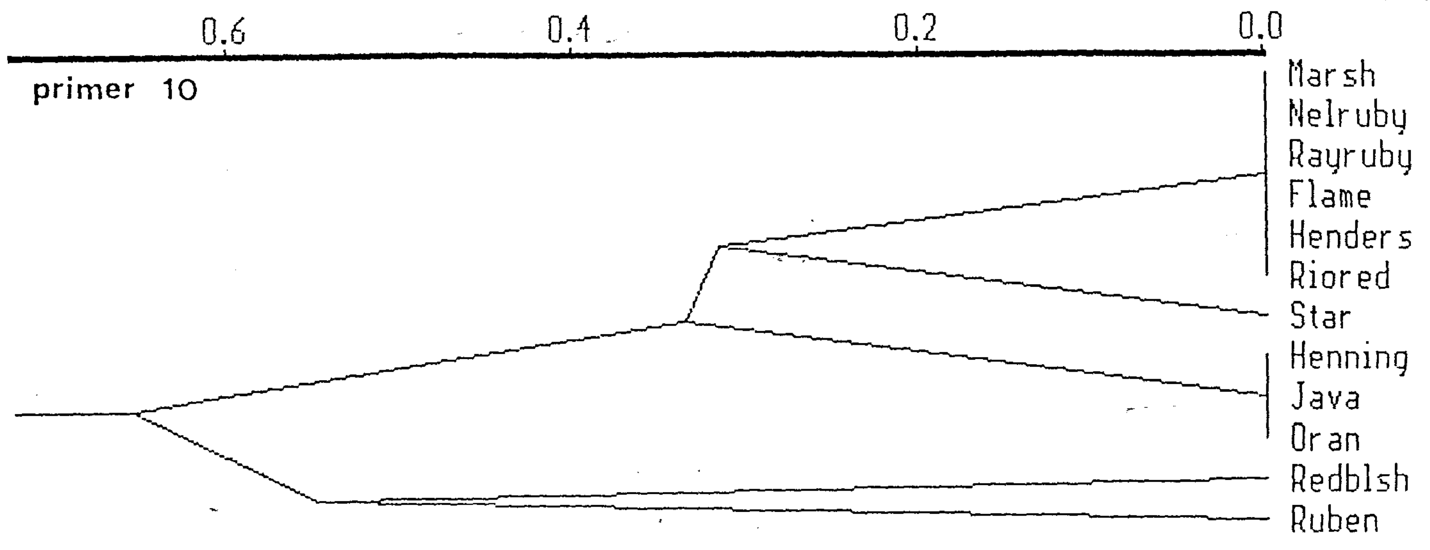
SAHN-UPGMA phylogenetic trees of Shaddocks with four primers.

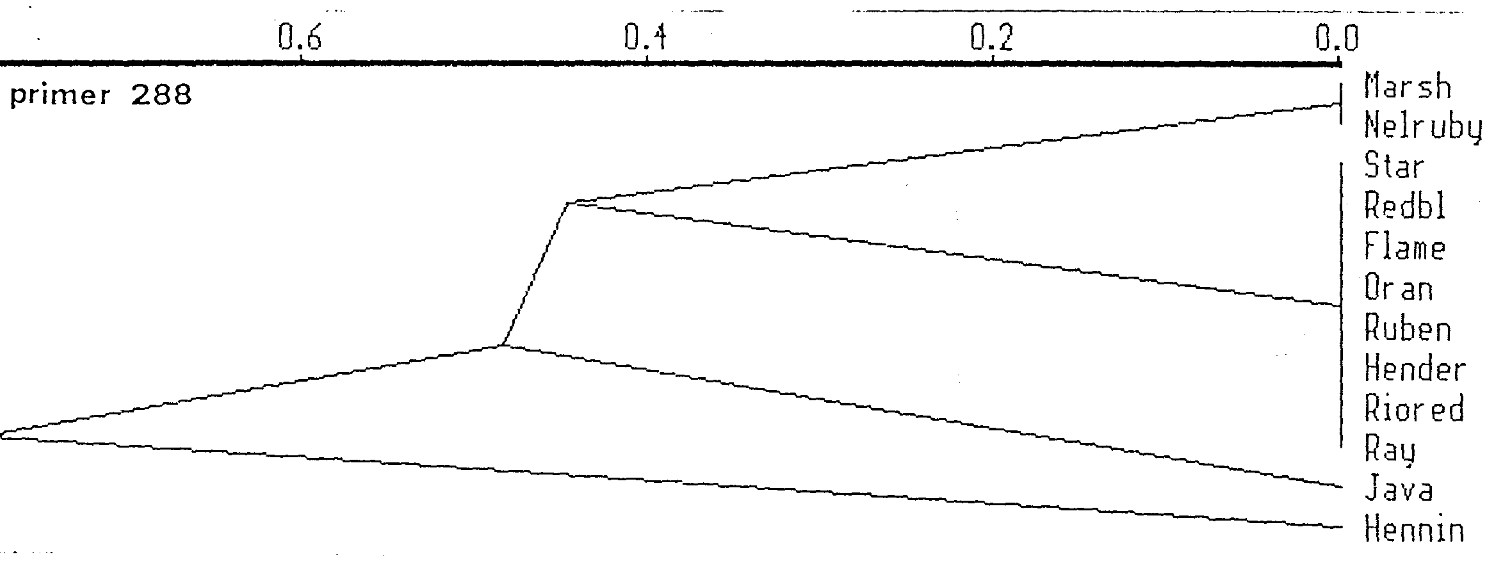


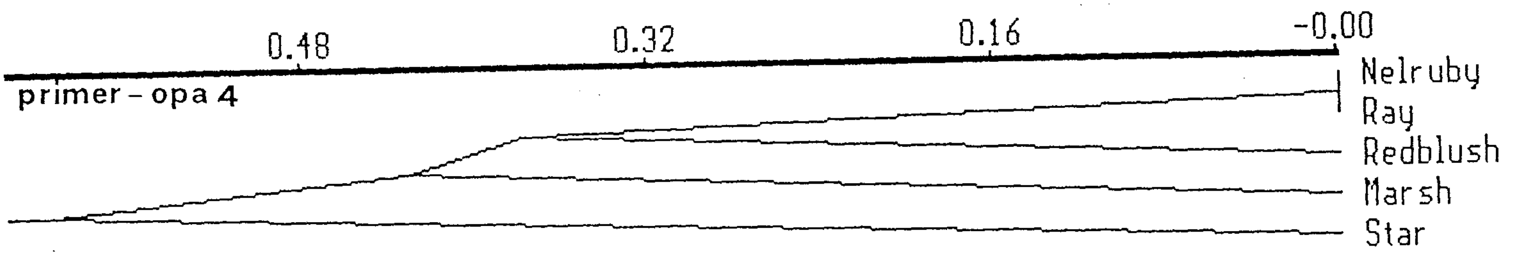




SAHN-UPGMA phylogenetic trees of MAndarins with three primers.







SAHN-UPGMA phylogenetic trees of Grapefruit with three primers.