Disclaimer: unless otherwise agreed by the Council of UPOV, only documents that have been adopted by the Council of UPOV and that have not been superseded can represent UPOV policies or guidance.

This document has been scanned from a paper copy and may have some discrepancies from the original document.

———

Avertissement: sauf si le Conseil de l'UPOV en décide autrement, seuls les documents adoptés par le Conseil de l'UPOV n'ayant pas été remplacés peuvent représenter les principes ou les orientations de l'UPOV.

Ce document a été numérisé à partir d'une copie papier et peut contenir des différences avec le document original.

———

Allgemeiner Haftungsausschluß: Sofern nicht anders vom Rat der UPOV vereinbart, geben nur Dokumente, die vom Rat der UPOV angenommen und nicht ersetzt wurden, Grundsätze oder eine Anleitung der UPOV wieder.

Dieses Dokument wurde von einer Papierkopie gescannt und könnte Abweichungen vom Originaldokument aufweisen.

———

Descargo de responsabilidad: salvo que el Consejo de la UPOV decida de otro modo, solo se considerarán documentos de políticas u orientaciones de la UPOV los que hayan sido aprobados por el Consejo de la UPOV y no hayan sido reemplazados.

Este documento ha sido escaneado a partir de una copia en papel y puede que existan divergencias en relación con el documento original.

UPOV

BMT/2/3
ORIGINAL : English
DATE : March 1, 1994

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA-PROFILING IN PARTICULAR

Second Session
Versailles, France, March 21 to 23, 1994

THE APPLICATION OF DNA PROFILING TO THE DETERMINATION OF DISTINCTNESS BETWEEN VARIETIES IN CITRUS AND IMPLICATIONS FOR VARIETAL IDENTIFICATION IN OTHER PLANT SPECIES

Document prepared by experts from Australia

THE APPLICATION OF DNA PROFILING TO THE DETERMINATION OF DISTINCTNESS BETWEEN VARIETIES IN *CITRUS* AND IMPLICATIONS FOR VARIETAL IDENTIFICATION IN OTHER PLANT SPECIES.

Document Prepared by Australia

Matthew K. Morell, Rod Peakall, Rudi Appels, Lynette R. Preston, C.D.S. Buller, Cooperative Research Centre for Plant Science, Australian National University, Canberra, A.C.T.
Mark R. Thomas and N. Steele Scott, CSIRO Division of Horticulture, Adelaide, GPO Box 350, Adelaide, SA 5001
Jamie M.H. Kijas and J.C.S. Fowler, Department of Biological Sciences, Flinders University of South Australia, Bedford Park, SA 5042
Lloyd,H.L., Plant Variety Rights Office, Canberra

I. Introduction

1. The First Session of the Working Group on Biochemical and Molecular Techniques, held at Geneva in April 1993, established four working units each with responsibility for collecting technical information on a particular plant group. It was agreed to concentrate on two types of DNA profiling methods, namely RFLP-like and RAPD-like methods, but that pertinent results from other methods should also be included.

2. Australia was given responsibility for preparing a paper on *Citrus*.

3. The purpose of this paper is to present the findings of the Australian Working Unit on *Citrus* and to extend those findings to broad issues of interest to UPOV. Specifically, this paper will:

(a) review the published literature describing the use of biochemical and molecular techniques to test for distinctness in *Citrus*.

(b) discuss procedures for DNA profiling in *Citrus*.

(c) discuss the application of DNA profiling to plant varietal identification in general.

(d) the paper will conclude with recommendations about the use of DNA profiling for *Citrus*, and more generally for varietal identification in the UPOV system.

(e) detailed appendices on profiling techniques for *Citrus*, DNA visualisation procedures and statistical analysis of DNA profiles are also attached.

II. The use of biochemical and molecular techniques to test for distinctness in Citrus.

4. A number of techniques have been used to test for distinctness in *Citrus*. These include:
(a) morphological comparison
(b) cytogenetic analysis
(c) the analysis of secondary metabolites
(d) isozyme analysis
(e) DNA profiling techniques, including RFLP and RAPD.

5. A bibliography listing publications describing the application of these techniques for demonstrating distinctness in *Citrus* is included in Appendix I. Morphological, cytogenetic and secondary metabolite analysis of *Citrus* are not considered further as these techniques are beyond the scope of this paper.

6. Isozymes have proven to be a useful method for *Citrus* identification (see Appendix I, section 5). Torres *et al.* (1982), Ashari *et al.* (1989) and Zubrzycki *et al.* (1990) have investigate the relationships between *Citrus* types using isozymes. Ashari *et al.* (1989) used 19 loci representing 16 enzymes to examine 19 mandarin and hybrid cultivars. Of the 19 loci, 12 were polymorphic and these were able to distinguish among 16 of the 19 cultivars. Zubrzycki *et al.* (1990) used a combination of PAGE and isozymes to identify *Citrus* varieties but warned that a complementary system of morphometric markers and biochemical characters were required to differentiate effectively. However, while isozyme systems have proven utility, their use is restricted by the limited number of enzyme systems that can be visualized and by the possibility that isozyme expression can be influenced by environmental conditions or management practices (Zubrzycki *et al*, 1990).

7. DNA based markers offer a number of advantages over isozymes and other biochemical methods for identifying distinctness. Firstly, the DNA sequence of an organism is independent of environmental conditions or management practices. Secondly, the presence of the same DNA in every living cell of the plant allows tests to be conducted on any tissue at any stage of growth (provided that DNA of sufficient purity can be isolated). Thirdly, new DNA profiling techniques enable us, for the first time, to quickly and easily scan large sections of the genome in search for polymorphisms that can be used to demonstrate distinctness.

8. The first DNA profiling technique to be widely applied in the study of plant variation was the Restriction Fragment Length Polymorphism (RFLP) assay. The use of this technique in *Citrus* is limited. Nevertheless, studies by Roose (1988), Durham *et al.*(1990,1992), Komatsu *et al.* (1993), Liou (1990), Yamamoto *et al. and* Matsuyama *et al.* (1992) have shown the utility of RFLP's in *Citrus*. Unfortunately, more extensive use of this technique in *Citrus* is limited by the availability of probes which are time consuming and costly to obtain. In contrast, in better known species

such as *Zea mays*, *Triticum aestivum*, or *Glycine max*, a large number of DNA probes are available and as a consequence extensive DNA profiling with RFLP's is very feasible in these groups.

9. The recent development of the polymerase chain reaction (PCR) has enabled promising new DNA profiling techniques that are simpler and faster to perform, and often require less development time than RFLP analysis. Given these considerations, and that there are relatively few RFLP probes available for *Citrus*, we have chosen to use PCR based techniques in our studies of *Citrus*.

III. The Polymerase Chain Reaction (PCR)

10. The Polymerase Chain Reaction was invented by Kary B. Mullis in 1985 (Saiki *et* al., 1985) and in a few years has revolutionised many areas of biological science. PCR can specifically amplify a single region of DNA in a complex genome, or it can be used to scan a genome for polymorphisms. Both of these approaches are relevant to the application of PCR as a tool for varietal identification. A brief background to the PCR process is necessary before the strengths and weaknesses of particular applications can be assessed.

11. PCR relies on the use of a specific class of enzymes, DNA polymerases, which all living cells possess and use to copy their own DNA. DNA polymerases copy single stranded DNA from the 3'OH end of double stranded DNA (Figure 1).

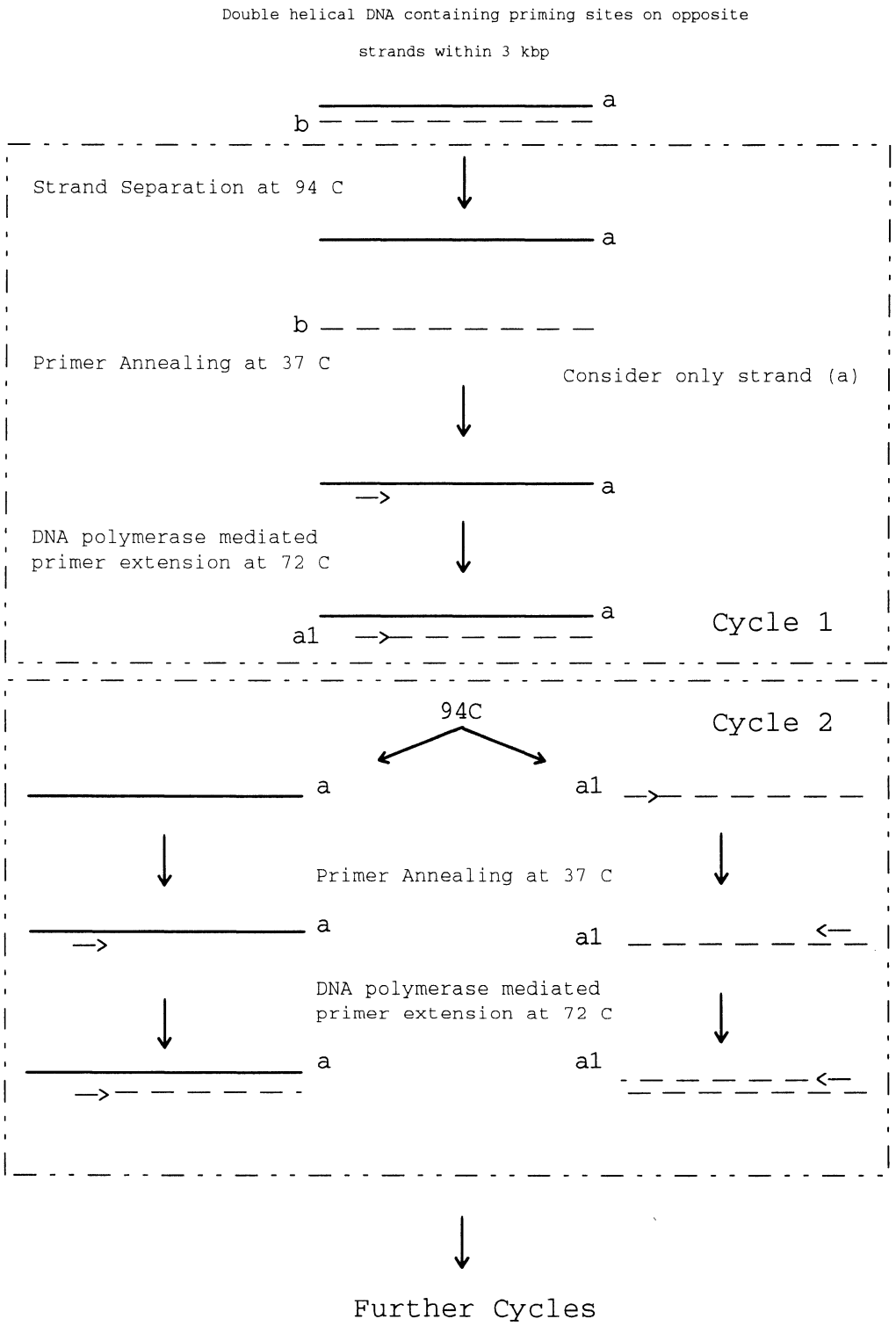
12. In PCR, these conditions are established by first heating the DNA to separate the double stranded DNA into single stranded molecules. Next, the temperature is lowered to allow short DNA molecules called 'primers' (typically of 8 to 20 base pairs in length) to 'anneal' to their complementary strand. These double-stranded complexes serve as starting points for the copying of single stranded DNA by the polymerase. By flanking a region of DNA with specific DNA primers and cycling the temperature to facilitate strand separation, primer annealing and primer extension, the PCR reaction can make billions of copies of DNA. The use of heat stable DNA polymerases which survive the lengthy exposure to high temperatures required by PCR, and the development of thermocyclers capable of cycling temperatures quickly and accurately have facilitated the automation of this process. Today PCR is used extensively world wide in many area's of biology and medicine.

BMT/2/3
page 5

Figure 1

The Polymerase Chain Reaction

DNA Amplification Using a Single Primer



13. While there are many applications of PCR in biological and medical research they can be broadly classified as two different approaches. In the first approach two different specific DNA primers that are complementary to opposite strands of conserved DNA sequence are used to amplify the region of DNA between the primers. Specific PCR is sometimes referred to as sequence-tagged-site PCR (STS) and we will use this acronym for convenience in this paper. One class of STS marker that promise to be particularly valuable for cultivar identification are sequence-tagged microsatellites (STMS) also known as simple sequence length polymorphisms (SSLP) (Thomas and Scott 1993). Microsatellites are highly polymorphic and abundant simple sequence repeats that can be detected by STS-PCR using specific primers that match the flanking microsatellite region (Morgante and Olivieri 1993; Goodfellow 1993). Although studies of microsatellite variation in plants have only just begun (e.g. Thomas and Scott 1993), the available evidence indicates that, as in animals, plant microsatellite loci can exhibit many alleles and are highly heterozygous. These features make microsatellites obvious targets for DNA profiling. However, because a knowledge of suitable DNA sequences from the organism to be studied, or a related species is required, the widespread application of STS-PCR in plants has been limited by the relatively few DNA sequences known. STS-microsatellite markers are presently being developed for *Citrus* and this will be described briefly later in the paper.

14. In 1990, Williams et al. (1990) and Welsh and McClelland (1990) developed the second approach to PCR. In this approach, short DNA primers, of known sequence, but chosen on an arbitrary basis, are used to amplify those regions of the genome where the primer(s) bind sufficiently close on opposite strands to allow amplification of the intervening DNA (Tingey and Del Tufo 1993; Newbury and Ford-Lloyd 1993). The advantage of this approach is that no prior knowledge of the DNA sequence is required. In addition, with appropriate primers the procedure produces polymorphic DNA profiles among and within species. Williams et al. (1990) proposed the term random amplified polymorphic DNA (RAPD) for this class of genetic marker, and the acronym is now widely used, although others exist for related protocols: AP-PCR for arbitrarily primed PCR (Welsh and McClelland 1990), DAF for DNA amplification fingerprinting (Caetano-Anolles et al. 1991; SS-RAPDs for silver-stained RAPDS (Huff and Barra 1993), and RAPD-DGGE for RAPDs separated via density-gradient-gel electrophoresis (Dweikat et al. 1993). We will use the acronym RAPD in this paper to include all of these protocols. However, one disadvantage of using this acronym is that it can be misleading in two ways. First it implies that primers are chosen randomly. This is not the case-primers are of arbitrary sequence, but are carefully chosen to have a 50% or greater GC content with no internal complementarity. Second, the acronym may be taken to imply that the technique is random, unreproducible and unreliable. As we will show below, this also is not the case. In this paper, we will illustrate the utility of the RAPD procedure for identifying *Citrus* varieties.

IV. The use of RAPDs to identify distinctness in *Citrus* varieties

15. The most important commercial cultivars of mandarins in Australia are Imperial, Ellendale and Murcott. Lesser cultivars include Hickson and Glen Retreat. These varieties have formed the foundation of mandarin hybrid breeding in Australia although a range of minor cultivars have more recently been included in breeding programs. In this section we describe our study of RAPD variation for all of the major Australian varieties and two new varieties; Monarch (Imperial x Hickson) and the triploid Eloise (Imperial x Murcott). This study was undertaken at the Co-operative Research Centre for Plant Science and the CSIRO Division of Plant Industry in Canberra, Australia and represents an extension of the study described by Preston *et al.* (1993).

Isolation and purification of Citrus DNA

16. The goal of the DNA isolation is to consistently yield DNA of sufficient quality for reproducible PCR analysis. We homogenised 1-2g of leaf tissue by grinding in liquid nitrogen and then added an extraction buffer containing detergents that dissolve membranes and denature proteins, and EDTA that binds to the metal cofactors of nucleases. This extract was treated with phenol and chloroform to remove excess proteins prior to precipitation of the DNA from the aqueous phase with ethanol. Contaminating RNA and proteins were subsequently removed with RNA and protein-degrading enzymes. The quantity and purity of the DNA was assessed in a spectrophotometer and by gel electrophoresis. Two protocols for DNA extraction that we found worked well for *Citrus* are given in the Appendix II.

PCR reaction conditions

17.The critical parameters in the PCR reaction are the sequence of the DNA primers, the template DNA, and the temperature profile maintained during the reaction. Other components such as the buffer, magnesium, the deoxynucleotide triphosphate concentration and the source of the DNA polymerase are also important variables which also need to be defined and controlled. Using the procedures of Weeden *et al*, (1992) and Tingey *et al.*(1992) as a basis, we performed titrations for these variables to identify the optimal conditions for amplification of *Citrus* DNA. We found reaction mixtures which contained 1 to 5 ng Mandarin DNA per 10 ul reaction, 1.75 to 2 mM $MgCl_2$, 0.3 to 0.7 M primer, 67 mM Tris-HCl (pH 8.8), 0.2 mM deoxynucleotide triphosphates and a number of other additives (outlined in Appendix III) gave reliable, reproducible results using the PCR protocol described in Appendix III. These conditions are similar to those empirically determined by other workers (eg: Tingey *et al.*, 1992; Rafalski et al., 1993; Weeden *et al.* 1992; Ellsworth *et al.* 1993).

18. After optimising the PCR reaction conditions, we screened 18 primers from the Operon Set A (Operon Technologies Inc, Alameda CA, USA) using agarose gel electrophoresis and ethidium bromide to visualize PCR bands under UV light. Ten primers were selected from

the set of 18 because they yielded clear, interpretable and reproducible bands.

19. An important consideration with RAPDs is reproducibility. We found these 10 primers gave reproducible patterns: (1) for DNA extracted by 3 different methods, (2) among multiple PCR reactions (at least 5), (3) between two operators working in separate laboratories in different organisations. Furthermore, the results were consistent between two collections of material obtained from different collections.

Data analysis and results

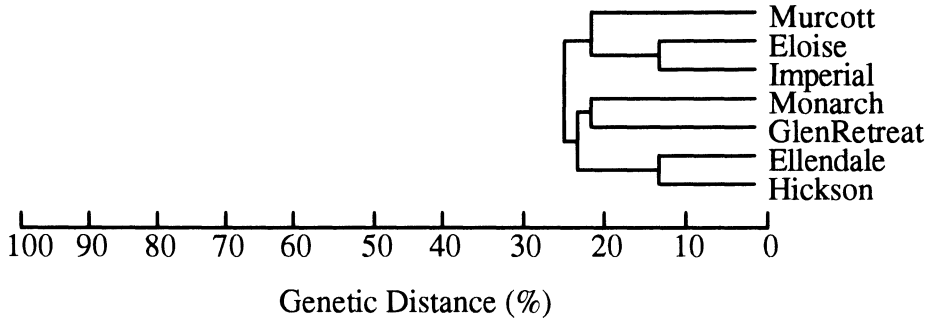
20. We scored the DNA profile manually, directly from photographs of the gels by assigning a value of 1 for band presence and a value of 0 for band absence. Thus for each sample, we converted the DNA profile into a binary vector (eg: 10111001). Next the vectors of band presence/absence were used to calculate a pairwise genetic distance matrix. This matrix was produced by comparing the vector of a given individual with that of every other individual and calculating a distance value then repeating the process for all individuals. Genetic distances were calculated by the formula of Nei and Li (1979):

$$1-F = \left[1 - \frac{2n_{xy}}{n_x + n_y} \right]$$

where $2n_{xy}$ equals the number of shared bands and n_x and n_y is the number of bands observed in individual x and individual y respectively. Finally, to visualize the patterns genetic relationship, we prepared a phenogram based on a UPGM cluster analysis of the genetic distance data.

21. Ninety one bands were revealed by the 10 primers, of which 51 were polymorphic. The pairwise genetic distances between varieties ranged from 13% for Hickson and Ellendale to 34% for Eloise and Hickson. Thus all varieties were identified by a unique fingerprint with the minimum number of band difference between varieties being 12. Furthermore, the use of just two primers, A3 and A4, generating 18 bands (11 polymorphic) produced banding patterns which could unambiguously identify all of the cultivars. The genetic relationships revealed by cluster analysis of the RAPD data are shown in the figure below. The grouping for Murcott, Eloise, and Imperial reflects their parentage, with the triploid Eloise the progeny of a cross between Murcott and Imperial.

Phenogram of Mandarin Varieties



22. In a similar study, Omura *et al* (1993) used RAPDs to examine genetic relationships among 16 mandarin species and cultivars from Japan, China, Algeria, India and Indonesia that show little or no difference in isozymes. In contrast to isozymes, the RAPD primers yielded polymorphisms which identified all 16 species. In addition, it was possible to identify clones within the species *C. leiocarpa, C. tachibana* and *C. clementina*. However, among seven cultivars of satsuma mandarin, no polymorphisms were detected. The authors were also able to confirm that the RAPD markers were: (1) inherited as Mendelian traits, (2) stable within clones and, (3) reproducible across seasons.

23. Carstens (1993) conducted a preliminary study to examine the feasibility of using RAPD as a method for discriminating 18 *Citrus* cultivars in South Africa. Fifteen primers were screened and most revealed polymorphisms between cultivars with one highly variable primer producing distinct profiles for all the cultivars tested. The study also examined navel oranges but found little differentiation among the cultivars, probably because most navel cultivars originate as somatic mutations.

24. Information about *Citrus* cultivars obtained using the RAPD technique would be acceptable under the Australian Plant Variety Rights Act provided that the applicant furnished comparative data on the claimed novel variety and the most similar varieties of common knowledge. The latter is a standard requirement for all morphological, biochemical or DNA-based methods to satisfy the DUS criteria. In the Australian PVR system, this comparative DNA data would currently be accepted with the status of supplementary data presented alongside a case based on the principle distinguishing morphological characteristics for *Citrus*.

V. The use of RAPDs to identify varieties in other plant groups.

25. We have shown that RAPDs can distinguish among varieties of *Citrus* effectively and reliably and it is now appropriate to compare our findings with those for other plant groups. Table 1 summarises the outcome of RAPD studies for a range of cultivars from *Allium* to Wheat. It is important to note that in almost all cases, RAPDs were able to distinguish among varieties including closely related varieties of wheat presumed to share over 94% of their genomes (Dweikat et al. 1993). Furthermore, all of the authors reported that the relationships revealed by RAPD analysis were generally consistent

with other types of evidence such as allozyme or protein markers, validating the value of the RAPD technique. However, in several plant groups very few genetic differences were detected among some varieties (Table 1). For example, in Celery one cultivar only differed from another by a single band differences, despite more than 300 bands being scored and in Brassica, 1-10 bands distinguished varieties although in this case only 40 bands were scored.

Table 1

Plant group	# var.	# primers	# bands	Genetic distance range[a]	# distinct bands[b]	GD[c]	Reference
Allium cepa	7	6 (6)	91	0.01-0.09	1-?	4	Wilkie et al. 1993
Barley	17	8	?	0.02-0.22	?	1*	Francisco-Ortega et al. 1993
Brassica oleracea var. capitata	8	25	200	0.06-0.20	7-24	1	Kresovich et al. 1992
B. oleracea var. costata	4	25	200	0.18-0.35	20-40	1	Kresovich et al. 1992
Broccoli	14	4 (4)	37	0.025-0.25	1-10	3	Hu and Quiros 1991
Cauliflower	12	4 (4)	40	0.025-0.40	1-18	3	Hu and Quiros 1991
Celery	19	28 (9)	309	0.035-0.48	1-13	3	Yang and Quiros 1993
Oats	17	8	?	0.04-0.18	?	1*	Francisco-Ortega et al. 1993
Papaya	10	11 (?)	102	0.05-0.30	5-	2*	Stiles et al. 1993
Theobroma cacoa	10	8	48	0.14-0.49	6	1	Wilde et al. 1992
Theobroma cacoa	25	9	75	0.05-0.35	3-48	1	Russell et al. 1993
Wheat	16	8	?	0.01-0.15	1-?	1*	Francisco-Ortega et al. 1993

[a]Genetic distance range for all pairwise comparisons of varieties.
[b]Range of band differences among varieties calculated as: minimum genetic distance x # variable bands. Note that some authors provided measures of genetic similarity rather than distance which have been converted to genetic distance by subtraction from 1.
[c]Minimum number of primers required to distinguish all cultivars.
[d]Genetic distance: 1=1-F of Nei and Li (1983), 2=Jaccard coefficient. 3=pairwise distance matrix as calculated by PAUP. 4=Rogers Genetic distance. *=converted to percentage genetic distance.

26. Although the data in Table 1 provides some clues, the important question 'How many RAPD primers and how many bands are needed to distinguish varieties?', is difficult to answer. Clearly, the answer will vary from group to group and will need to be empirically determined. Kresovich et al. (1992) suggest that for *Brassica* there are limited returns after screening 10-12 primers or 50-60 bands. Furthermore, they concluded that one well selected primer and/or 5-25 randomly chosen fragments would be sufficient to discriminate among varieties. Demeke et al. (1992) in another study of *Brassica* concluded that at least 10 primers and approximately 100 bands are needed, corroborating the suggestion of Kresovich et al. (1992). On

the other hand, in groups with a narrow genetic base, a larger number of primers than usually required, may need to be screened in order to find differences among very closely related varieties. Alternatively, procedures that improve the resolution of the DNA profile such as polyacrylamide electrophoresis and silver staining (DAF: Caetano-anolles et al. 1990) or denaturing-gradient-polyacrylamide electrophoresis (RAPD-DGGE: Dweikat et al. 1993) may be necessary to uncover suitable polymorphisms. However, if the genetic basis for varietal differences merely reflects some point mutation(s) (as likely in some Navel oranges), genetic differences may be very difficult to find regardless of the DNA profiling technique employed. However, RAPDs presently offer the greatest chance of detecting small genetic differences, since a larger component of the genome can be scanned than in other systems. Indeed, the extensive coverage of the genome by RAPD primers has been used to find gene loci of interest by bulk segregant analysis (Michelmore et al. 1991).

27. In summary, our findings on *Citrus* and those for other plant groups point to the usefulness of RAPDs as supplementary evidence for varietal identification. But we add a cautionary note that DNA profiling in general may fail to detect differences among very closely related varieties. Furthermore, we will argue in a later section that despite the appeal and simplicity of RAPDs, there are some limitations that may prevent RAPD data being used as primary DUS data.

VI. The use of STS to identify varieties of *Citrus*.

28. The CSIRO Division of Horticulture in Adelaide, Australia has developed and patented an STS based methodology for the reliable identification of grapevine varieties (Scott et al 1992; Thomas et al 1993a,b; Thomas et al 1994). More recently, the Flinders University of South Australia has mounted a collaborative study with the CSIRO to develop a suitable technique for *Citrus*. This work has concentrated on the rapid isolation of microsatellite sequences suitable for use in STSs. It is the isolation and identification of suitable STS sequences which makes this method laborious compared with the RAPD procedure. An isolation technique has been developed (Kijas et al. 1994) in which biotinylated oligonucleotides are bound to the fragmented Citrus genome and those fragments containing microsatellite sequences are collected using streptaviden coated magnetic particles thus resulting in an enrichment of microsatellite sequences suitable for cloning, sequencing and use as STSs.

29. Microsatellites isolated in this way are now being tested in *Citrus* populations for use both as markers in *Citrus* genotype mapping and for identifying individual *Citrus* genotypes. Although data for *Citrus* are not yet available, the effectiveness of the STS method can be judged by reference to the grapevine work cited above. With four or five microsatellite based STS markers all grapevine varieties tested so far can be separated (Thomas et al 1993). This has enabled a data base of grapevine STS DNA types to be established such that individual grapevine DNA tests can be referred to the data base to allow direct identification of any previously established grapevine variety for which a DNA type has been accredited in the data base.

VII. General Considerations

30. The choice of the appropriate DNA profiling technology is dependent on the aims of the testing being undertaken. Under UPOV, the formal obligation is to demonstrate the distinctness, stability and uniformity of a plant variety. The practical considerations are that the test must also be inexpensive, technically straightforward, reliable, reproducible and capable of unambiguous analysis. The cost of the development and conducting of the test must also be justified by the economic importance of the species or variety. In this section we will consider the general application and role of DNA profiling for identifying varieties including some discussion on the advantages and limitations of each technology. First we consider the question of which DNA profiling technology: RFLP vs PCR?. Next we compare RAPD and STS markers. Finally we conclude with recommendations for UPOV's consideration. A detailed discussion on data collection and analysis, which we believe is also of general interest to UPOV, are presented in Appendices IV and V.

RFLP vs PCR

31. Several reviews comparing PCR and RFLP procedures for detecting genetic polymorphism in both animals and plants are available (Arnheim 1990; Weber 1990; Tingey and del Tufo 1993 etc.). In these reviews there is general agreement that PCR based DNA profiling offers a number of important advantages over DNA profiling with RFLP's including:

- PCR only requires small amounts of DNA and often crude miniprep procedures yield DNA of sufficient quantity and quality.
- The PCR process involves fewer steps than RFLP's and is therefore faster to perform.
- PCR is technically straightforward once the PCR conditions have been established.
- PCR does require the use of radioactivity to visualize polymorphisms.
- PCR can be readily automated at all stages from DNA extraction to data collection and analysis.
- The vast range of potential primer sequences that can be used gives the technique great diagnostic power.

32. Given these advantages, it is generally concluded that DNA profiling via PCR will also be much cheaper than via RFLP's. However, careful comparisons of the costs of RFLPs vs RAPDs by Ragot and Hoisington (1993) suggest while RAPDs are clearly cheaper for small sample sizes, RFLP analysis becomes cost effective for larger sample sizes. However, these calculations did not include the cost of developing RFLP probes. Similarly, cost estimates for STSs must include the cost of developing suitable specific primers. Thus any comparisons of cost among PCR and RFLP must take into account the type of PCR based marker to be used, costs of developing the technique and the sample sizes to be screened. While considerations of cost may not always weigh in favour of PCR we believe that collectively, PCR based DNA profiling will prove to be the most

efficient and probably the most cost effective technology to develop in plant groups where RFLP assays are not already established. On the other hand, when RFLP analysis is well established this technology remains a viable alternative. However, even in these cases it may ultimately prove worthwhile to turn useful RFLP polymorphisms into PCR assays for routine screening.

Which PCR marker?

33. In preceding sections we have shown that both RAPD and STS markers can be used to assess varietal distinctiveness. However, both types of assay offer advantages and disadvantages that need to be considered when choosing a PCR based to use for plant varietal registration.

RAPD's

34. As noted earlier, because specific DNA sequence information is not required, RAPDs enable the study of anonymous genomes with PCR. However, one of the inevitable trade offs with the RAPD technique is that amplification is performed under conditions of low stringency. Consequently, some of the products formed are the result of mismatching and this can result in poor reproducibility for some primers and bands. Nevertheless, our study in *Citrus* and that of many others researchers have shown that reproducible RAPD bands can be found by a careful selection of primers, optimisation of PCR conditions for the target species and replication to ensure that only reproducible bands are scored.

35. Another problem reported for RAPDs, is a low incidence of non-inherited bands which are probably PCR artefacts. For example, Heun and Helentjaris (1993) report this problem for a small percentage of RAPD bands in maize and note that similar patterns have been found in other plants. While the great majority of RAPD bands are known be inherited as mendelian markers, nonetheless, it is apparent that care needs to be taken when drawing conclusions based on a small number of band differences.

36. The problem of reproducibility amongst laboratories is another issue of relevance to RAPDs. Penner et al. (1993) have experimentally compared RAPD results among laboratories for two oat cultivars. Some differences in DNA profiles were found among labs, but it was concluded that if the overall temperature profiles of the PCR reactions are identical among laboratories then RAPD fragments are reproducible for appropriately chosen primers. Thus, some differences can be expected among laboratories using different thermocylers. However, for the purposes of varietal registration all comparisons should be made within laboratories regardless of the DNA profiling technique.

37. One limitation of the RAPD technique that cannot be overcome is that of band dominance. This means that heterozygotes are rarely detected and there are usually only two states for a polymorphism, either present or absent. As a consequence, RAPD and other

multilocus profiles provide less genetic data than for single locus codominant markers such as STS.

38. In summary, we believe that many of the limitations of RAPDs can be overcome. However, no matter how careful the study, in cases of litigation, it is possible that the potential for these problems to exist may be used as justification for disregarding the genetic data. Therefore, while RAPDs can provide valuable supplementary evidence, they are unlikely to be suitable as a primary DUS characteristic.

STS via RAPDs

39. STS markers overcome many of the inherent limitations of RAPDs, providing highly reproducible and informative, single locus codominant markers. An important consideration in favour of codominant STS loci over RAPDs is that a positive result, that is a DNA band, is always achieved in a successful STS reaction. This control is not possible in all RAPD reactions. The disadvantage of STS is that the development of specific primers is a costly and time consuming process, albeit using standard procedures. Furthermore, primers may not be transferable to other species or genera. However, once suitable primers are available, specific PCR markers can be screened as easily as RAPDs.

40. One way to speed the development of these specific PCR assays is to use RAPDs as a means of identifying useful polymorphisms. Subsequently, the RAPD fragment(s) of interest are sequenced and specific primers are constructed to assay a specific polymorphism sometimes called a 'sequence characterized amplified region' (SCAR). This and other related assays have recently been reviewed by Rafalski and Tingey (1993). SCARS can be further assayed for variation with restriction enzymes to reveal RFLP's. This approach has been called cleaved amplified polymorphic sequence (CAPS). Another related method is allele specific PCR (AS-PCR), where PCR amplification of specific alleles at a locus is made possible by designing primers that partially overlap the allelic sequence differences.

STS Microsatellites

41. If microsatellites are found to be as ubiquitous, polymorphic and heterozygous as predicted, STS microsatellite markers may prove to be the DNA profiling method of choice for primary DUS criteria. As we have revealed in this paper, new methods are now available for minimizing the development time and undoubtedly other time saving procedures will be emerge in the near future. For important crops, it is likely that the cost of developing microsatellite primers will ultimately be outweighed by the high information content and reliability of this marker system.

42. In conclusion, recent technological developments have provided for the first time a series of markers that can be used to demonstrate distinctiveness, stability and uniformity of a plant variety. We have shown that RAPDs provide accessible and relatively inexpensive supplementary data and are useful entry point for plant groups that are poorly known at the DNA level. The conversion of

RAPDs into STSs or the alternative development of STS markers such as microsatellites provide robust markers that may ultimately prove useful as primary data. The development of STS markers should be encouraged where the cost of the development and conducting of the tests can be justified by the economic importance of the species or variety.

VIII. Recommendations

43. Recommendations

- that UPOV continue to consider DNA profiling methods for *Citrus* for use as supplementary DUS information.
- that the RAPD technique provides information which is of a standard for member countries to treat as important supplementary DUS data.
- that the Sequence-Tagged-Sites methodology also provides information which is of a standard for member countries to treat as supplementary DUS information and holds the promise, following further development and testing, of providing data which UPOV may consider in the future to satisfy the requirements for acceptance as a principle DUS character.
- that UPOV consider the question of the statistical treatment of DNA data and standardisation of the procedures.
- that UPOV continue to explore the potential that DNA profiling offers to assist UPOV to develop more rigorous DUS test guidelines.

Appendix I

Bibliography

1. General References

Alexander, DM. (1983) Some *Citrus* species and varieties in Australia. CSIRO, Melbourne,

Arnheim N, White T, Rainey WE (1990) Application of PCR: organismal and population biology. Bioscience 40:174-182.

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and Sensitive Silver Staining of DNA in Polyacrylamide Gels. Anal Biochem 196: 80-83

Bowman, FT. (1956) *Citrus*-growing in Australia. Angus and Robertson, Sydney,

Caetano-Anolles G, Bassam BJ, Gresshoff PM (1991) DNA Amplification fingerprinting: A strategy for genome analysis.. Plant Mol Biol Rep 9: 294-307

Citrus breeding workshop. Proceedings of a workshop held at Merbein, Victoria, 27-29 July 1987 on *Citrus* breeding in Australia and the use of new breeding technologies. CSIRO, (1988) Melbourne, Australia, pp 1-153

Dawes SN, Martin PJ (1989) Early maturing satsuma mandarins show promise. Orchardist of New Zealand 62: 11-12

Dawes SN, Martin PJ (1989) New mandarins appear on the horizon. Orchardist of New Zealand 62: 20-21

Dawes SN, Martin PJ (1990) Parent Navel beats other orange selections. Orchardist of New Zealand 63: 16-17

Demeke T, Adams RP, Chibbar R (1992) Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study in Brassica. Theor Appl Genet 84: 990-994

Dweikat I, Mackenzie S, Levy M, Ohm H (1993) Pedigree assessment using RAPD-DGGE in cereal crop species. Theor Appl Genet 85: 497-505

Ellsworth DL, Rittenhouse KD, Honeycutt RL (1993) Artifactual Variation in Randomly Amplified Polymorphic DNA Banding Patterns. BioTechniques 14: 214-217

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA halotypes: application to human mitochondrial DNA restriction sites.. Genetics 131: 479-491

Francisco-Ortega J, Newbury HJ, Ford-Lloyd BV (1993) Numerical analyses of RAPD data highlight the origin of cultivated tagasaste (Chamaecytisus proliferus ssp.palmensis) in the Canary Islands. Theor. Appl. Genet. 87:264-270.

Goodfellow PN (1993) Viewpoint: Microsatellites and the new genetic maps. Current Biology 3: 149-151

Gresshoff PM, Sayavedra-Soto L, Landrau-Ellis D, Culpepper J, Caetano-Anolles G, Bassam BJ (1990) DNA Fingerprinting - The Powerful New Tool for the Genetic Detective. Tennessee Farm and Home Science 155: 4-10

Grosser JW, Gmitter FG, Sesto F, Deng XX, Chandler JL (1992) Six new somatic *Citrus* hybrids and their potential for cultivar improvement. J Am Soc Hortic Sci 117: 169-173

Handa T, Ishizawa Y, Oogaki C (1986) Phylogenetic study of Fraction I protein in the genus *Citrus* and its closely related genera. Japan J Breed 61: 15-24

Heun M, Helentjaris T (1993) Inheritance of RAPDs in F1 hybrids of corn. Theor Appl Genet 85: 961-968

Hillis, DM., Larson, A., Davis, SK., and Zimmer, EA. (1990). Protocol 2: DNA isolation from plants, fungi, and algae. Chapter 9. Nucleic Acids III: Sequencing. In 'Molecular Systematics' (Ed. Hillis, D.M. and Moritz, C.) pp 340-341. Sineauer Associates Inc. Massachusetts.

Hirai M, Mitsue S, Kita K, Kajiura I (1990) A survey and isozyme analysis of wild mandarin, tachibana (*Citrus* tachibana (Mak.) Tanaka) growing in Japan. Journal of the Japanese Society for Horticultural Science 59: 1-7

Hodgson RW (1967) Horticultural varieties of *Citrus*. In CA Berkeley ed, The *Citrus* industry, University of California Press, pp 431-591

Hu J, Quiros CF (1991) Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Reports 10: 505-511

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing Buffalograss (Buchloe dactyloides (Nutt. Engelm.). Theor. Appl. Genet. 86:927-934.

Jorgensen KR (1980) *Citrus* breeding, taxonomy and the species problem. In Anonymous ed, Proceedings of the International Society of Citriculture, Griffith, Australia, 1978, Griffith, Australia, pp 51-57

Kijas, J.M.H. , Fowler, J.C.S.&Thomas, M.R. 1994. Enrichment of microsateelite from the Citrus genome using biotynilated oligonucleotides bound to Strepavidin coated magnetic particles.Biotechniques, in press.Adams RP, Demeke T, Abulfatih HA (1993) RAPD DNA fingerprints and terpenoids; clues to past migrations of Juniperus in Arabia and east Africa. Theor. Appl. Genet. 87:22-26.

Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992) Characterisation of genetic identities and relationships of Brassica oleracea L. via a random amplified polymorphic DNA assay. Theor Appl Genet 85: 190-196

Loukas M (1984) Phylogenetic relationships among *Citrus* species. Georgike Ereuna 8: 263-277

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis : A rapid method to detect markers in specific genomic regions by using segregating populations. PNAS 88: 9828-9832

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3: 175-182

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

Nei M (1972) Genetic distance between populations. Amer. Nat. 106:283-392.

Nei M (1977) F-Statistics and Analysis of Gene Diversity in Subdivided Populations. Ann. Human Genet. 41:225-233.

Newbury HJ, Ford-Lloyd BV (1993) The use of RAPD for assessing variation in plants. Plant Growth Reg 12: 43-51

Penner GA, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles G, Motnar SJ, Fedak G (1993) Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. PCR Methods and Applications 2:341-353.

Penner GA, Chong J, Levesque-Lemay M, Molnar SJ, Fedak G (1993) Identification of a RAPD marker linked to the oat stem rust gene Pg3. Theor Appl Genet 85: 702-705

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. TIG 9:275.

Ragot M, Hoisinton DA (1993) Molecular markers for plant breeding comparisons of RFLP and RAPD genotyping costs. Theor. Appl. Genet. 86:975-984.

Russell J , Hosein F, Johnson E, Waugh R, Powell W (1993) Genetic differentiation of cocoa (Theobroma cacao L.) populations revealed by RAPD analysis.. Molecular Ecology 2: 89-97

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic Amplification of b-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anaemia. Science 230: 1350-1354

Saunt, J (1990) *Citrus* varieties of the world. An illustrated guide. Sinclair International Ltd, Norwich, UK, pp 1-126

Scora RW (1989) Biochemistry, taxonomy and evolution of modern cultivated *Citrus*. In R Goren, K Mendel, eds, Citriculture : proceedings of the Sixth International *Citrus* Congress, Middle-East, Tel Aviv, Israel, March 6-11, 1988, Balaban Publishers, Rehovot, Israel, pp 277-289

Scott NS, Thomas MR (1992) DNA Fingerprinting Aust Prov Patent Application No PL 3330/92, PCT lodged July 1993

Scott NS, Thomas MR (1992). DNA Fingerprinting and the Identification of Grapevines. Wine 7, 221-223.

Sedgley M, Ashari S (1990) Biochemistry, taxonomy and evolution of modern cultivated *Citrus*. In Anonymous ed, Laboratory Identification of Plant Varieties.: Proceedings of Workshop, Burnley Plant Research Institute, 17 May 1990, Plant Variety Rights Office, Canberra, Australia,

Stiles JI, Lemme C, Sondur S, Morshidi MB, Manshardt R (1993) Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85: 697-701

Thomas MR, Matsumoto S, Cain P, Scott NS (1993). Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. Theoret.& Appl. Gen. 86, 173-180

Thomas MR, Scott NS (1993). Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequenced tagged sites (STSs). Theoret.& Appl. Gen. 86:985 -990.

Thomas MR, Cain P, Scott NS. (1994). DNA typing of grapevines: A methodology and database for describing cultivars and evaluating genetic relationships. Plant Mol Biol: in press.

Tingey SV, Del Tufo JP (1993) Genetic analysis with random amplified polymorphic DNA markers. Plant Physiol 101: 349-352

Tingey SV, Rafalski JA, Williams JGK (1992) Genetic Analysis with RAPD Markers. In M Neff ed, Applications of RAPD Technology to Plant Breeding, American Society for Horticultural Science, Minneapolis, Minnesota, pp 3-8

Weber J (1990) Informativeness of Human (dC-dA)n . (dG-dT)n Polymorphisms.. Genomics 7: 524-530

Weeden NF, Timmerman GM, Hemmat M, Kneen BE, Lodhi MA (1992) Inheritance and Reliability of RAPD Markers. In M Neff ed, Applications of RAPD Technology to Plant Breeding, American Society for Horticultural Science, Minneapolis, Minnesota, pp 12-17

Welsh J, McClelland M (1990) Fingerprinting Genomes using PCR with Arbitrary Primers. Nucleic Acids Research 18: 7213-7218

Wilde J, Waugh R, Powell W (1993) Genetic Fingerprinting of Theobroma clones using randomly amplified polymorphic DNA markers. Theor Appl Genet 83: 871-877

Wilkie SE, Isaac PG, Slater RJ (1993) Random amplified polymorphic DNA (RAPD) markers for genetic analysis in Allium. Theor. Appl. Genet. 86:497-504.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18: 6531-6535

Wright (1951) The genetical structure of populations. Ann. Eugenics 15:=323-354.

Yang X, Quiros C (1993) Identification and classification of celery cultivars with RAPD markers. Theor Appl Genet 86: 205-212

Zhang WC, Shao ZY, Lo JH, Deng CH, Deng SS, Wang F (1989) Investigation and utilisation of *Citrus* varietal resources in China. In R Goren, K Mendel, eds, Citriculture : proceedings of the Sixth International *Citrus* Congress, Middle-East, Tel Aviv, Israel, March 6-11, 1988. Volume, Balaban Publishers, Rehovot, Israel, pp 291-294

2. Morphology

Barrett HC, Rhodes AM (1976) A numerical taxonomic study of affinity relationships in cultivated *Citrus* and its close relatives. Syst Bot 2: 105-136

Calvarano M, Houjiu WU, Calvarano I, Crescimanno FG, Germana MA, Chironi G (1991) Research on Morphological and Biochemical Fruit Characteristics of 12 *Citrus* Aurantium L. Clones. In H Bangyan, H Huibai, H Mingdu, K Chunyen, T Xingjie, Y Qian, S Guifen, eds, Proceedings of the International *Citrus* Symposium, Guangzhou, China, November 5-8, 1990, International Academic Publishers, Guangzhou, China, pp 837-843

Cottin R (1988) Numerical taxonomy. Application to *Citrus*. Fruits Paris 43: 721-733

Dawes SN, Martin PJ (1990) Comparison of four early-maturing satsuma mandarins with the standard industry cultivar 'Silverhill'. New Zealand Journal of Crop and Horticultural Science 18: 23-29

Development of new technology for identification and classification of tree crops and ornamentals. (1986) Fruit Tree Research Station, Ministry of Agriculture, Forestry and Fisheries, Ibaraki-Ken, Japan, pp 1-116

Geraci G, Manzocchi LA, Tusa N, Occorso G, Radogna L, Pasquale F (1982) Comparison of different methods for identifying zygotic and nucellar seedling in *Citrus*. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 1-4

Germana MA, Crescimanno FG, Pasquale F, Ying WY (1992) Androgenesis in 5 cultivars of *Citrus* limon L. Burm. f. Acta Hortic 300: 315-324

Gogorcena Y, Ortiz JM (1989) Morphometric and biochemical characterisation of Spanish sour orange cultivars. In R Goren, K Mendel, eds, Citriculture : proceedings of the Sixth International *Citrus* Congress, Middle-East, Tel Aviv, Israel, March 6-11, 1988, Balaban Publishers, Rehovot, Israel, pp 215-223

Gogorcena Y, Ortiz JM (1989) Characterisation of sour orange (*Citrus* aurantium) cultivars. Journal of the Science of Food and Agriculture 48: 275-284

Handa T, Oogaki C (1985) Numerical taxonomic study of *Citrus* and Fortunella using morphological characters. J Jap Soc Hortic Sci 54: 145-154

Kozaki I, Hirai M (1986) Pollen ultrastructure of *Citrus* for taxonomic identification. In K Kitaura, et al, eds, Development of new technology for identification and classification of tree crops and ornamentals, pp 11-17

Kozaki I, Hirai M (1991) Pollen ultrastructure of *Citrus* cultivars. In H Bangyan, H Huibai, H Mingdu, K Chunyen, T Xingjie, Y Qian, S Guifen, eds, Proceedings of the International *Citrus* Symposium, Guangzhou, China, November 5-8, 1990, International Academic Publishers, Guangzhou, China, pp 19-22

Moore GA, Castle WS (1988) Morphological and isozymic analysis of open-pollinated *Citrus* rootstock populations. Journal of Heredity 79: 59-63

Reforgiato-Recupero G, Starrantino A (1982) Characterisation with SEM of the pollen of some Italian lemon cultivars. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 17-19

Richardson A, Anderson P, Harty A, Sutton P, Machin T (1991) Satsumas - ten cultivars compared. Orchardist of New Zealand 64: 22-25

Tisserat B, Galletta P, Jones D (1990) Carpel Polymorphism in *Citrus* Fruit. Bot Gaz Chicago 151: 54-63

Ye YM, Kong Y, Zheng XH (1982) Studies on Pollen Morphology of *Citrus* Plants. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 23-25

3. Cytogenetic Analysis

Guerra M (1993) Cytogenetics of Rutaceae. V. High chromosomal variability in *Citrus* species revealed by CMA/DAPI staining. Heredity 71: 234-241

4. The Analysis of Secondary Metabolites

Abkhazava DM, Kharebava LG (1989) The essential oil composition of grapefruit fruits of different cultivars. Subtropicheskie Kul'tury No. 6, 111-117

Abkhazava DM, Tsilosani MV (1990) Studies on the biochemical properties of the fruits of different grapefruit cultivars. Subtropicheskie Kul'tury 2: 111-113

Bade JB, Gmitter FG, Jr., Bowman KD (1991) Simplified method of volatile leaf oil analysis for identification of *Citrus* cultivars. HortScience 26: 186-188

Bagaturiya NS, Bziyava RM, Tsanava NG, Lominadze SD (1990) The effect of mineral nutrition on the yield, quality and composition of the essential oils in lemon and orange fruits. Subtropicheskie Kul'tury 2: 68-76

Dellacassa E, Rossini C, Menendez P, Moyna P, Verzera A, Trozzi A, Dugo G (1992) *Citrus* essential oils of Urugary. I. Composition of oils of some varieties of mandarin. J Essent Oil Res 4: 265-272

Nishiura M, Kamiya S, Esaki S (1971) Flavonoids in *Citrus* and related genera. II. flavonoid pattern and *Citrus* taxonomy. Agric Biol Chem 35: 1691-1706

Ortiz JM, Tadeo JL, Guerri J, Forner JB (1982) Distinction between hybrid and nucellar *Citrus* trees by analysis of their biochemical compounds. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 4-7

Ortiz-Marcide JM, Tadeo-Lluch JL, Diaz-Llanos FJ, Estelles-Adam A (1983) Examination of the essential oils from leaves of the mandarin clementine group. Use in taxonomy *Citrus* clementina. Fruits Paris 38: 125-131

Sawamura M, Kuwahara S, Shichiri K, Aoki T (1990) Volatile constituents of several varieties of pummelos and a comparison of the nootkatone levels in pummelos and other *Citrus* fruits. Agric Biol Chem 54: 803-805

Scora RW (1972) Application of chemical characteristics in plant taxonomy. [*Citrus*]. Umschau 21: 694-696

Sugisawa H, Yamamoto M, Tamura H, Takagi N (1989) Comparison of odour quality of volatiles in peel oils of four kinds of navel orange. J Jap Soc Food Sci Tech 36: 543-550

Ulubelde M, Tan A (1986) A numerical Taxonomic Study of *Citrus* species based on Leaf Phenolic Compounds. Acta Hortic 182: 349-358

Wan RX, Ma JQ, He YZ (1990) A comparison of the physicochemical characters of three lemon cultivars. China *Citrus* 19: 17-18.

Witt W, Mauk CS, Yelenosky G, Bausher MG, Mayer RT (1988) The physiology of cold hardiness in *Citrus* genotypes. Effect of cold-hardening temperatures and plant growth regulators on glycoprotein and isoenzyme profiles. Angew Bot 62: 311-323

Zhu LW (1988) Numerical chemotaxonomical study of *Citrus* in China. Acta Phytotaxonomica Sinica 26: 353-361

Zubrzycki HM, Gogorcena AY, Ortiz Marcide JM (1991) Identification and Characterisation of Varieties and Hybrids of Tangerines by Biochemical and Morphometric Characters. In H Bangyan, H Huibai, H Mingdu, K Chunyen, T Xingjie, Y Qian, S Guifen, eds, Proceedings of the International *Citrus* Symposium, Guangzhou, China, November 5-8, 1990, International Academic Publishers, Guangzhou, China, pp 221-231

5. Isozyme Analysis

Almansa MS, Rio LA, Alcaraz CF, Sevilla F, Del-Rio LA (1989) Isoenzyme pattern of superoxide dismutase in different varieties of *Citrus* plants. Physiologia Plantarum 76: 563-568

Anderson CM, Castle WS, Moore GA (1991) Isozymic identification of zygotic seedlings in swingle citrumelo *Citrus* paradisi X Poncirus trifoliata nursery and field populations. J Am Soc Hortic Sci 116: 322-326

Ashari S, Aspinall D, Sedgley M (1988) Discrimination of zygotic and nucellar seedlings of five polyembryonic *Citrus* rootstocks by isozyme analysis and seedling morphology. J Hort Sci 63: 695-703

Ashari S, Aspinall D, Sedgley M (1989) Identification and Investigation of Relationships of Mandarin Types Using Isozyme Analysis. Scientia Hortic 40: 305-315

Esen A, Scora RW (1977) Amylase polymorphism in *Citrus* and some related genera [Genetic, taxonomic, and phylogenetic relationship]. Am J Bot 64: 305-309

Button J, Vardi A, Spiegel-Roy P (1976) Root peroxidase isoenzymes as an aid in *Citrus* breeding and taxonomy. Theor Appl Genet 47: 119-123

Frias-de-Fernandez AM, Lozzia-de-Canelada ME, Cristobal-de-Hinojo ME, Foguet JL (1989) Isoenzyme variability in a population of Poncirus trifoliata. Lilloa 37: 155

Gogorcena Y, Zubrzycki H, Ortiz JM (1990) Identification of mandarin hybrids with the aid of isozymes from different organs. Scientia Horticulturae 41: 285-291

Gogorcena Y, Zubrzycki HM, Ortiz JM (1989) Identification of mandarins and hybrid mandarins using biochemical characteristics. Investigacion Agraria, Produccion y Proteccion Vegetales 4: 317-331

Gonzalez C, Gonzalez JA (1982) Study of rootstocks for 'Tahiti' lime (*Citrus* aurantifolia Tan.). VII. Isoenzymatic characterisation. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan, International Society of Citriculture, Shimizu, Japan, pp 132-134

Gonzalez C, Roman MI (1982) Electrophoretic analysis in polyacrylamide gel for different types of Poncirus trifoliata, citranges and citrumelos. Ciencia y Tecnica en la Agricultura, Citricos y Otros Frutales 5: 15-24

Hirai M, Kajiura I (1987) Genetic analysis of leaf isozymes in *Citrus*. Japanese Journal of Breeding 37: 377-388

Hirai M, Kozaki I (1986) Isozyme of *Citrus* leaves. In K Kitaura, T Akihama, H Kukimura, K Nakajima, M Horie, I Kozaki, eds, Development of new technology for identification and classification of tree crops and ornamentals, Fruit Tree Research Station, Ministry of Agriculture, Forestry & Fisheries, Yatabe, Japan, pp 73-76

Hirai M, Kozaki I (1991) Isozyme of *Citrus* leaves. In H Bangyan, H Huibai, H Mingdu, K Chunyen, T Xingjie, Y Qian, S Guifen, eds, Proceedings of the International *Citrus* Symposium, Guangzhou, China, November 5-8, 1990, International Academic Publishers, Guangzhou, China, pp 10-13

Hirai M, Kozaki I, Kajiura I (1986) Isozyme analysis and phylogenic relationship of *Citrus*. Japanese Journal of Breeding 36: 377-389

Iglesias L, Lima H, Simon JP (1974) Isoenzyme identification of zygotic and nucellar seedlings in *Citrus*. [Breeding]. J Hered 65: 81-84

Jarrell DC, Roose ML, Traugh SN, Kupper RS (1992) A genetic map of *Citrus* based on the segregation of isozymes and RFLPs in an intergeneric cross. Theor Appl Genet 84: 49-56

Kapanadze BI (1985) Isoenzymes and identification of nucellar and hybrid seedlings in *Citrus* crops. Subtropicheskie Kul'tury 6: 121-126

Kapanadze BI (1985) Aspartate amino transferase and isocitrate dehydrogenase isoenzymes in *Citrus* crops. Stabil'nost' i izmenchivost' genoma 89-96

Kirai M, Kozaki I (1982) Isozymes of *Citrus* leaves. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 10-13

Li WP, He SW, Liu GF (1987) A study of *Citrus* in Hunan province by analysis of leaf peroxidase isozymes. Acta Horticulturae Sinica 14: 153-160

Liao YF (1988) A study of esterase isozymes from related *Citrus* plants. Acta Botanica Sinica 30: 163-168

Lima H (1982) Enzymatic polymorphism in *Citrus*. II. Amylases. Ciencia y Tecnica en la Agricultura, Citricos y Otros Frutales 5: 79-97

Marin E, Lima H (1986) Electrophoretic study of isoenzymes with peroxidase activity in seven mandarin cultivars. Ciencia y Tecnica en la Agricultura, Citricos y Otros Frutales 9: 33-39

Niedz RP, Bausher MG, Hearn CJ (1991) Detection of *Citrus* leaf and seed glycoproteins using biotinylated lectin probes. HortScience 26: 910-913

Protopapadakis EE (1988) Effect of rootstocks on isoenzymic composition of *Citrus*. In R Goren, K Mendel, eds, Citriculture : proceedings of the Sixth International *Citrus* Congress, Middle-East, Tel Aviv, Israel, March 6-11, 1988, Balaban Publishers, Rehovot, Israel, pp 609-614

Protopapadakis EEP, Protopapadakes EE (1987) Identification by isoenzymes of five cultivars of *Citrus* medica grafted on four rootstocks. Journal of Horticultural Science 62: 413-419

Reforgiato-Recupero G (1984) Leaf isoenzymes in the *Citrus* breeding programme of the I.S.A. at Acireale. Annali dell'Istituto Sperimentale per l'Agrumicoltura 17-18: 119-126

Reforgiato-Recupero G, De-Leonardis W, Piccione V, Zizza A (1989) A study on the identification of some lemon biotypes with leaf isozymes analysis and pollinic morphobiometry. In R Goren, K Mendel, eds, Citriculture : Sixth international *Citrus* congress, Middle-East, Tel Aviv, Israel, 6-11 March 1988. Volume 1, Balaban Publishers, Rehovot, Israel, pp 167-174

Sevilla F, Almansa MS, Hellin E, Alcaraz CF (1989) Influence of the *Citrus* species and varieties on the isozyme profile of iron-superoxide dismutases. In R Goren, K Mendel, eds, Citriculture : Sixth international *Citrus* congress, Middle-East, Tel Aviv, Israel, 6-11 March 1988. Volume 1, Balaban Publishers, Rehovot, Israel, pp 561-570

Sharma KK, Jawanda JS (1985) Identification of nucellar and zygotic seedlings in *Citrus* rootstock species through isozyme analysis. Journal of Research, Punjab Agricultural University 22: 277-284

Soost RK, Torres AM (1982) Leaf isozymes as genetic markers in *Citrus*. In K Matsumoto ed, Proceedings of the International Society of Citriculture, November 9-12, 1981, Tokyo, Japan. Volume I, International Society of Citriculture, Shimizu, Japan, pp 7-10

Soost RK, Williams TE, Torres AM (1980) Identification of nucellar and zygotic seedlings of *Citrus* with leaf isozymes. Hortsci 15: 728-729

Torres AM (1983) Fruit trees [Genetic analysis, isozyme techniques, apples, peach, pear, figs, olives, *Citrus*, dates, palms, avocados, mango]. Dev Plant Genet Breed 1B: 401-421

Torres AM (1983) Fruit Trees. In SD Tanksley, TJ Orton, eds, Isozymes in Plant Genetics and Breeding, Elsevier Science Publishers, Amsterdam, pp 401-422

Torres AM (1989) Isozyme analysis of tree fruits. In DE Soltis, PS Soltis, eds, Isozymes in plant biology, Dioscorides Press, Portland, Oregon, USA, pp 192-205

Torres AM, Mau-Lastovicka T, Williams TE, Soost RK (1985) Segregation distortion and linkage of *Citrus* and Poncirus isozyme genes. J Hered 76: 289-294

Torres AM, Soost RK, Diedenhofen U (1978) Leaf Isozymes as Genetic Markers in *Citrus*. Am J Bot 65: 869-881

Torres AM, Soost RK, Mau-Lastovicka T (1982) *Citrus* Isozymes: Genetics and distinguishing nucellar from zygotic seedlings. J Hered 73: 335-339

Ueno I (1976) Application of zymography to *Citrus* breeding. II. Variations in peroxidase isozymes for species, varieties and strains of *Citrus* and its relatives. Bull Fruit Tree Res B3: 9-24

Ueno I, Nishiura M (1976) Application of zymography to *Citrus* breeding. III. Studies of peroxidase isozyme in some graft hybrids. Bull Fruit Tree Res B3: 25-32

Wu MT, Jwo JM (1986) Studies on the isoenzymes of *Citrus* leaves. I. Gel electrophoretic analysis of six important *Citrus* species in Taiwan and the effect of leaf age. Journal of Agricultural Research of China 35: 484-494

Yamashita K (1983) Chimerism of Kobayashi-mikan (*Citrus* natsudaidai X unshiu) judged from isozyme patterns in organs and tissues [Cultivars]. J Jap Soc Hortic Sci 52: 223-230

6. DNA Profiling Techniques

Carstens K (1994) Differentiation of *Citrus* Cultivars by means of Random Amplified Polymorphic DNA. UPOV BMT Meeting

Durham, RE. (1990) Mapping genes involved in freezing tolerance in a backcross of *Citrus* and poncirus using a linkage map of isozymes and restriction fragment length polymorphisms. pp 1-112

Durham RE, Liou PC, Gmitter FG, Moore GA (1992) Linkage of Restriction Fragment Length Polymorphisms and Isozymes in *Citrus*. Theor Appl Genet 84: 39-48

Fortarnau A, Hernandez-Yago J (1982) Characterisation of Mitochondrial DNA in *Citrus*. Plant Physiol 70: 1678-1682

Galun E (1989) Application of molecular methods to modern *Citrus* taxonomy. In R Goren, K Mendel, eds, Citriculture : Sixth international *Citrus* congress, Middle-

East, Tel Aviv, Israel, 6-11 March 1988. Volume 1, Balaban Publishers, Rehovot, Israel, pp 295-301

Green RM, Vardi A, Galun E (1986) The plastome of *Citrus*. Physical map, variation among *Citrus* cultivars and species and comparison with related genera. Theor Appl Genet 72: 170-177

Kapanadze BI, Shevchenko VA (1988) Method of identifying nucellar and hybrid *Citrus* plants. USSR Patent A.s.1376989.:

Komatsu A, Akihama T, Hidaka T, Omura M (1993) Identification of Poncirus Strains by DNA Fingerprinting. In T Hayashi, et al, eds, Techniques on Gene Diagnosis and Breeding in Fruit Trees, FTRS, Japan, pp 88

Liou, PC. 1990 A Molecular Study of the *Citrus* Genome through Restriction Fragment Length Polymorphism and Isozyme Mapping. Florida, USA,

Matsuyama T, Motohashi R, Akihama T, Omura M (1992) DNA Fingerprinting in *Citrus* Cultivars. Japan J Breed 42: 155-159

Omura M, Hidaka T, Nesumi H, Yoshida T, Nakamura I (1993) PCR Markers for *Citrus* Identification and Mapping. In T Hayashi, et al, eds, Techniques on Gene Diagnosis and Breeding in Fruit Trees, FTRS, Japan, pp 66-73

Preston LR, Appels R, Lee LS, Morell MK (1993) Differentiation of Mandarin varieties by DNA profiling techniques. In Focused Plant Improvement, Tenth Australian Plant Breeding Conference, 1993, Vol 2, Brisbane, Australia , pp 231-232

Roose ML (1989) Isozymes and DNA Restriction Fragment Length Polymorphisms in *Citrus* Breeding and Systematics. In R Goren, K Mendel, eds, Citriculture : proceedings of the Sixth International *Citrus* Congress, Middle-East, Tel Aviv, Israel, March 6-11, 1988, Balaban Publishers, Rehovot, Israel, pp 155-165

Roose ML, Traugh SN (1988) Identification and performance of *Citrus* trees on nucellar and zygotic rootstocks. Journal of the American Society for Horticultural Science 113: 100-105

Uematsu C, Omura M, Hidaka T, Ogihara Y (1990) Strain diagnosis of *Citrus* species by means of RFLPs analysis. Japan J Breed 40 (Suppl. 2): 452-453

Yamamoto M, Kobayashi S, Nakamura Y, Yamada Y (1993) Phylogenetic Relationships of *Citrus* Revealed by Diversity of Cytoplasmic Genomes. In T Hayashi, et al, eds, Techniques on Gene Diagnosis and Breeding in Fruit Trees, FTRS, Japan, pp 39-46

Appendix II

Methods for the isolation of *Citrus* DNA.

Method 1

1. Grind 1g fresh plant tissue in a mortar under liquid nitrogen.
2. Add about 4 volumes of extraction buffer, mix well and then add an equal volume of phenol:chloroform (1:1).
3. Mix by inverting the tube sporadically over 10min - 30min.
4. Separate the phases by centrifugation for 5 min at about 3000RPM.
5. Remove the upper, aqueous phase to a fresh tube with the wide bore of a 10ml glass pipette and extract once more with phenol:chloroform.
6. Separate the phases and transfer the aqueous phase to a corex tube (or other centrifugable glass tube).
7. Add 0.5 vol Ammonium Acetate (7.5M) and mix. Layer two volumes of ethanol over the top, mix gently (Optional: transfer to -20°C overnight).
8. Recover the DNA by centrifugation at 3-10,000 RPM for 10-20 min.
9. The DNA is then washed 2 times with 70% ethanol, dried under vacumn, and resuspended in 2ml distilled deionised water or TE if proceeding with additional purity steps.

Final. Check concentration and purity by measuring the $OD_{260/280}$ of a 1:50 dilution and/or running 2-5 µl on a gel.

Additional Purity Steps

i. RNAseA treat the DNA using 25 µg of RNAseA. Incubate at 37°C for 30min.
ii. Proteinase K treat the sample by adding SDS to 0.5% and 50 mg/ml of a 20 mg/ml stock of proteinase K and incubate for 30min at 37°C.
iii. Phenol : Chloroform extract the sample
iv. Ethanol Precipitate
  a. Layer 2 volumes of ethanol and mix gently
  b. Add 0.5 vol Ammonium Acetate (7.5M) and mix, allow to precipitate at room temp for at least 30 min (Optional: transfer to -20°C overnight).
  c. Centrifuge for 15-30 min at 3-10 000 RPM
  d. Discard the supernatant and briefly dry the DNA pellet
v. The DNA is then washed 2 times with 70% ethanol, dried under vacumn, and resuspended in 0.5 ml distilled deionised water.

RNAseA Prepare from powder at a conc. of 10 mg/ml. Boil for 10 min and store at -20°C in small aliquots which should not be refrozen.

Proteinase K Prepare from powder at a concentration of 20 mg/ml in double distilled water. Store at -20°C.

Extraction buffer

Final Conc	Amounts of Stock for 500 ml
0.1M Na_2SO_3	6.302g of powder (mw 126.04)
4% Sarkosyl	100 ml of 20% solution
0.1M Tris-HCl, pH 8.0	50ml of 1M solution
10 mM EDTA	10ml of 0.5M solution

Method 2.

1. Add b-ME to 2xCTAB to a concentration of 0.2% and preheat to 60°C for 5 min (the temperature is dependant on the material to be examined)
2. Grind 0.5g leaf material under liquid nitrogen
3. Using a syringe, transfer 2.5 ml CTAB to another mortar, add the ground leaf material, and mix thoroughly with a pestle.
4. Transfer with a spatula to a 15 ml corex tube and place in the 60°C water bath for 45 min.
5. Add 2.5ml chloroform and extract by inversion for 5-10 min
6. Centrifuge for 10 min at 3000 RPM
7. Remove the aqueous layer to another tube
8. Add 2 volumes of ethanol and allow to precipitate at room temperature for at least 30 min
9. Centrifuge for 15-30 min at 3000 RPM
10. Discard the supernatant and briefly dry the DNA pellet
11. Redissolve in 1 ml distilled deionised water.
12. Reprecipitate DNA with 500 µl NH_4OAc and 2.5 ml ethanol at room temperature for at least 30min
13. Centrifuge for 15-30 min at 3000 RPM
14. Discard the supernatant and briefly dry the DNA pellet under vacumn

It may be necessary to repeat the precipitation to remove all traces of CTAB and chloroform.

15. Redissolve in 1 ml distilled deionised water or TE if proceeding with additional purity steps.

Final. Check concentration and purity by measuring the $OD_{260/280}$ of a 1:50 dilution and/or running 2-5µl on a gel.

Additional Purity Steps

i. RNAseA treat the DNA using 25µg of RNAseA. Incubate at 37°C for 30min.

ii. Proteinase K treat the sample by adding SDS to 0.5% and 50 mg/ml of a 20mg/ml stock of proteinase K and incubate for 30min at 37°C.

iii. Phenol : Chloroform extract the sample

iv. Ethanol Precipitate

a. Layer 2 volumes of ethanol and mix gently

b. Add 0.5 vol Ammonium Acetate (7.5M) and mix, allow to precipitate at room temp for at least 30 min (Optional: transfer to -20°C overnight).

c. Centrifuge for 15-30 min at 3-10 000 RPM

d. Discard the supernatant and briefly dry the DNA pellet

v. The DNA is then washed 2 times with 70% ethanol, dried under vacumn, and resuspended in 0.5ml distilled deionised water.

RNAseA Prepare from powder at a conc. of 10 mg/ml. Boil for 10 min and store at -20°C in small aliquots which should not be refrozen.

Proteinase K Prepare from powder at a concentration of 20 mg/ml in double distilled water. Store at -20°C.

2xCTAB Extraction Buffer

10g	CTAB
140ml	5M NaCl
25ml	2M Tris-HCl pH8.0
20ml	0.5M EDTA

The use of glass corex tubes encourages the preferential precipitation of DNA rather than DNA plus RNA.

The use of NH_4OAc and room temperature gives somewhat better yields of DNA than NaOAc and -20°C

Appendix III

***Citrus* RAPD Primers and PCR conditions**

Experimental Material

Fresh leaves from actively growing mandarin trees were supplied on two separate occasions and isolated using different methods.

DNA primers

The DNA primers were supplied by Operon

Code	Sequence - 5' TO 3'
A1	CAGGCCCTTC
A3	AGTCAGCCAC
A4	AATCGGGCTG
A5	AGGGGTCTTG
A9	GGGTAACGCC
A11	CAATCGCCGT
A13	CAGCACCCAC
A17	GACCGCTTGT
A18	AGGTGACCGT
Histone	GTCACCGCCATGG

Experimental Procedure

DNA was isolated from the leaves using a Sulphite/Sarkosyl method (Peter Langridge pers. comm.) or a modification of the method of Hillis DM et al (1990). This DNA was used as template with a selection of arbitrary sequence 10 oligonucleotide long primers in PCR reactions (using the Corbett capillary PCR machine) (Williams, JGK. et al 1990) with cycling conditions and reaction mix modified to suit the Corbett machine. The DNA products were identified by gel electrophoresis and scored manually.

1. DNA samples were prepared at a concentration of 50 ng/2µl in distilled deionised water.
2. Stock reactions were made up as shown with reagents added in order, and kept on ice.

	per sample	7x stock mix
H_2O(distilled deionised)	4.0µl	28.0
25mM $MgCl_2$	0.8µl	5.6
5x buffer	2.0µl	14.0
Primer(1/10)	1.0µl	7.0
Taq Polymerase(5U/µl)	0.2µl	1.4
	8.0µl	

*A control sample containing no DNA was included in all cases.

3. The stock mix was vortexed and spun briefly to return all the solution to the base of the tube.
4. 8µl of stock reaction was dispensed to each sample eppendorf
5. 2µl of the prepared DNA sample was added to the side of the tube and spun to mix.
6. The sample was drawn into a positive displacement tube and heat sealed.
7. The sample was then returned to ice until all samples were prepared.

8. Tubes were placed in the Corbett PCR machine and the following programme was run.

Cycle	Step	Temp. (°C)	Time (min)	TTC
1	1	92	2.00	
	2	35	2.00	
	3	72	1.30	1
2	1	92	0.10	
	2	35	2.00	
	3	72	1.30	4
6	1	92	0.10	
	2	40	0.25	
	3	72	1.30	35
41	1	92	0.10	
	2	40	0.20	
	3	72	5.00	
	4	25	1.00	1

9. The tubes were then removed from the thermocycler and the samples run on 1% agarose gels containing 0.5 µg/ml Ethidium bromide.

5x PCR buffer

Component	Stock	Volume	Final Conc.
Tris-HCl pH 8.8	1M	3.35 ml	335 mM
$(NH_4)_2SO_4$	1M	830µl	83 mM
dNTP's	100 mM each	100µl each	1 mM
gelatin	10 mg/ml	1.0 ml	0.1% (1mg/ml)
Triton X-100	10%	2.25 ml	2.25%
H_2O(dist. deionised)		2.17 ml	

Appendix IV

Data Collection

The systems used for the collection and analysis of PCR data vary in sensitivity and in their ability to resolve closely migrating DNA bands.

The most familiar method to the molecular biologist is the use of agarose gel electrophoresis to resolve the PCR products on the basis of length and to visualise the position of the DNA bands by staining with ethidium bromide and viewing or photographing the gel under UV light. This system can resolve bands in the range from several hundred base pairs to several thousand base pairs, although the errors in assigning sizes may be as much as +/- 5%. The ethidium bromide staining method is capable of detecting bands containing 15 pmol of DNA/200 bp band. The agarose gel electrophoresis/ethidium bromide system has the advantage that it is relatively inexpensive and technically straightforward.

The use of silver staining of DNA in polyacrylamide gels provides a system which is capable of greater resolution of DNA fragments, particularly in the range from 25 bp to 1 kbp. The silver staining procedure is also significantly more sensitive, being capable of detecting bands containing 0.15 pmoles of DNA/200 bp band of DNA (Bassam, BJ. *et al.*, 1991).

A significantly more highly resolving and sensitive system involves the use of a DNA sequencer with a laser detection system. The system produced by Applied Biosystems and using Genescan software requires the use of DNA primers which have been labelled at the 5' end with a fluorescent dye molecule. The PCR products are separated in a polyacrylamide gel and in the range from 50 bp to 350 bp the sizing precision is greater than 99%, allowing the resolution of alleles which differ by as little as 1 bp. This system is capable of detecting bands containing less than .1 pmol/ 200 bp band of DNA. A major advantage of the system is that internal standards are run in each lane, allowing very accurate sizing by minimising any artefacts arising from the electrophoresis system. The system also allows the products of different PCR reactions to be electrophoresed in the same lane if the PCR primers have been labelled with different dyes. This allows for extremely accurate comparison of samples from different sources.

Appendix V

Data Analysis

Laboratory procedures provide the means for identifying genetic differences between varieties. It is then essential that the genetic differences be quantified using appropriate statistical procedures. While in general many laboratories follow similar procedures for the statistical analysis of genetic data, there are nevertheless various combinations and permutations of data analysis and presentation that make it difficult to compare results among labs. We believe that it is in the interests of UPOV to standardise some of the more important statistical procedures so as to allow meaningful comparisons among varietal registrations. Ultimately, within certain groups of plants it might be feasible to actually specify genetic criteria (eg. no. primers, no. bands etc) and perhaps minimum genetic distances acceptable among varieties.

The way one proceeds with data analysis depends on the type of data to be analysed. Multilocus data which is generated by RAPDs, other fingerprinting methods or multilocus RFLP probes are analysed differently to single locus data obtained for sequence tagged sites such as microsatellites or single locus RFLP probes. In the next section of the paper we will focus primarily on the statistical procedures that are particularly applicable to RAPD analysis but include a brief discussion on single locus data analysis and procedures common to both. Next we will look at the question of genetic distance between varieties before turning to the problem of outcrossing species. Finally we will make some recommendations for standardising statistical procedures. A number of suitable computer programs are presently available, mostly from the public domain, those that are known to us are listed in the appendix.

Steps in the analysis of multilocus profiles

1. Scoring the profile

The first step in data analysis is to score the profile. With small numbers of samples this is often done manually from gel photographs with the aid of a ruler or callipers. With larger data sets it is more convenient to digitise the profiles and analyse them with the various computer packages available. Automatic sequencers of course automatically produce a digitised output. Digitised profiles avoid human subjectivity and error and are therefore preferred where possible. Regardless of the way the DNA profile is stored the goal of scoring is to convert the profile for each sample into a vector of 1's and 0's representing band presence or absence respectively. Thus ultimately, each individual DNA profile is represented by a vector (eg: 10111001) which can then be used to calculate genetic distances.

The scoring of bands as merely present or absent does not take into account differences in band intensity. Some workers have reported that these qualitative differences in band intensity are reproducible and thus may constitute an additional character. Therefore, a number of labs have attempted to score band intensity in addition to band presence or absence. The value of this approach will be reconsidered when we turn to multivariant analysis below but these qualitative differences are not easily incorporated into genetic distance calculations.

2. Calculating genetic distances

Once vectors of band presence/absence are available the next step is to calculate a pairwise genetic distance matrix. This matrix is produced by comparing the vector of a given individual with that of every other individual and calculating a distance value based on a given formula. This process is repeated for all individuals. Thus for n individuals a matrix of n^2-n genetic distances will be produced.

Various formulae for calculating genetic distances or dissimilarity are available. One of the most widely used is that of Nei and Li (1979):

$$1-F = \left[1 - \frac{2n_{xy}}{n_x + n_y} \right]$$

where $2n_{xy}$ equals the number of shared bands and n_x and n_y is the number of bands observed in individual x and individual y respectively.

A second genetic distance is the Euclidean distance of Excoffier et al. (1992) as presented in Huff et al. (1993):

$$E = n \left[1 - \frac{2n_{xy}}{2n} \right]$$

where $2n_{xy}$ equals the number of shared bands and n equals the total number of banding positions.

This is the preferred formula for calculating genetic distances for subsequent Analysis of Molecular Variance (AMOVA). Also, because this measure is actually a tally of band differences between individuals with a range from 0 to n, it is more comprehensible than 1-F which ranges from 0-1.

For both these genetic distances, shared bands include both 00 and 11 comparisons, ie. both the presence or the absence of bands in two individuals is counted as a shared. This is not the case for the Jaccard coefficient which has been used for RAPD data by Stiles et al. (1993). For this coefficient the absence of bands in two individuals is not counted as shared. When comparing results among labs, it is important to note that often workers will present there data as genetic similarity which is simply 1 minus the genetic distance. However, because UPOV is primarily interested in distinctness, genetic distance rather than genetic similarity seems more appropriate.

3. Visualizing genetic relationships

While the presentation of a pairwise genetic distance matrix may be informative for small sample sets, it is usual to present the results as a dendrogram. Furthermore, a dendrogram often allows one to infer the patterns of genetic relationship among the taxa. Dendrograms are produced by running the distance matrix through a computer program that performs a cluster analysis. Various algorithms for clustering data are available but the UPGM methods is the most commonly used.

For inbred or clonal varieties that exhibit little genetic variation within the taxon, production of a dendrogram often completes the data analysis. On the other hand, for outcrossing varieties that exhibit variation within taxa, further analysis is appropriate and may indeed be essential to enable varieties to be identified. Further procedures include multivariant analysis and the Analysis of Molecular Variance which are discussed further in later sections.

Steps in the analysis of single locus profiles

1. Scoring the profile

While there are presently, few studies that have employed DNA profiling at single loci in plants, procedures for analysing this type of data are well established because of the large body of data on allozymes. Nevertheless, new methods of statistical analysis are anticipated since single locus DNA profiling is fast becoming an essential tool for the analysis of identity and paternity in humans (Weber 1990). The process of scoring single locus profiles is different to multilocus data. Individual alleles are identified and genotypes can then be assigned to each locus. Alleles are usually scored by numbering alleles 1 to n from the fastest to the slowest to migrate on the gel. Homozygotes for the alleles are scored as 11, 22 ... nn, while heterozygotes are scored as 12, 13, 23, 34 etc. When multiple loci are scored, the genotypes of each locus can be combined to produce a multilocus genotype. For selfing or inbred varieties a multilocus genotype may be diagnostic of the variety and little further analysis may be necessary.

2. Calculating allele frequencies

The second step in the analysis of single locus data is to calculate the allele frequencies for the data set. Formulae for calculating allele frequencies are presented in any standard genetic text and will not be reproduced here.

3. Calculating genetic distances

Allele frequencies subsequently form the basis of genetic distance and other statistics. Genotypic data can also be combined over loci to give a multilocus genotype. There are a series of genetic distances formulae available the most popular are those of Nei (1972).

4. Visualizing genetic relationships

Pairwise genetic distances can be used to generate a dendrogram via cluster analysis as for multilocus data.

Dealing with outcrossing varieties?

Most inbred, clonal or predominant selfing varieties will exhibit little genetic variation within the taxon. Therefore, the DNA profile of one to several individuals will be representative of the variety and variation within the taxon will rarely need to be considered. In these cases, cultivars are characterised by fixed differences for at least one to many bands or loci and these differences are readily discernible on a gel with the naked eye. However, in outcrossing varieties, there may be few fixed differences among varieties with frequency differences accounting for most of the genetic divergence. Consequently, these differences may not be discernible by eye and statistical analysis is essential. It follows, that for outcrossing species many more samples may need to be analysed since it is necessary to document the level of variation within varieties before differences between varieties can be determined. The procedure described below offers a new approach to data analysis for outcrossers that is applicable to both single and multilocus data.

AMOVA

As indicated in the preceding discussion, identification of outcrossing varieties requires an analysis of genetic variation at two levels: within and among varieties. The patterns of genetic variation within and among hierarchical groups of interest have traditionally been analysed by Wrights Hierarchical F-statistics or Nei's analogous Gene diversity statistics (Wright 1951; Nei 1977). Recently, Excoffier et al. (1992) have introduced an Analysis of Molecular Variance (AMOVA) approach which produces analogs to Wrights F statistics. AMOVA is a powerful procedure for the analysis of genetic variation when there is an hierarchical structure in the data set. For example, a data set consisting of individuals within varieties, varieties within species, species within a genus etc. AMOVA performs an analysis of variance within and among the different hierarchical levels of the data set and produces significance values based on random permutation at all levels of analysis. The ability to test for statistical significance in the data set is not provided by other types of analysis. Furthermore, by using random permutation the analysis is not dependant on assumptions about the distributions of the data. These assumptions imposed by many statistical procedures are frequently violated, as for example the assumption of normality which applies to standard analysis of variance. Another feature of AMOVA is that it can accommodate different pairwise genetic distance matrices. Therefore, the procedure is applicable to the analysis of both multi locus and single locus data. For example, Peakall et al. (In review) have compared the patterns of allozyme (single locus data) and RAPD (multilocus data) variation in a turf grass using AMOVA. In this study, the ability to use AMOVA for the analysis of both types of data has permitted explicit comparisons between the different data types because a common procedures and statistics are used. Because PVR applications of outcrossing species will frequently present genetic data obtained from different types of genetic marker (eg: RAPD vs RFLP), we believe the use of AMOVA will enable meaningful comparisons to be made both among different genetic markers and among registrations and is a particularly valuable tool for defining distinctness among outcrossing varieties.

Multivariate Analysis

Another powerful way to analyse DNA profiles is by multivariate analysis. While there are many different approaches to this type of analysis, all enable one to visualize clusters in 3 dimensional space. This is particularly valuable for analysing outcrossing species. Adams et al. (1993) describe a 3-D ordination procedure for RAPDs and have made their program freely available.

A computer program for RAPD data analysis

A flexible computer program for the analysis of RAPD data will shortly be freely available from the Australian National University. The program can accept data directly, or from a spreadsheet and calculates various genetic distances which are output in formats compatible with other programs such as WINAMOVA and various tree

drawing programs. The program is written in C and versions for both IBM PC compatibles and the Macintosh will be available.

Further information contact:

Dr Rod Peakall

Division of Botany and Zoology

The Australian National University, Canberra ACT 0200 Australia

Ph (616) 249 0022, Fax (616) 249 5573, Email: rod.peakall@anu.edu.au

[End of document]