Theor Appl Genet (1995) 91: 1248–1252

S. J. Orford · N. Steele Scott · J. N. Timmis

A hypervariable middle repetitive DNA sequence from citrus

Received: 2 February 1995 / Accepted: 9 February 1995

Abstract The use of hypervariable sequences for DNA typing of plants is focussed on microsatellites and on amplification of regions defined by random (RAPD) or defined (AFLP) primers for PCR analysis of genomes. A hypervariable length of middle repetitive DNA has been isolated from citrus that contains no obvious hypervariable structures. The fingerprinting probe was shown to have an important commercial application in the separation of zygotic from nucellar progeny. A somatic variant of the sequence within one orange tree suggests that somatic variation in hypervariable markers may be a common event.

Key words DNA fingerprinting · Citrus · Hypervariable DNA · Cultivar typing · Somatic variation

Introduction

Eukaryotic genomes contain repetitive heterogeneous classes of non-coding DNA that are found within and between genes. Apart from highly repetitious satellite DNA, a second class of coding or non-coding DNA is referred to as middle-repetitive DNA. Middle-repetitive DNA includes interspersed sequences such as transposable elements, inverted repeats and tandem repeat sequences that are variable in number (VNTRs) or "minisatellites" (Schaal et al. 1991). Some middle-repetitive DNA sequences are highly polymorphic (hypervariable). Detection, by Southern blots, of hypervariable regions in the genomes of many eukaryotes may produce a genotype-specific hybridisation pattern. The complex set of DNA fragments produced by this method has been termed a DNA "fingerprint" (Jeffreys et al. 1985).

Communicated by H. F. Linskens

S. J. Orford (✉) · J. N. Timmis
Department of Genetics, The University of Adelaide, Adelaide, South Australia 5005

N. Steele Scott
CSIRO Division of Horticulture, GPO Box 350, Adelaide, South Australia 5001

Many classes of genetic markers and DNA-based diagnostic techniques are available to plant geneticists and breeders. Jeffreys' minisatellite probes have been used to detect cultivar-specific patterns in the genome of rice *Oryza sativa* (Winberg et al. 1993). Probes derived from sequences within the M13 bacteriophage genome have been used to detect polymorphism in many plant genomes including barley (Rogstad et al. 1988), roses (Tzuri et al. 1991) and apples (Nybom and Schaal 1990). More recently, polymerase chain reaction (PCR) techniques have been used in microsatellite detection and random amplified polymorphic DNA (RAPDs) have provided genetic markers for applications such as genome mapping (Morgante and Olivieri 1993; Rafalski and Tingay 1993) and cultivar typing (Tancred et al. 1994). Genetic mapping with RAPD markers has been initiated in citrus (Cai et al. 1994).

DNA fingerprinting has the capacity to become a powerful tool in the breeding of new citrus varieties. Most citrus species, including sexually sterile hybrids, are able to reproduce asexually by means of adventitious (nucellar) embryony (Cameron and Frost 1968; Frost and Soost 1968). The seeds resulting from most crosses are both zygotic and nucellar in origin, the latter being genetically uniform and identical to the seed parent, and breeding programmes must distinguish between these two seed types. Distinction between nucellar and zygotic progeny can sometimes be achieved using morphological characters (Frost and Soost 1968), but such features may not appear until the tree has reached maturity. Isozymes have been used routinely in the separation of zygotic from nucellar progeny (Spiegel-Roy et al. 1977; Torres et al. 1978), but electrophoretic analysis of several peroxidase enzymes must be carried out concurrently in order to distinguish between cultivars. Isozyme variation is not truly representative of genotypic variation, and gel patterns may vary between various organs of the same plant and between plants maintained under different growth conditions (Scandalios 1974). The amount of work and costs involved could be much reduced by a reliable discrimination method applied at the seed or seedling stage of growth.

Accurate plant and plant product identification is crucial to many aspects of modern agriculture. DNA fingerprinting could provide a simple, reliable and stable method of varietal identification in citrus. DNA fingerprinting allows cultivars to be resolved by a single analysis independent of environmental and developmental factors. Citrus species are particularly amenable to fingerprinting techniques since, while they are genetically diverse plants, varieties are propagated asexually. Consequently, all individuals of a particular variety are clonal, and intravarietal variation should be minimal.

This article describes the identification and characterisation of a hypervariable sequence in the DNA of citrus and its use in the production of cultivar-specific DNA fingerprints.

Materials and methods

Citrus leaf samples

The *Citrus sinensis* (L.) Osbesk leaf samples supplied by CSIRO Division of Horticultural Research, Merbein, Victoria were 'Washington Navel Orange' (five trees), 'Biggs Early Navel Orange' and 'Lowes Late Navel Orange'. Samples from the *Citrus reticulata* Blanco ('Cleopatra Mandarin') and *C. sinensis* × *P. trifoliata* ('Carrizo Citrange') cross included 'Cleopatra Mandarin' and one nucellar seedling, 'Carrizo Citrange' and one nucellar seedling and two hybrid progeny. Other samples supplied were *Citrus aurantium* L. ('Sour Orange'), *Citrus limon* (L.) Burm. f. (Burm.) Mewill ('Karna Lemon' and 'Rough Lemon'), *Citrus paradisi* Macf. ('Davis Grapefruit'), *Citrus maxima*, ('Pummelo'), *Citrus aurantifolia* (Christm.) Swingle ('West Indian Lime'), *Poncirus trifoliata* (L.) Raf. ('Williams Trifoliate Orange'), *Citrus indica* ('Wild Indian Orange') and an Australian native, *Microcitrus virgata* (a hybrid between *Microcitrus australasica.* (F. Muell.) swing. and *Microcitrus austrans* (Planch.) swing).

Samples of leaf material from a 40-year-old *Citrus sinensis* ('Navel Orange') tree were obtained from Wittunga Botanic Garden, Blackwood, South Australia.

Samples of *Citrus limon* leaf material from 'Eureka Lemon', 'Lisbon Lemon', 'Variegated Lisbon Lemon' and 'Meyer Lemon' were obtained from Caudles Citrus Nursery, Old Noarlunga, South Australia.

Isolation of plant nucleic acids

Total plant genomic DNA was isolated by the method of Scott and Possingham (1980).

A small-scale method was used to isolate DNA from the leaves of Washington Navel Orange trees. Approximately 400 mg of plant tissue was ground to a powder in liquid nitrogen and thawed into 2 ml of extraction buffer [2×SSC, 50 m*M* EDTA pH 8.0, 2%(w/v) sarcosyl]. The resulting suspension was twice extracted with phenol/chloroform and NaCl added to 1.5 *M*. Plant lysate (200 µl) was loaded onto a 1 ml Sepharose CL-4B column (Bio-Rad) which had been equilibrated with 1.5 *M* NaCl, 50 m*M* EDTA, 100 m*M* Tris-Cl pH 7.2. The first 8 drops to be eluted were discarded, while the next 6 drops, containing the DNA, were collected and precipitated with an equal volume of isopropanol. The DNA was resuspended in 50 µl 1×TE. Between 3 µg and 5 µg of restrictable DNA was isolated per preparation.

Southern blotting and hybridisation

DNA samples were restricted with restriction endonucleases under conditions recommended by the manufacturer (Boehringer Mannheim) and electrophoresed on 0.8%(w/v) agarose (Sigma) gels. Following electrophoresis the DNA was transferred to a Hybond-C Extra membrane (Amersham) following the manufacturer's instructions.

Plasmid inserts requiring radioactive labelling were purified from an agarose gel using the freeze-squeeze method (Thuring et al. 1975) using Ultrafree-MC 0.45-µm filter units (Millipore). Probes were labelled with α-[^{32}P]dATP (Bresatec) using the Hexaprime DNA Labelling Kit (Bresatec).

Prehybridisation, hybridisation and washes were carried out under conditions recommended by Amersham.

Purification of plasmid DNA

Plasmid DNA for sequencing and Southerns was purified using the alkaline lysis method of Sambrook et al. (1989), followed by centrifugation through a Sepharose CL-6B column (Murphy and Kavanagh 1988).

Preparation of insert DNA

Five bands were excised from a *Sau*3A digest and three bands from an *Alu*I digest of total citrus ('Eureka Lemon') DNA. High-molecular-weight genomic DNA (>6 kb) was sheared by agitation through a Pasteur pipette and vortexing. Single-stranded ends were removed by treatment with Mung Bean Nuclease (Pharmacia) under conditions recommended by Sambrook et al. (1989).

Ligations

Ligations were performed using 200 ng of insert DNA and an equal amount of vector DNA, catalysed by T_4 DNA Ligase (New England Biolabs) and conducted under conditions recommended by the manufacturer. Blunt-end ligations were performed in the presence of 15%(w/v) PEG 8000.

Transformation of competent cells

Preparation of competent Epicurian coli PLK-F' cells (Stratagene) was performed using the calcium chloride method as per Sambrook et al. (1989).

Transformation of PLK-F' cells utilised the heat shock method described by Sambrook et al. (1989). Cells were plated on blue/white colour selection medium (Sambrook et al. 1989) supplemented with 100 µg/ml ampicillin and grown O/N at 37°C.

Generation of nested deletions

Exonuclease III-nested deletions of clone pS5j were made in both directions (Henikoff 1984). Cloned DNA (10 µg) was restricted to completion with *Sma*I and *Sac*I (forward primer end) or *Sal*I and *Sph*I (reverse primer end). Following phenol/chloroform extraction, 5 µg samples were treated with 100 units of Exonuclease III (Pharmacia), with 1 µg samples being removed at 1 min intervals. The reaction was stopped with an equal volume of 10 m*M* Tris-Cl, 10 m*M* EDTA pH 8.0, followed by heating at 65°C for 10 min. Samples were treated with Mung Bean Nuclease (Pharmacia) before half of each was restricted with *Eco*RI and the sizes checked by gel electrophoresis. Time samples in the desired size range were ligated and transformed into PLK-F' cells. Insert sizes were screened by small-scale plasmid preparation (He et al. 1990) and excision of the inserts.

Sequencing

Double-stranded DNA templates were sequenced by the dideoxy chain termination method of Sanger et al. (1977) using a Sequenase

Version 2.0 kit (United States Biochemical Corp) and α-[^{32}P]dATP (Bresatec). DNA sequencing reactions were electrophoresed in a 5%(w/v) acrylamide, 7 *M* Urea and 1×TBE (0.9 *M* Tris-borate, 0.002 *M* EDTA) denaturing polyacrylamide gel using a BRL sequencing gel electrophoresis system. Sequence analysis was done using the Genetics Computer Group Sequence Analysis Software Package version 6.1 (Devereaux et al. 1984).

Results

Restriction enzymes which have 4-base pair (bp) recognition sequences would be expected to cut, on average, once every 256 bp in a random sequence. *Sau*3A and *Alu*I restrictions of 'Eureka Lemon' (*Citrus limon*) genomic DNA produce, in addition to a majority of low-molecular-weight fragments, a small number of larger than expected fragments of DNA, visible as bands on agarose gels (result not shown). These fragments contain atypical DNA, which for some reason lacks a particular recognition sequence. A repeat sequence lacking such a site could be dispersed throughout the genome. A probe specific for such regions would be expected to generate DNA fingerprints because of variation in the genomic regions flanking the repeat.

Cloning of this atypical DNA produced four clones, pS5a, pS5c, pS5i and pS5j, each of approximately 3 kb in length. A Southern blot of these clones was probed with total genomic DNA from Eureka Lemon in order to assess the degree of repetition of the cloned sequences in the genome. The strong hybridisation signal detected for the pS5a insert implied high genomic copy number, while the signals of the other clones were weaker in intensity, characteristic of less highly repeated genomic sequences (result not shown).

One clone (pS5j) appeared to have potential as a fingerprinting probe. This clone produced variable banding patterns when used as a probe to Southern blots of *Eco*RI-digested citrus genomic DNAs (Fig. 1), resulting in species-specific DNA fingerprints. Discrete bands were produced, with 12–20 discernible bands per variety. Banding patterns were not, however, unique to each cultivar. 'Eureka Lemon' (track 1) and 'Variegated Lisbon Lemon' (track 3) exhibited identical fingerprints, as did the 3 'Navel Orange' cultivars (tracks 6–8). 'Meyer Lemon' (track 2) is thought to be a hybrid of lemon and orange (Alexander 1983) and produced patterns distinct from the lemons and oranges. 'Williams Trifoliate Orange' (track 5) is a cultivar of *Poncirus trifoliata*, and it produced a pattern highly divergent from those of any *Citrus* species and from *Microcitrus virgata* (track 15). Similarly, *Citrus indica*, the 'Wild Indian Orange' (track 13), is a grapefruit hybrid (Hodgson 1967), and its DNA produces a fingerprint that is different from that of 'Davis Grapefruit' (track 12). Longer exposure of the Southern hybridisation revealed additional bands of lower molecular size that were potentially useful in further distinguishing the citrus species.

Citrus species are particularly amenable to DNA fingerprinting techniques because they are propagated clonally, therefore intravarietal variation is not expected. To

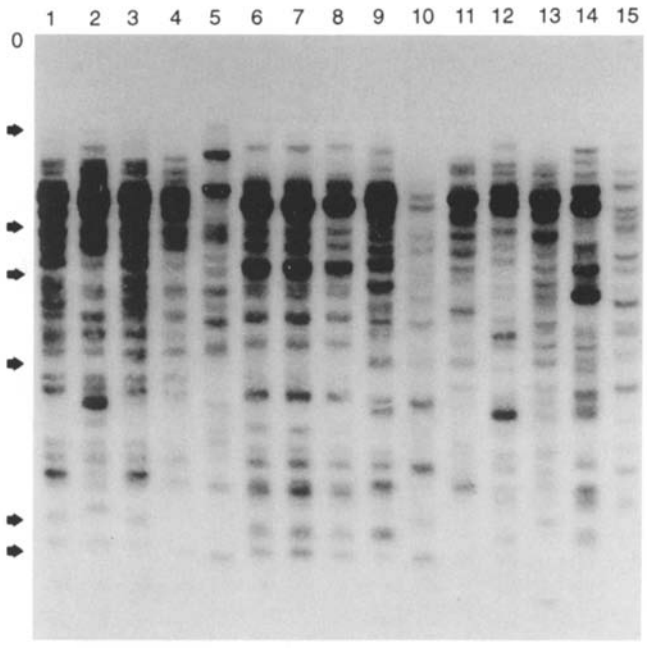

Fig. 1 Hybridisation of the pS5j probe to citrus genomic DNAs. Each track was loaded with 3 µg of DNA restricted with *Eco*RI. Markers on the *left* represent molecular weights of 23.1 kb, 9.6 kb, 6.6 kb, 4.4 kb, 2.3 kb and 2.0 kb. Tracks are: *1* 'Eureka Lemon', *2* 'Meyer Lemon', *3* 'Variegated Lisbon Lemon', *4* 'Rough Lemon', *5* 'Williams Trifoliate Orange', *6* 'Lowes Late Navel', *7* 'Biggs Early Navel', *8* 'Washington Navel', *9* 'Sour Orange', *10* 'Cleopatra Mandarin', *11* 'Tahitian Lime', *12* 'Davis Grapefruit', *13 Citrus indica*, *14* 'Pummelo', 15 *Microcitrus virgata*

test this, variation between five clonally propagated individuals of 'Washington Navel Orange' (*Citrus sinensis*) was assessed using probe pS5j. Banding patterns were indistinguishable in the DNA from each of the trees analysed (result not shown).

DNA fingerprints derived from different tissues of the same plant should be identical if the probe is to be of use in horticulture. To examine the possibility of somatic variation, pS5j was used to probe DNA from several branches of a 40-year-old 'Navel Orange' tree (Fig. 2a). One major difference between the branches was apparent. Branch 1a could be distinguished from the remaining branches by the presence of an additional band of about 3 kb in size. Pattern differences between branches were not apparent on the ethidium bromide-stained gel.

The pS5j DNA fingerprinting probe may have important practical applications, for example in distinguishing zygotic from nucellar embryos. Restricted DNA from 'Cleopatra Mandarin' (maternal parent), 'Carrizo Citrange' (paternal parent), together with two previously characterised nucellar and two hybrid progeny were probed with pS5j (Fig. 2b). DNA from the nucellar progeny exhibited hybridisation bands identical to those of their respective maternal parent. In contrast, fingerprints of the zygotic progeny clearly differed from each other and from both parents. Each band in the zygotic progeny can be attributed to one or other of the parents, indicating that the al-

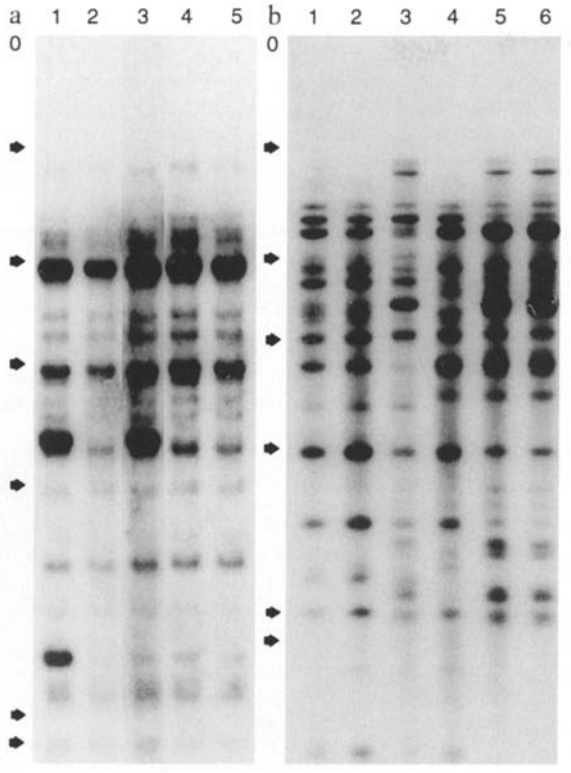

Fig. 2a, b Further analysis of a citrus DNA fingerprinting probe. Markers on the *left* represent molecular weights of 23.1 kb, 9.6 kb, 6.6 kb, 4.4 kb, 2.3 kb and 2.0 kb.
a Hybridisation of the pS5j probe to genomic DNA from the branches of a 40-year-old Navel Orange tree. Each track contains 3 μg of DNA digested with *Eco*RI and the tracks are: *1* branch 1a, *2* branch 1b, *3* branch 2, *4* branch 3, *5* branch 4. **b** Southern blot of *Eco*RI-restricted DNAs from parents and progeny of a citrus cross hybridised with the pS5j insert. Each track contains 3 μg of genomic DNA. The tracks are: *1* 'Cleopatra Mandarin' nucellar progeny, *2* 'Cleopatra Mandarin' parent, *3* hybrid seedling 80-05-05, *4* hybrid seedling 80-05-13, *5* 'Carrizo Citrange' parent, *6* 'Carrizo Citrange' nucellar progeny

leles detected by probe pS5j are inherited in a Mendelian fashion and are of nuclear origin.

The finding of hypervariable regions in the genomes of citrus species prompted further characterisation of the informative probe. A partial restriction map of clone pS5j was constructed (result not shown), and the size of the insert was estimated to be 2.8 kb. The sequencing of nested deletions in both directions produced a complete sequence of 2986 bp (not shown). The GC content was estimated at 26.6%, which is consistent with its resistance to digestion by *Sau*3A. The sequence data indicated an absence of microsatellites, obvious repeat units or significant open reading frames. Screening of databases for sequence homology revealed no significant homology with any known DNA sequences.

Discussion

A dispersed middle repetitive sequence has been isolated from the genome of *Citrus limon* ('Eureka Lemon'). The sequence had homologous counterparts in all of the citrus genomes tested, detecting sufficient variation for the production of citrus species-specific DNA fingerprints. This is the first hypervariable sequence to be characterised from a citrus genome.

The potential applications of probe pS5j and extensions of this technology include determination of taxonomic relationships between citrus species, screening of germ plasm resources and the policing of Plant Variety Rights (1987). Absence of variation between several 'Navel Orange' cultivars (Fig. 1, tracks 6–8) and two Lemon cultivars (Fig. 1, tracks 1 and 3) suggests that pS5j is not applicable for citrus identification at the cultivar level.

Banding patterns were stable between clonally propagated plants and between cultivars of the same species. However, somatic instability was detected within an individual 'Navel Orange' tree. Fingerprint pattern consistency throughout one plant is an important criterion if a probe is to be used for horticultural purposes. A possible explanation for the observed variation is somatic mutation, which is a common phenomenon in citrus (Cameron and Frost 1968). Repeated sequences are commonly involved in rapid genome changes, and parts of the plant genome are claimed to be rapidly modulated in response to stress or differing environmental conditions (Cullis 1985). The age and size of the tree implies that there have been many rounds of DNA replication at which such mutational events could occur. Examination of different leaf clusters on branch 1a could indicate the timing of the putative mutational event.

The presence of variability in DNA fingerprints between the branches of an individual tree suggests that pS5j may be applicable to fine-scale identification of citrus varieties, but this result contrasts with the apparent lack of variation between cultivars. The extent of somatic instability will become apparent with more extensive screening of clonally propagated individuals. Any sequence suitable for fingerprinting must be sufficiently variable to allow taxonomic differentiation but not so variable that reproducible and reliable results become difficult to obtain.

The pS5j fingerprinting probe was shown to have a potentially useful practical application. Sufficient polymorphism was detected to distinguish between zygotic and nucellar progeny of a 'Cleopatra Mandarin' × 'Carrizo Citrange' cross. Combined with refinement of the small-scale method for plant genomic DNA isolation, the Southern analysis depicted in Fig. 2b could greatly reduce the time and cost involved in distinguishing nucellar from zygotic progeny in citrus crosses. Biochemical and morphological tests could be rendered obsolete by a molecular genetic analysis undertaken with a small amount of leaf material at the seedling stage of plant development.

The hybridisation patterns indicate stable Mendelian inheritance of pS5j-associated polymorphisms. Therefore, the fingerprinting probe is unlikely to be cytoplasmic (mitochondrial or chloroplast) in origin, as these genomes show uniparental inheritance. The AT-rich sequence does not contain obvious hypervariable sequences, such as minisatellites or microsatellites. There are many categories of dis-

persed repeat sequences in plant genomes into which the isolated sequence could fall, including SINEs (short interspersed repeat elements), LINEs (long interspersed repeat elements) and transposable elements. The latter are noted for their ability to cause mutations at high frequencies in plant genomes (Vedel and Delseny 1987).

Despite exhibiting some somatic instability, the DNA fingerprinting probe isolated provides a potentially powerful applied tool in citrus breeding and in species identification.

References

Alexander DMcE (1983) Some citrus species and varieties in Australia. CSIRO Division of Horticultural Research, Adelaide, South Australia

Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in citrus using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theor Appl Genet 89:606–624

Cameron JW, Frost HB (1968) Genetics, breeding and nucellar embryony. In: Reuther W, Webber HJ, Batchelor LD (eds) The citrus industry, vol. 2. University of California Press, Berkeley, Calif., pp 325–370

Cullis CA (1985) Sequence variation and stress. In: Hohn B, Dennis ES (eds) Genetic flux in plants. Springer, Berlin Heidelberg New York, pp 157–168

Devereaux J, Haerberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Frost HB, Soost RK (1968) Seed reproduction: development of gametes and embryos. In: Reuther W, Webber HJ, Batchelor LD (eds) The citrus industry, vol. 2. University of California Press, Berkeley, Calif., pp 290–324

He M, Wilde A, Kaderbhai MA (1990) A simple single-step procedure for small-scale preparation of *Escherichia coli* plasmids. Nucleic Acids Res 18:1660

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359

Hodgson RW (1967) Horticultural varieties of citrus. In Reuther W, Webber HJ, Batchelor LD (eds) The citrus industry, vol. 1. University of California Press, Berkeley, Calif., pp 431–591

Jeffreys AJ, Wilson V, Thein SL (1985) Individual specific "fingerprints" of human DNA. Nature 316:76–79

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

Murphy G, Kavanagh T (1988) Speeding-up the sequencing of double-stranded DNA. Nucleic Acids Res 16:5198

Nybom H, Schaal BA (1990) DNA "fingerprints" applied to paternity analysis in apples (*Malus×Domestica*). Theor Appl Genet 79:763–768

Plant Varieties Act 1987, No.2 of 1987, as enacted by the Queen, and the Senate and the House of Representatives of the Commonwealth of Australia

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275–280

Rogstad SH, Patton JC, Schaal BA (1988) M13 repeat unit probe detects DNA minisatellite like sequences in gymnosperms and angiosperms. Proc Natl Acad Sci USA 85:9176–9178

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Sanger F, Nickleu S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Scandalios JG (1974) Isozymes in development and differentiation. Annu Rev Plant Physiol 25:225–258

Schaal BA, O'Kane SL, Rogstad SH (1991) DNA variation in plant populations. Trends Ecol Evol 6:329–333

Scott NS, Possingham JV (1980) Chloroplast DNA in expanding spinach leaves. J Exp Bot 31:1081–1092

Spiegel-Roy P, Vardi A, Shani A (1977) Peroxidase isoenzymes as a tool for early separation of nucellar and zygotic *Citrus* seedlings. Proc Intl Soc Citricult 2:619–624

Tancred SJ, Zeppa AG, Graham GC (1994) The use of PCR-RAPD technique in improving the Plant Variety Rights description of a new Queensland apple (*Malus domestica*) cultivar. Aust J Exp Agric 34:665–667

Thuring RWJ, Sanders JPM, Borst P (1975) A freeze-squeeze method for recovering long DNA from agarose gels. Anal Biochem 66:213–220

Torres AM, Soost RK, Diedenhofen U (1978) Leaf isozymes as genetic markers in citrus. Am J Bot 65:869–881

Tzuri G, Hillel J, Lavi U, Haberfeld A, Vainstein A (1991) DNA fingerprint analysis of ornamental plants. Plant Sci 76:91–97

Vedel F, Delseny M (1987) Repetitivity and variability of higher plant genomes. Plant Physiol Biochem 25:191–210

Winberg BC, Zhou Z, Dallas JF, Mcintyre CL, Gustafson JP (1993) Characterisation of minisatellite sequences from *Oryza-sativa*. Genome 36:978–983