Theor Appl Genet (1997) 94: 249–254

P. Guilford · S. Prakash · J. M. Zhu · E. Rikkerink
S. Gardiner · H. Bassett · R. Forster

Microsatellites in *Malus X domestica* (apple): abundance, polymorphism and cultivar identification

Received: 20 February 1996 / Accepted: 24 May 1996

Abstract Screening of an apple genomic library with (GA)15 and (GT)15 probes demonstrated that these repeats are abundant, occurring about every 120 and 190 kb, respectively. Microsatellites isolated from a small insert library enriched for (GA) repeats contained numbers of repeats ranging from 7 to 39. Primers to these microsatellite loci were able to direct the amplification of the repeats in 21 different cultivars. The majority of markers were highly polymorphic, diploid, and showed simple Mendelian inheritance, although about 25% of markers generated complex banding patterns consistent with the amplification of more than one locus. As few as three microsatellite markers were sufficient to differentiate between all 21 cultivars.

Key words Microsatellites · Apple · Identification

Introduction

The difficulty of rapidly and reliably identifying apple cultivars is a considerable problem for the apple industry. Simple, definitive identification of cultivars would enable questions of identity to be eliminated prior to vegetative propagation, provide a means for identifying parents of chance seedlings, and establish a framework for the registration of new cultivars.

Traditional methods for cultivar identification are based on objective descriptions of tree and fruit characteristics. However, these observations are time consuming and error-prone due to environmental variations affecting the expression of these characteristics. Biochemical markers, such as isozymes, are more consistent (Weeden and Lamb 1985) but are not sufficiently informative to be widely applicable.

The development of DNA markers has, however, provided a wealth of polymorphisms to enable the rapid identification of cultivars and the construction of saturated genetic maps. Restriction fragment length polymorphisms (RFLPs) have been used with some success for apple cultivar identification and parentage analysis (Nybom 1990; Nybom and Schaal 1990; Nybom et al. 1992). However, they are technically laborious and have been surpassed by randomly amplified polymorphic DNA (RAPD) markers (Koller et al. 1993; Mulcahy et al. 1993; Landry et al. 1994). RAPDs are time and cost effective, but can be difficult to interpret due to their dominant mode of inheritance, and the complexity and occasional variability of the banding patterns (Lavi et al. 1994). In comparison, simple-sequence repeat markers (microsatellites) have the potential to provide a more reliable method for DNA fingerprinting of apples because of their high repeatability, relatively simple banding patterns, and co-dominant inheritance.

Microsatellites have become the genetic markers of choice in mammalian and many plant systems due to their abundance, high degree of polymorphism and amenability to automation (Weber and May 1989). The first genetic map of apple has recently been published (Hemmat et al. 1994). However, this map consists largely of RAPD markers which are not easily transported to genetic maps of unrelated cultivars. The co-dominant mode of inheritance of microsatellites not only simplifies analysis but facilitates the transfer of markers between genetic maps derived from different crosses. Therefore microsatellite markers should enable the integration of apple genetic maps from various sources.

The present study characterizes microsatellites in apple and demonstrates their abundance, polymorphism, and transportability between cultivars. We show that microsatellite markers can be used for the rapid and reliable DNA typing of apple cultivars with as few as three highly poly-

Communicated by J. Beckmann

P. Guilford (✉) · S. Prakash · J. M. Zhu · E. Rikkerink · R. Forster
Horticulture and Food Research Institute of New Zealand,
Mt Albert Research Centre, Private Bag 92 169,
Auckland, New Zealand

S. Gardiner · H. Bassett
Bachelar Research Centre, Private Bag 11 030,
Palmerston North, New Zealand

morphic microsatellites being capable of distinguishing between over 20 apple cultivars.

Materials and methods

Plant material and DNA purification

Apple leaves used for DNA purification were obtained from the *Malus* collection held by the Horticulture and Food Research Institute of New Zealand at Havelock North. Leaves (approximately 0.3 g) were frozen in liquid nitrogen then ground in 2.5 ml of extraction buffer [150 mM sorbitol, 20 mM Tris pH 7.5, 20 mM EDTA, 0.8 M NaCl, 0.8% (w/v) CTAB, 1% (w/v) N-lauroylsarcosine, 1% polyvinylpyrrolidone]. The slurry was mixed with 400 µl of chloroform:isoamyl alcohol (24:1) and incubated at 65°C for 30 min. After cooling on ice the mix was centrifuged at 15 000 *g* for 10 min and the supernatant precipitated with an equal volume of ice-cold propanol. The precipitate was then pelleted, washed in 70% ethanol, and resuspended in TE. After digestion with RNase A, contaminating polysaccharides were removed by passage through Sephacryl S-1000 (Pharmacia) and a Promega 'Wizard' column.

Screening of a phage lambda library for microsatellite repeats

A genomic *Malus X domestica* cv 'Granny Smith' DNA library cloned into lambda GEM11 was plated to a density of about 300 plaques/plate on a lawn of *E. coli* strain KW251 using standard techniques. Plaques were transferred to Hybond N+ filters (Amersham) and hybridized at 65°C in 0.5 M sodium phosphate pH 7.2, 1 mM EDTA and 7% SDS. Probes were oligonucleotides which had been end-labelled with ^{32}P-γATP. Washes were carried out in 2×SSC, 0.1% SDS at 65°C.

Identification, cloning and sequencing of microsatellite repeats

Small insert genomic libraries, enriched for GA- and GT-repeat microsatellite sequences derived from *Malus X domestica* cv 'Royal Gala' DNA, were produced by hybrid selection with immobilized oligonucleotides (GA)15 and (GT)15, respectively, essentially as described by Karagyozov et al. (1993). Selected DNA was cloned into pUC19 and used to transform *E. coli* DH5α according to standard techniques (Sambrook et al. 1989). Colonies were picked into 96-well microtiter plates and screened for the presence of repeats by colony hybridization to $(GA)_{15}$ and $(GT)_{15}$ which had been end-labelled with ^{32}P-γATP (Sambrook et al. 1989). Hybridizations were carried out at 65°C in 0.5 M sodium phosphate pH 7.2, 1 mM EDTA and 7% SDS (Church and Gilbert 1984). Washes were carried out twice for 20 min in 2×SSC, 0.1% SDS at 65°C.

Plasmid was purified from positive colonies using the Promega Wizard plasmid purification system and sequenced using T7 polymerase (Pharmacia) according to the manufacturer's instructions. To sequence PCR products, the products were first eluted from an agarose gel using Bresaclean (Bresatec Ltd.) and ligated into pGEM-T (Promega).

Microsatellite primer design and PCR

Primers were designed using PRIME from the GCG package. PCR was carried out under the following conditions; 50 ng DNA template, 1 µM each primer, 200 µM dNTPs, 0.5% Tween 20, 45 mM Tris HCl pH 8.8, 11 mM $(NH_4)_2SO4$, 6.7 mM 2-mercaptoethanol, 4.5 µM EDTA, 0.25 mM spermidine, 1 U *Taq* polymerase (Gibco BRL). After denaturation at 96°C for 5 min and addition of the *Taq* polymerase at 94°C (hot start), the reactions underwent 35 cycles of: 94°C, 40 sec; 50°C or 55°C (annealing), 40 s; and 72°C, 20 s; followed by a single step at 72°C for 10 min; 2.5 vol of 80% formamide containing 0.5% (w/v) bromophenol blue and xylene cyanol were then added to the PCR products. Samples were denatured at 94°C for 40 s, chilled on ice, then loaded on a 6% polyacrylamide/7 M urea sequencing gel. Following electrophoresis, the DNA was transferred to Hybond N+ membrane (Amersham) and hybridized with $(GA)_{15}$ which had been end-labelled with ^{32}P-dCTP using terminal deoxytransferase as previously described (Hazan et al. 1992). Autoradiography was carried out using Kodak XOmat K film at –70°C. In some PCR reactions, one primer was end-labelled with ^{33}P-γATP using T4 polynucleotide kinase (Sambrook et al. 1989) and the reaction products were analysed by direct autoradiography of the dried gel.

Results

Frequency of dinucleotide and trinucleotide repeats in the apple genome

To estimate the frequency of microsatellite repeats in the apple genome, 5000 plaques from a lambda genomic library of *Malus X domestica* cv 'Granny Smith' were hybridized under stringent conditions with the dinucleotide repeats $(GA)_{15}$ and $(GT)_{15}$. The frequency of hybridization of (AT) and (GC) dinucleotide probes was not determined because of the self-complementarity of these probes. The frequencies of two nominally chosen trinucleotide repeats, $(AAG)_{10}$ and $(CAG)_{10}$, were also estimated. The GA probe hybridized to 13% of plaques and the GT probe to 8% (Table 1). The trinucleotide repeats $(AAG)_{10}$ and $(CAG)_{10}$ occurred infrequently, hybridizing to fewer than 0.5% of the plaques. Since the average insert size in this library is about 15 kb, the GA and GT repeats occur in the apple genome about every 120 and 190 kb, respectively, and the different trinucleotide repeats about every 3 Mb.

Isolation and characterisation of microsatellite markers

To facilitate the isolation of GA-repeat microsatellites from *Malus X domestica* cv 'Royal Gala', an enriched, small-insert genomic library was constructed. Colonies containing GA repeats were identified by colony hybridization with ^{32}P-end labelled $(GA)_{15}$ at 65°C. Approximately 40% of colonies hybridized with the $(GA)_{15}$ probe. However, subsequent sequencing and hybridization with end-labelled primers demonstrated that the library was enriched with a strong bias to a limited number of unique sequences.

Table 1 The estimated occurrence of microsatellite sequence repeats in the genome of *Malus X domestica* cv 'Granny Smith'

Repeat	% Positive plaques	Average distance between microsatellites (kb)
$(GA)_n$	13	120
$(GT)_n$	8	190
$(CAG)_n$	<0.5	>3000
$(AAG)_n$	<0.5	>3000

Table 2 Oligonucleotide primer sequences and heterozygosity of apple microsatellites. PCR product lengths were calculated from the sequence of the cv 'Royal Gala' allele which was isolated from the enriched library. (a) Sequences of single locus microsatellites. (b) Sequences of 'complex', multi-loci microsatellites. Heterozygosity (H) was calculated using the formula $H=1-\Sigma p_i^2$, where p_i the frequency of the *i*th allele in the 21 cultivars examined. The PCR annealing temperature was 55°C, except for markers 01d7, 22c6, 23g4, 28f4, 03c1, 17e6 and 26c6 which were annealed at 50°C. N.D.=not determined

Marker	Primer sequences (5′–3′)	Cloned repeat	Predicted product length (bp)	No. of alleles detected	Heterozygosity
a					
01a6	agg att gct gga aaa gga gg tta gac gac gct act tgt cct	$(ga)_{21}$	136	6	0.75
01d7	ttc cat cct tct ccc ttc cac cta ctc tcc acc aca	$(ga)_{18}$	117	1	0
01d12	cgg act ata atc aac ctg aaa gt cgg ttt aac atg aca tca caa	$(ga)_{19}$	126	1	0
02b1	ccg tga tga caa agt gca tga atg agt ttg atg ccc ttg ga	$(ga)_{14}$	238	7	0.75
04h11	ctt cca tcg aga ttg cat cat a cga att gag agg tcg tcg tt	$(ga)_{23}$	225	6	0.65
05g8	cgg cca tcg att atc tta ctc tt gga tca atg cac tga aat aaa cg	$(ga)_{18}$	121	6	0.76
22c6	gac ctt tcc ctc tcc tga ctg gat atg att att gca ga	$(ga)_{12}$	111	3	0.57
23fl	gga aga gtg cga agg caa tgt ccc aac cac cgc ata	$(ga)_7$	105	2	0.36
23g4	ttt ctc tct ctt tcc caa ctc agc cgc ctt gca tta aat ac	$(ga)_{19}$	88	9	0.76
28f4	tgc ctc cct tat ata gct ac tga gga cgg tga gat ttg	$(ga)_{18}$	112	4	0.67
b					
03cl	gct ctc atc ttc aca gat aa aga ccc gga aaa ttc tat	$(ga)_{25}$	168	ND	ND
04f3	caa aac cac cct cat cct cga a ccc caa gca gac ctg aag aaa	$(ga)_{14}$	111	ND	ND
17e6	aac acg cca tca cac atc ctg ttt gct aga aga gaa gtc	$(ga)_{21}$	109	ND	ND
26c6	gac gaa gaa ctc gcc gga gc cga gga cca acc cac aca caa	$(ga)_{10}$	136	ND	ND

The sequence of 44 (GA)-containing inserts was determined. Twenty-eight contained perfect GA repeats ranging in number from 7 to 39, with a mean of 20 repeats. Eight positives colonies contained compound $(GA)_n(GT)_n$ repeats and one contained a $(GAAA)_{10}(GA)_{20}$ compound repeat. Seven positive clones contained imperfect (GA) repeats.

The enrichment procedure was repeated using immobilized (GT)15. The majority of (GT) repeat-containing colonies sequenced contained long, imperfect repeats. In many, the margins of the repeat were too poorly defined to enable specific primer design. As a consequence, the (GT) repeats were not analysed for their level of polymorphism.

Transportability and polymorphism of (GA) microsatellite loci

Fourteen primer pairs, first selected for their ability to amplify discrete, specific products in PCR reactions containing 'Royal Gala' DNA, were tested for their ability to direct amplification of microsatellite repeat sequences in different apple cultivars. PCR reactions were carried out with DNA from 21 randomly chosen apple cultivars (10 were 'modern' and 11 'old fashioned'). All of the markers tested successfully amplified the corresponding (GA) repeat in each of the 21 cultivars.

The majority of markers were polymorphic, with up to nine different alleles being observed in the 21 cultivars (Table 2a). Most individual cultivars were heterozygous for each of the markers shown in Table 2a. For example, only three and six of the 21 cultivars are homozygous for the markers 23g4 and 02b1, respectively (Fig. 1). Heterozygosity values, which were calculated using the genotypes of the 21 cultivars, averaged 0.66 for the polymorphic markers.

The segregation of the alleles of the markers shown in Table 2a was analysed in the cross of cv 'Royal Gala' and clone A172-2. Inheritance of the genotypes in the progeny of this cross was consistent with simple diploid, Mendelian inheritance. The segregation of the alleles of marker 02b1 is shown in Fig. 2.

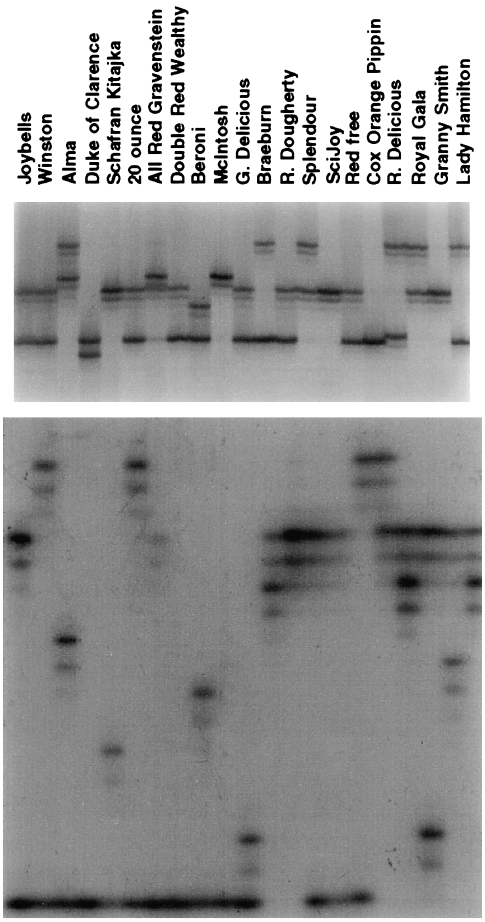


Fig. 1 Genotypes of 21 apple cultivars at the microsatellite loci 02b1 (upper panel) and 23g4 (lower panel)

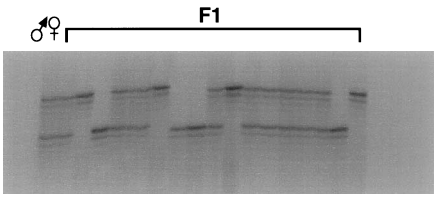
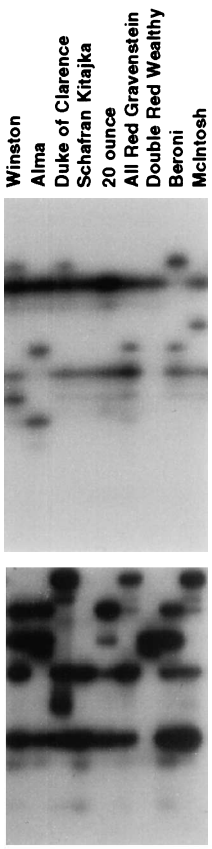


Fig. 2 Example of the simple Mendelian segregation of microsatellite alleles in apple. Parental lines and progeny from the cross cv 'Royal Gala' and clone A172-2 were amplified with marker 02b1

Since most commercial cultivars of *Malus X domestica* are functionally diploid, no more than two alleles would be expected for each cultivar. However, approximately 25% of (GA)-repeat markers gave more than two distinct PCR products for individual cultivars (Table 2b). The genotypes of two of these markers (03c1 and 26c6) with nine different cultivars are shown in Fig. 3. Both markers gave up to four distinct PCR products per cultivar. Segregation analysis of another complex marker, 04f3, which gave two clearly separated sets of biallelic products, was consistent with two independently segregating diploid markers (data not shown). These complex banding patterns may reflect ancestral chromosome duplication events.

Fig. 3 Example of polymorphic banding patterns with 'complex' apple microsatellite markers. Upper panel, 26c6; lower panel, 03c1



Identification of cultivars

The high degree of polymorphism and the simple Mendelian segregation of apple microsatellite markers provide a potential system for the identification and parentage analysis of apple cultivars. The genotypes of the 21 cultivars for six markers with heterozygosities greater than 0.65 is listed in Table 3. The allele lengths in Table 3 were determined by comparison to the lengths of the alleles of cv 'Royal Gala' and exclude the additional dATP nucleotide added to the 3' termini of PCR products by *Taq* polymerase. The allele lengths of cv 'Royal Gala' were determined by cloning and sequencing the microsatellite PCR products. Since 'slippage' of the *Taq* polymerase on the microsatellite repeat may have occurred during cloning of the microsatellites, the possibility of discrepencies in the absolute allele lengths given in Table 3 cannot be eliminated. Consequently, comparison of the allele lengths of unknown cultivars with those in Table 3 should be carried out using the amplification products of known apple cultivars as internal length standards.

Forty percent of the cultivars tested could be distinguished using the single marker 23g4. Complete differen-

Table 3 Cultivar identification of apple using microsatellites. The lengths (in base pairs) of the alleles of cv 'Royal Gala' were determined by cloning the PCR products of each of the markers into vector pGEM-T (Promega) and sequencing. The allele lengths of the other cultivars were derived by comparison with the 'Royal Gala' allele lengths on a 6% polyacrylamide sequencing gel using a pUC19 sequencing reaction for size reference

Cultivar	Genotype (bp)					
	01a6	02b1	04h11	05g8	23g4	28f4
'Joybells'	132/130	227/216	225/225	115/115	110/84	110/98
'Winston'	138/130	227/216	225/203	115/115	116/84	110/104
'Alma'	130/120	238/230	225/201	127/127	102/84	112/98
'Duke of Clarence'	132/130	216/212	225/225	121/121	84/84	110/98
'Schafran Kutajka'	138/132	227/227	225/207	123/121	94/84	112/98
'Twenty Ounce'	132/130	227/216	225/201	127/121	116/84	112/112
'All Red Gravenstein'	132/120	230/230	207/203	141/121	110/84	112/110
'Double Red Wealthy'	132/130	227/216	225/225	147/143	84/84	98/98
'Beroni'	132/130	223/216	225/201	121/121	98/84	98/98
'McIntosh'	132/132	230/230	225/205	127/123	84/84	112/110
'Golden Delicious'	132/130	227/216	207/203	121/121	88/84	112/98
'Braeburn'	138/136	238/216	225/207	141/139	110/106	112/104
'Red Dougherty'	130/128	227/216	225/207	121/121	110/108	112/98
'Splendour'	136/130	238/227	225/207	141/121	110/84	98/98
'SciJoy'	136/130	227/227	225/207	141/123	110/84	98/98
'Red Free'	130/128	227/216	225/225	123/123	116/84	112/110
'Cox's Orange Pippin'	130/128	216/216	225/207	123/123	116/110	112/110
'Red Delicious'	136/120	238/217	225/201	141/121	110/106	112/98
'Royal Gala'	136/132	238/227	225/207	121/121	110/88	112/112
'Granny Smith'	130/130	227/227	233/207	123/123	110/100	112/98
'Lady Hamilton'	138/120	238/216	225/207	141/139	110/106	112/104

tiation between the 21 different cultivars tested could be achieved by genotyping as few as three microsatellites (e.g. 23g4, 02b1 and 01a6).

Discussion

Since all apples of the same cultivar have originated from a single ancestor by vegetative propagation, analysis of polymorphic DNA sequences should produce banding patterns which are conserved among members of the same cultivar, but differ between cultivars. The transportability of apple microsatellites between cultivars and their high degree of polymorphism makes them an ideal marker system for the identification of apple cultivars. We show that three highly polymorphic microsatellites are capable of differentiating between all cultivars in our collection of 21 apple DNAs, and present a panel of six markers with an average heterozygosity of over 0.7 which will enable differentiation between the majority of cultivars.

Di- and tri-nucleotide microsatellites occur in the apple genome at a frequency comparable with other plant species. For example, the $(GA)_n$ motif occurs about every 120 kb, compared to about every 225 and 100 kb in rice and tropical tree genomes, respectively (Condit and Hubbell 1991; Wu and Tanksley 1993). The $(GT)_n$ motif is less common in apple, occurring about every 190 kb. The greater abundance of the GA motif compared to GT in apple has been observed in other plant species (Wu and Tanksley 1993; Lagercrantz et al. 1993) and contrasts with mammals in which the CA motif predominates (Beckmann and Weber 1992). The abundance, polymorphism and transportability of di-nucleotide microsatellites in apple will enable them to be readily integrated into genetic maps derived from diverse crosses (Hemmat et al. 1994; King 1994) and will facilitate the merging of these different maps.

Apples and crab apples (subfamily *Maloideae*) form one of the four subfamilies of the family *Rosaceae*. Compared to the other subfamilies *Prunoideae*, *Spiraeoideae* and *Rosoideae*, which have haploid chromosome numbers of 8, 9 and 7, respectively, the haploid chromosome number of the *Maloideae* is unusually high (x=17). It has been proposed that the basic haploid chromosome number in the *Rosaceae* is seven and that the members with haploid chromosome numbers of eight or nine have arisen by duplication of individual chromosomes (Challice 1974). The apple genome may have then derived from allopolyploidy between primitive members of the subfamilies *Prunoidae* and *Spiraeoideae* (Stebbins 1958). Therefore, it is probable that apple contains several chromosomes that are derived from the same common ancestor. Analysis of microsatellite markers will provide a means to identify the genome fusion and chromosome duplication events which may have occurred during the evolution of the apple genome. In this regard, it will be of interest to test the transportability of the markers described here to other members of the *Rosaceae* and examine the chromosomal localizations of the microsatellite markers which amplify more than one locus.

Acknowledgements This work was funded by FRST research grant 95-HRT-S06. We thank Leslie Beuning for the 'Granny Smith' genomic library.

References

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. Genomics 12:627–631

Challice JS (1974) Rosaceae chemotaxonomy and the origins of the Pomoideae. Bot J Linn Soc 69:239–259

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Condit R, Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34:66–71

Hazan J, Dubay C, Pankowiak M-P, Becuwe N, Weissenbach J (1992) A genetic linkage map of human chromosome 20 composed entirely of microsatellite markers. Genomics 12:183–189

Hemmat M, Weeden NF, Mangaris AG, Lawson DM (1994) Molecular marker linkage map for apple. J Hered 85:4–11

Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res 21:3911–3912

King GJ (1994) Progress in mapping agronomic genes in apple (The European Apple Genome Mapping Project). Euphytica 77:65–69

Koller B, Lehmann A, McDermott JM, Gessler C (1993) Identification of apple cultivars using RAPD markers. Theor Appl Genet 85:901–904

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and invertebrates. Nucleic Acids Res 21:1111–1115

Landry BS, Li RQ, Cheung WY, Granger RL (1994) Phylogeny analysis of 25 apple rootstocks using RAPD markers and tactical gene tagging. Theor Appl Genet 89:847–852

Lavi U, Cregan P, Schaap T, Hillel J (1994) Application of DNA markers for identification and breeding of perennial fruit crops. In: Janick J (ed) Plant breeding reviews. John Wiley and Sons, Inc, New York. vol 12: 195–226

Mulcahy DL, Cresti M, Sansavini S, Douglas GC, Linskens HF, Bergamini Mulcahy G, Vignani R, Pancaldi M (1993) The use of random amplified polymorphic DNAs to fingerprint apple genotypes. Scien Hort 54:89–96

Nybom H (1990) Genetic variation in ornamental apple trees and their seedling (*Malus*, Rosaceae) revealed by DNA fingerprinting with the M13 repeat probe. Hereditas 113:17–28

Nybom H, Schall BA (1990) DNA 'fingerprints' applied to paternity analysis in apples (*Malus X domestica*). Theor Appl Genet 79:763–768

Nybom H, Gardiner S, Simon CJ (1992) RFLPs obtained from a rDNA probe and detected with enhanced chemiluminescence in apples. HortScience 27:355–356

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Stebbins GL (1958) On the hybrid origins of the angiosperms. Evolution, 12:267–270

Weber JK, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–397

Weedon NF, Lamb RC (1985) Identification of apple cultivars by isozyme phenotypes. J Am Soc Hort Sci 110:509–515

Wu K-S, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241: 225–235