Theor Appl Genet (1997) 95: 911–921 

**D. Sharon · P. B. Cregan · S. Mhameed**
**M. Kusharska · J. Hillel · E. Lahav · U. Lavi**

# An integrated genetic linkage map of avocado

Received: 2 April 1996 / Accepted: 28 February 1997

**Abstract** An avocado genomic library was screened with various microsatellite repeats. $(A/T)_n$ and $(TC/AG)_n$ sequences were found to be the most frequent repeats. One hundred and seventy-two positive clones were sequenced successfully of which 113 were found to contain simple sequence repeats (SSR). Polymerase chain reaction primers were designed to the regions flanking the SSR in 62 clones. A GenBank search of avocado DNA sequences revealed 1 sequence containing a $(CT)_{10}$ repeat. A total of 92 avocado-specific SSR markers were screened for polymorphism using 50 offspring of a cross between the avocado cultivars 'Pinkerton' and 'Ettinger'. Both are standard avocado cultivars which are normally outcrossed and highly heterozygous. Fifty polymorphic SSR loci, 17 random amplified polymorphic DNA (RAPD) and 23 minisatellite DNA Fingerprint (DFP) bands were used to construct the avocado genetic map. The resulting data were analyzed with various mapping programs in order to assess which program best accommodated data from progeny of heterozygous parents. The analyses resulted in 12 linkage groups with 34 markers (25 SSRs, 3 RAPDs and 6 DFP bands) covering 352.6 cM. This initial map can serve as a basis for developing a detailed genomic map and for detection of linkage between markers and quantitative trait loci.

**Key words** Linkage groups · Mendelian inheritance · Microsatellites · Minisatellites · Simple sequence repeats · Variable number of tandem repeats

Contribution from the Agricultural Research Organization, The Volcani Center, Bet Dagan, Israel. No. 1865-E, 1996 series

Communicated by A. L. Kahler

D. Sharon · S. Mhameed · E. Lahav · U. Lavi (✉)
Department of Fruit Trees, Genetics and Breeding, Agricultural Research Organization, P.O. Box 6, Bet-Dagan 50250, Israel

P. B. Cregan
USDA, ARS, Soybean and Alfalfa Research Laboratory, Beltsville MD 20705, USA

M. Kusharska
Department of Genetics, The Silesian University, Jagiallonska 28, Katovice 40032, Poland

J. Hillel
Department of Genetics, Faculty of Agriculture, The Hebrew University of Jerusalem, P.O. Box 12, Rehovot 76100, Israel

## Introduction

Avocado (*Persea americana* Mill.) is a sub-tropical fruit-tree, with 24 chromosomes (n = 12) and a haploid genome size of $8.83 \times 10^8$ bp (Arumuganatham and Earle 1991). Genome analysis and breeding of sub-tropical fruit trees, especially avocado, are quite difficult mainly because of the size of the trees, a long juvenile phase, and a lack of satisfactory genetic knowledge.

In the last few years, four types of genetic markers have been used in avocado (reviewed by Lavi et al. 1994b). Degani et al. (1990) used mainly isozyme markers to study the level of selfing and outcrossing in the cultivar 'Fuerte'. Furnier et al. (1990) used two restriction fragment length polymorphism (RFLP) loci to study the relationships between *Persea* species and between avocado cultivars. DNA fingerprints (DFP), resulting from the use of multilocus minisatellite and microsatellite probes, were applied for identification of cultivars and races (Lavi et al. 1991) and for detection of linkage between a locus affecting fruit skin color and a specific DFP band (Mhameed et al. 1995). Simple sequence repeat (SSR) markers have been shown to be highly polymorphic and to generally segregate according to Mendelian laws of inheritance (Lavi et al. 1994a). Both DFP and SSR markers were used to estimate the heterozygosity level in the avocado genome (Mhameed

et al. 1996) and to detect genetic relationships in the *Persea* genus (Mhameed et al. 1997). None of these markers has thus far been used to construct a genetic linkage map of the avocado genome.

The construction of genetic linkage maps is an essential step for several purposes, such as studies of genome structure, map-based cloning and quantitative trait locus (QTL) mapping. Botstein et al. (1980) proposed systematic molecular genetic mapping with RFLP markers in humans. The genomes of a large number of species have been mapped using DNA markers (Akkaya et al. 1995; Bell and Ecker 1994; Bernatzky and Tanksley 1986; Helentjaris et al. 1986; Hudson et al. 1995). Mapping the genome of fruit trees is complicated in most species due to the high level of heterozygosity and the inability to create inbred lines. Genetic maps of only a few fruit trees have been constructed, including that of *Prunus* which has 9 linkage groups and 107 markers (RFLPs and isozymes) covering 800 cM (Foolad et al. 1995). Additionally, the map of apple (Hemmat et al. 1994) has 24 linkage groups with 253 isozyme and RFLP markers and a total length of 950 cM while that of *Citrus* contains 11 linkage groups with 62 isozyme and RFLP loci (Durham et al. 1992).

Tautz (1989) suggested the use of simple sequence repeat markers (microsatellites) detected via the polymerase chain reaction (PCR) as single-locus, highly polymorphic markers. These two characteristics and the high reproducibility of SSR markers have made them the preferred marker in many different organism, including human (Hudson et al. 1995), animals (Dietrich et al. 1992) and plants. In plants, the use of SSR markers was first demonstrated by Akkaya et al. (1992), but as yet they have not been applied on a scale comparable to their usage in the human or other mammalian species. Rongwen et al. (1995) used 7 SSR loci to fingerprint 96 soybean genotypes, while Akkaya et al. (1995) integrated 40 SSR markers into a soybean linkage map. Senior and Heun (1993) developed 6 SSR markers from GenBank data in maize, studied their polymorphism and confirmed the co-segregation of the SSR loci with the isozyme coded by the same locus. More recently, Senior et al. (1996) mapped 42 distinct GenBank-derived SSR loci in maize. Becker and Heun (1995) mapped 5 SSR loci in the barley RFLP map, Bell and Ecker (1994) reported the assignment of 30 microsatellite loci to the linkage map of *Arabidopsis* and Veilleux et al. (1995) used 5 SSR markers to characterize the genetic composition of anther-derived potato. Preliminary studies of SSR markers in fruit trees have also been reported. Kijas et al. (1995) developed 2 SSR markers and found them to be polymorphic between *Citrus* and related species. Thomas and Scott (1993) isolated 5 microsatellite loci from grapevine and found that the heterozygosity level of these loci ranged between 69% and 88%. Lavi et al. (1994a) generated 30 avocado dinucleotide microsatellite markers and demonstrated the high level of polymorphism associated with 2 of them.

In this report we present the generation of avocado-specific SSR markers. These markers were integrated with random amplified polymorphic DNA (RAPD) and DFP markers to construct the first genetic linkage map of avocado. The map construction of outcrossing species is presented and discussed.

## Materials and methods

### Plant material

Trees were grown at the Akko Experiment Station and the Agricultural Research Organization, Bet Dagan, Israel. The cross between the female parent 'Pinkerton' and the pollen donor 'Ettinger' was made by caging the two parents under a net during the flowering season in the presence of a beehive. Leaves were taken from the parents and 50 offspring of this cross for DNA isolation.

### DNA isolation

DNA was isolated from very young leaves using the CTAB (cetyltrimethyl ammonium bromide) method (Murray and Thompson 1980), with modifications as described by Lavi et al. (1991).

### DNA fingerprinting (DFP) with the minisatellite probe 22.3

Ten micrograms of DNA were digested with *Hin*fI restriction endonuclease according to the manufacturer's recommendations (New England Biolabs, Beverly, Mass.). Electrophoresis was carried out on a 20-cm-long 0.8% agarose gel in TBE (0.045 *M* TRIS borate, 0.001 *M* EDTA pH = 8.0) at 1.5 V/cm for 40 h. Gels were Southern-blotted onto Hybond-$N^+$ membranes (Amersham Int, Amersham, UK). Lambda *Hin*dIII fragments (New England Biolabs, Beverly, Mass.) served as DNA fragment size standards.

Membranes were hybridized to the double-stranded probe 22.3, a 2-kb probe that includes probe R18.1 (Haberfeld et al. 1991). Probe 22.3 was labeled with $\alpha$-[$^{32}$P]-dCTP by random priming (Feinberg and Vogelstein 1983). Pre-hybridization and hybridization were carried out at 65°C for 3–5 h and 16 h, respectively, in 0.263 *M* Na-phosphate buffer with 7% sodium dodecyl sulfate (SDS), 1 m*M* EDTA and 1% bovine serum albumin (BSA).

Membranes were washed with 0.263 *M* Na-phosphate and 1% SDS for 20 min, then twice with 2 × saline sodium citrate (SSC), and 0.1% SDS for 20 min each and twice with 1 × SSC and 0.1% SDS for 20 min each. Washes were carried out at 65°C. Membranes were exposed to X-ray film for 1–10 days at − 80°C with intensifying screens.

### RAPD markers

Reaction mixes contained 30 ng genomic DNA, 1.5 m*M* $Mg^{2+}$, 0.2 μ*M* of primer, 250 μ*M* of each nucleotide, 1X Taq buffer [containing 20 m*M* $(NH_4)SO_4$, 75 m*M* TRIS-HCl pH 9.0, 0.1% Tween] and 0.5 unit of *Taq* DNA polymerase (Advanced Biotechnologies, Leatherhead, Surrey, UK) in a total volume of 25 μl. The reaction mixes were overlaid with 40 μl of mineral oil. Reactions were denatured at 94°C for 3 min before cycling, which consisted of a 1-min denaturation at 94°C, 1-min annealing at 36°C and a 1-min extension at 72°C for 44 cycles on a MJ Research model PTC-100 thermocycler (MJ Research, Watertown, Mass.). The cycles were followed by an extension step for 5 min at 72°C. Amplification

products were analyzed by electrophoresis on 1.5% agarose gels and detected by staining with ethidium bromide. The RAPD primers used are from set no. 7 (numbered 601-700) of the University of British Columbia (Vancouver, B.C., Canada).

### Genomic library construction and screening for SSR-containing clones

Genomic DNA of avocado cv 'Pinkerton' was cut with the restriction enzyme *Sau*3AI (Boehringer Mannheim, Mannheim, Germany). The DNA fragments were first size-selected (400–600 bp) on agarose gels, then cloned in pBluescript II $KS^+$ (Stratagene, LaJolla, Calif.) and transformed into *E. coli* Hinv-α strain. Transformants were grown in LB media with 50 μg/ml ampicillin. White colonies were transferred into microtiter plates and stamped onto Hybond-$N^+$ membranes (Amersham Int, UK). Colonies were grown, lysed, denatured and neutralized. Membranes were hybridized to end-labeled γ-[$^{32}$P]-oligonucleotide probes (Fritsch et al. 1989). The probes were: $(A)_{30}$, $(CT)_{10}$, $(GT)_{10}$, $(CAC)_6$, $(TCC)_6$, $(TCT)_6$, $(GATA)_4$, $(GACA)_4$ and $(GGAT)_4$.

### Sequencing of positive clones

Plasmid DNA was isolated from positive clones by WIZARD columns (Promega, Madison, Wis.). Clones were sequenced with the KS or SK primer using an ABI 373 DNA Sequencer (Applied Biosystems Int, Foster City, Calif.). In a few cases clones were sequenced from both sides of the insert because sufficiently good sequence data were not obtained for both flanking regions of the SSR as a result of sequencing from only one side of the insert.

### SSR definition

Sequences of positive clones were searched for any type of repeat. A repeat was identified as an SSR if the number of repeats (*n*) was more than 14 for a mononucleotide repeat, more than 6 for dinucleotide repeats, more than 3 for trinucleotide repeats and more than 2 for tetranucleotide (or more) repeats.

### GenBank search

Avocado sequences in the GenBank database were searched for the presence of all possible mono-, di-, and trinucleotide SSRs. The cellulase gene (accession no. X59944) sequence (Cass et al. 1990) was found to contain a $(CT)_{10}$ repeat in the 5′ region. The sequence data in the regions flanking the SSR were used to design polymerase chain reaction (PCR) primers as described below.

### Selection and synthesis of primers

SSR-containing clones were analyzed using Primer Detective Software (Clonetech Labs., Palo Alto, Calif.) or PRIMER 0.5 (Whitehead Institute for Biomedical Research, Cambridge, Mass.). Primers having melting temperatures of at least 60°C and which would amplify products with predicted lengths of 75–210 bp were synthesized on an Applied Biosystems International (ABI) 391 DNA synthesizer.

### PCR of SSR markers

Each pair of PCR primers selected from the regions flanking SSRs were examined in PCR amplification reactions using avocado genomic DNA as template. Reaction mixes contained 30 ng of genomic DNA, 1.5 m*M* $Mg^{2+}$, 0.15 μ*M* of 3′ and 5′ end primers, 100 μ*M* of each nucleotide, 200 mg/ml BSA, 0.1 μl of 111 TBq/mmol α-[$^{32}$P]-dCTP, 1X Taq buffer (containing 50 m*M* TRIS-HCl pH 9, 0.1% Triton X-100), and one unit of *Taq* DNA polymerase in a total volume of 10 μl. The reaction mixes were overlaid with 20 μl of mineral oil. Reactions were denatured at 94°C for 30 s before cycling, which consisted of a 15-s denaturation at 95°C, a 25-s annealing at 45°C or 50°C (depending on the primers), and a 25-s extension at 68°C for 32 cycles on a MJ Research model PTC-100 thermocycler (MJ Research, Watertown, Mass.). Cycles were followed by an extension step for 2 min at 68°C. Ten microliters loading buffer (95% formamide, 0.02 *M* EDTA pH = 8, 0.1% bromophenol blue, 0.1% xylene cyanol and 10 m*M* NaOH) was added, and after denaturation (2 min at 95°C), 3 μl of the reaction was loaded on a DNA sequencing gel containing 6% polyacrylamide, 8 *M* urea and 1 × TBE at 50 watts constant power for 3–4 h. Gels were dried and exposed overnight to Fuji X-ray film. The length of the SSR alleles was determined using a sequence ladder of M13 as a size standard. This procedure is essentially identical to that described by Cregan et al. (1994).

### Computer programs for map construction

A computer program with the following characteristics was needed to map the avocado genome with DNA markers: (1) analysis of up to 4 alleles at a locus in a single family, (2) analysis of one full-sib family with 50 offspring or more, (3) multi-point linkage analysis and (4) analysis of unknown phase of parental alleles. The most common program for map construction in plants is MAPMAKER, which analyses data from controlled crosses with a known linkage phase. MAPMAKER will only accommodate $F_2$, backcross or recombinant inbred data and cannot analyze loci with three or four alleles per locus. The LINKEM program can analyze up to four alleles at a locus but cannot perform multi-point analysis. This program is capable of determining the linkage phase (coupling or repulsion) in the parents. It also performs two-point, goodness-of-fit analysis which is more suitable for markers with segregation that deviates from Mendelian expectation. The JOINMAP program uses a modified LOD score which is based on the $\chi^2$-test for independence of segregation, resulting in different LOD score values for markers with segregation distortion. Other linkage programs such as LINKAGE, CRI-MAP and MULTIMAP were also examined. These programs (used by human geneticists) can analyze up to four alleles per locus with unknown linkage phase but are not capable of analysis of a single family with 50 offspring.

### Map construction

Each of the two parents and 50 offspring were genotyped using 50 SSR loci, 17 RAPD bands and 23 DFP bands. The segregation of the alleles from each marker was examined for deviation from Mendelian expectations using the LINKEM program (Vowden et al. 1995). Each marker was classified into one of four groups (Groover et al. 1994): MI (Maternal Informative) – loci which are informative for the maternal gametes only (the paternal genotype is homozygous): PI (Paternal Informative) – loci which are informative for the paternal gametes only (the maternal genotype is homozygous); BI (Both Informative) – loci which are heterozygous in both parents with the same two alleles; FI (Fully Informative) – both parents are heterozygous and the number of alleles is three or four. The following maps were constructed: (1) maternal map – contains markers of groups MI, BI and FI; (2) paternal map – contains markers of groups PI, BI and FI; (3) combined map – contains markers of all groups. For each map, two-point analysis of all possible combinations was performed using the computer programs:

LINKEM, LINKAGE 5.1 (Lathrop et al. 1984), CRI-MAP (Lander and Green 1987), MAPMAKER/EXP 3.0 (Lander et al. 1987) and JOINMAP (Stam and Van Ooijen 1995). In order to make the avocado mapping data (two to four alleles per locus) compatible with MAPMAKER input (two alleles per locus and specific cross-$BC_1$ or $F_2$), the allelic phase of each linked marker-pair was determined using the LINKEM program and the family was analyzed as a backcross family ($BC_1$). Linkage was determined when the LOD score was greater than 3.0. Linkage analysis of loci that showed significant deviation from Mendelian expectation was tested using the $\chi^2$ goodness-to-fit test in addition to the LOD score calculation. The final map was drawn using MAPMAKER results and was verified by JOINMAP, LINKAGE and MULTIMAP (Matise et al. 1994).

## Results

### Abundance of microsatellites in the avocado genome

An avocado genomic library containing about 10,000 clones with a mean insert size of 500 bp was screened with various SSR repeats (Table 1). This library represents about 0.57% of the avocado genome ($10^4$ clones × 500 bp/883 Mbp in the avocado genome). The frequency assessment of SSRs, based upon positive hybridization signal, indicates that $(A/T)_n$ and $(TC/AG)_n$ sequences are the two most frequent microsatellites, occurring at 1 per 40 kb and 1 per 68 kb, respectively, in the avocado genome. Estimates based upon hybridization may overestimate positive clones. Sequence data suggested that $(A/T)_n$ and $(TC/AG)_n$ microsatellites occur at 1 per 86 and 1 per 128 kb, respectively. These are conservative estimates because 53 of the 225 clones that were sequenced did not produce reliable sequence data. Re-isolation of plasmid and re-sequencing would probably detect additional SSRs among this group of clones. Based upon the hybridization data, we identified 236 positive clones, suggesting the presence of 45,200 microsatellite sequences in the avocado genome (Table 1).

### Generation of SSRs

Of the 225 positive clones sequenced (from one end), 172 gave reliable sequence data (Table 2). Of these 172 clones, 12 (7%) did not contain any type of repeat, the sequences of 47 clones (27%) contained short repeats and 113 clones (66%) contained an SSR. Suitable primers could not be selected in 51 (45%) of the SSR-containing clones (Table 3) due to the location of the repeat near the cloning site (41 clones) or a sub-optimal melting temperature detected during PCR primer selection (10 clones), leaving 62 (55% of the SSR containing clones) SSR primer pairs (Table 3).

**Table 1** Frequency of nine repeat types in the avocado genomic library

| Repeat type | Number of screened clones | Number of positive clones (%) | Predicted ratio of SSRs per kb[a] | Predicted no. of SSRs in the genome[b] |
|---|---|---|---|---|
| (A/T)n | 10,000 | 125 (1.25) | 1:40 | ~ 22,000 |
| (TC/AG)n | 10,000 | 74 (0.74) | 1:68 | ~ 13,000 |
| (GT/CA)n | 10,000 | 27 (0.27) | 1:185 | ~ 4,800 |
| (CAC/GTG)n | 5,000 | 3 (0.06) | 1:833 | ~ 1,050 |
| (TCC/AGG)n | 5,000 | 2 (0.04) | 1:1,250 | ~ 700 |
| (TCT/AGA)n | 5,000 | 2 (0.04) | 1:1,250 | ~ 700 |
| (GATA/CTAT)n | 3,000 | 0 (0) | < 1:1,500 | < 590 |
| (GACA/CTGT)n | 3,000 | 0 (0) | < 1:1,500 | < 590 |
| (GGAT/CCTA)n | 3,000 | 3 (0.1) | 1:500 | < 1750 |
| Total | 10,000 | 236 (2.36) | | ~ 45,200 |

[a] Predicted ratio of SSRs per kilobase is based on positive clones and was calculated as follows: (no. of screened clones × insert average length in kb)/no. of positive clones
[b] Predicted number of SSRs in the genome is based on positive clones and was calculated as follows: ($8.8 \times 10^5$ kb × no. of positive clones)/(no. of screened clones × insert average length in kb)

**Table 2** From positive clones to SSR-containing clones

| | A/T | TC/AG | GT/CA | Others[a] | Compound[b] | Total | Percent |
|---|---|---|---|---|---|---|---|
| Number of positive clones | 119 | 66 | 21 | 9 | 10 | 225 | |
| Good sequence | 89 | 51 | 15 | 9 | 8 | 172 | 100 |
| No repeat | 2 | 2 | 2 | 5 | 1 | 12 | 7 |
| Short repeat | 29 | 10 | 5 | 1 | 2 | 47 | 27 |
| SSR-containing clones[c] | 58 | 39 | 8 | 3 | 5 | 113 | 66 |

[a] Various types of repeats (such as AT, TCC, TCT)
[b] More than one type of repeat in a clone
[c] SSR was defined by sequencing analysis if the number of repeats ($n$) was > 14 for mononucleotide repeats, > 6 for dinucleotide repeats, > 3 for trinucleotide repeats and > 2 for tetranucleotide repeats

**Table 3** From SSR-containing sequences to primers

| | A/T | TC/AG | GT/CA | Others[a] | Compound[b] | Total | Percent |
|---|---|---|---|---|---|---|---|
| SSR-containing sequences[c] | 51 | 33 | 6 | 9 | 14 | 113 | 100 |
| SSR near the cloning site | 15 | 14 | 2 | 4 | 6 | 41 | 36 |
| Primers not found | 6 | 2 | 2 | 0 | 0 | 10 | 9 |
| Primers were found | 30 | 17 | 2 | 5 | 8 | 62 | 55 |

[a] Various types of repeats (such as AT, TCC, TCT)
[b] More than one type of repeat in a clone
[c] Based on sequencing analysis (Table 2)

**Table 4** Typing the progeny of the cross 'Pinkerton' × 'Ettinger'

| | A/T | TC/AG | GT/CA | Others[a] | Compound[b] | Total | Percent |
|---|---|---|---|---|---|---|---|
| Currently available[c] | 29 | 36 | 3 | 5 | 19 | 92 | 100 |
| No PCR product | 0 | 3 | 0 | 1 | 1 | 5 | 5 |
| Non-reliable product | 7 | 4 | 0 | 0 | 3 | 14 | 15 |
| Non-informative | 9 | 7 | 2 | 1 | 6 | 25 | 27 |
| Typed | 13 | 22 | 1 | 3 | 9 | 48 | 52 |

[a] Various types of repeats (such as AT, TCC, TCT)
[b] More than one type of repeat in a clone
[c] The currently available avocado SSR markers consist of 62 markers detailed in Table 3, 29 SSR markers from a previous avocado library (Lavi et al. 1994a) and 1 SSR from the 5′ region of the cellulase gene (GenBank data)

In addition to the 62 SSR primer sets developed in this study, 29 primer sets from a previous avocado library (Lavi et al. 1994a) and 1 from the 5′ region of the cellulase gene (GenBank data) were used to select polymorphic markers for avocado linkage analysis. Each parent of the mapping population ('Pinkerton' and 'Ettinger') was screened with each of the 92 primer pairs. Five SSRs gave no PCR product (Table 4), 14 gave non-reliable or multiple products, 25 gave non-informative patterns (non-polymorphic, or both parents homozygous for different alleles). Forty-eight of 92 SSRs were used for typing the 50 progeny of the mapping population. Two primer pairs amplified products from 2 different loci each, and thus a total 50 loci were scored. The SSR locus names, type of repeat(s), sequence of the primers, annealing temperature, allele size, and group are listed in Table 5. If more than one type of repeat was amplified, the repeat type was referred to as 'Compound'. The primers are 18–26 bases long and the lengths of the PCR products vary between 76 to 224 bp. About 24% of the loci are maternal-informative (12 loci), 36% are paternal-informative (18 loci), 14% are both-informative (7 loci) and 26% are fully-informative (13 loci). All 50 SSR loci can be used for the combined map construction, while 38 can be used for the construction of the paternal map and 32 loci can be used for construction of the maternal map.

## Avocado linkage map

The genotypes of the two parents ('Pinkerton' and 'Ettinger') and 50 offspring were determined using 50 SSRs, 17 RAPDs and 23 DFP bands (Mhameed et al. 1995). Of the DFP bands 13 are maternal-specific and 10 are paternal specific. The segregation of alleles in the mapping family using the marker AVAG07 is presented in Fig. 1. AVAG07 is a full-informative marker – both parents are heterozygotes and carry different alleles. All 90 loci were used for two-point linkage analysis using the following programs: CRI-MAP, JOINMAP, LINKAGE, LINKEM and MAPMAKER. The two-point analysis of LINKAGE and CRI-MAP programs resulted in identical LOD scores and *theta* values, and the two-point analysis of LINKEM and MAPMAKER resulted in identical LOD scores and *theta* values. The last two programs gave higher values of LOD score (about 10% higher compared to the results obtained with LINKAGE and CRI-MAP). The two-point analysis of JOINMAP and MAPMAKER resulted in similar LOD score values (except for markers with distorted segregation). The paternal, maternal and combined maps are presented in Fig. 2. The combined map consists of 12 linkage groups and 34 markers. The number of markers per linkage group varies between two and five, and the length of the 12 linkage groups varies from 14.7 cM to 68.7 cM. Linkage groups 2 and 3 combined data from both parents, linkage groups 4, 6, 9 and 12 are maternal-specific and linkage groups 1, 5, 7, 8, 10 and 11 are paternal-specific. Linkage group 1 consists of DFP loci only, linkage groups 2, 3, 4, 5, 9, 10, 11 and 12 consist of SSR loci, linkage groups 7 and 8 consist of SSR and RAPD markers and linkage group 6 consists of DFP and RAPD markers. The length of linkage group 3 varies significantly between the paternal and maternal maps (18.7 cM vs. 50.4 cM). Analysis of marker distribution

**Table 5** List of the avocado SSR markers

| Serial no. | Locus name | Repeat type | Primers sequence | | Annealing temperature[a] | Allele size in basepairs | | Group[b] |
|---|---|---|---|---|---|---|---|---|
| | | | Forward (5′ → 3′) | Reverse (5′ → 3′) | | Pinkerton | Ettinger | |
| 1 | AVAG01 | $(TC)_{24}$ | TTCTATGTCAATGCGGTCAG | GTTCCTTAGCTGTTTCATAT | 40 | 163,179 | 169,181 | FI |
| 2 | AVAG02 | $(AG)_{25}$ | CATCATGGTGTTTGAATGCC | TGGTGACCTTAATCTACCCTCC | 45 | 107,143 | 147,149 | FI |
| 3 | AVAG03 | $(TC)_{17}$ | GCACTTCCTAAACTTGCAGGT | CTGAACATCCAATGACAAACATCC | 45 | 98,106 | 96 | MI |
| 4 | AVAG04 | $(CT)_{14}$,$(CT)_{12}$ | TCGACTGATTTCTCAGAAGCCC | ACATAGAGACGTGAGCCTACC | 45 | 147,161 | 141,147 | FI |
| 5 | AVAG05 | $(AG)_{10}$ | GGATCTGATGTGTGGGGGAG | CCTGTCGGAAAAGACTATGCG | 45 | 95,99 | 93,103 | FI |
| 6 | AVAG06 | $(CT)_{18}$ | CGACCTCTTCTTATACTC | GTACCTCTGATAATGAGCAT | 40 | 77,85 | 81,87 | FI |
| 7 | AVAG07 | $(TC)_{15}$ | ATCCAAAATGCACAAGGTGAGG | TGTCGCTATGTCCAAAATGTGG | 50 | 110,113 | 108,120 | FI |
| 8 | AVMIX01 | $(AT)_{7}$,$(AG)_{12}$ | GCCGCAAGGACAACAGCT | TGCCATGCCGCAATGAAC | 45 | 183 | 181,183 | PI |
| 9 | AVAG08 | $(TC)_{19}$ | CAGTTAATCTAGCCATTC | TCATGTTTGGTGGACTGAG-TTTTCAA | 45 | 113,117 | 113,121 | FI |
| 10 | AVAG09 | $(AG)_{12}$ | TGGCTAGGTCTAGGTCTAGGG | TCCTGCTTTTGACGGCTG | 55 | 84 | 76,84 | PI |
| 11 | AVMIX02 | $(TC)_{6}$,$(TCC)_{4}$ | GAGTCACGCTCGTAGGCT | TATAAATTCAAATGACAC | 40 | 153,171 | 153,171 | BI |
| 12 | AVGC01 | $(GC)_{12}$ | GAACTTAGTATGAGAGAA | AAATGGAAACCCATCGGTC | 40 | 196 | 192,196 | PI |
| 13 | AVAG10 | $(CT)_{22}$ | GAATTACAAAGCACTAGAG | GTAGAAAGTGGGCACACAT | 45 | 149,153 | 153 | MI |
| 14 | AVMIX03 | $(TG)_{16}$,$(AG)_{20}$ | GATATTCCTGTTGTCACTGC | AATGTTCCCCATGAAAGTCTCC | 50 | 146,170 | 146,174 | FI |
| 15 | AVMIX04 | $(AG)_{12}$, $(CAA)_{5}$ $(ACAG)_{10}$ | CCGTTTGCTTCCTGTATC | GTTATCCCTTCCACTTTC | 50 | 180 | 164,170 | PI |
| 16 | AVAG11 | $(AG)_{20}$ | AGCGATGAACATTACCCA | ATTTCTTCAACCCATCTGTC | 50 | 107 | 109,115 | PI |
| 17 | AVAG13 | $(CT)_{18}$ | CTGCGATAACAACTGGAC | AACTAGGACCTGAAACCG | 50 | 130 | 134,140 | PI |
| 18 | AVAC01 | $(TG)_{15}$ | CTGGTTGCTCTCTTGTCTACATAATA | CGGTTTTGTAAGTTGATAG | 40 | 117 | 113,117 | PI |
| 19 | AVAG14 | $(TC)_{15}$ | CTAGACGACCCTGACACT | CGGTTGTAGAGGAATAAG | 45 | 142 | 142,154 | PI |
| 20 | AVTCT01 | $(TCT)_{9}$ | GATTACATCCAAGGTTGG | AGATCGTTCCTTATACCAGTGG | 45 | 198,210 | 210 | MI |
| 21 | AVAG151 | $(CT)_{20}$ | CAAAGTTGCCTCCTCATTTACA | TACACACAACCACGTCCTCTAA | 45 | 120 | 120,128 | PI |
| 22 | AVAG1511 | | | | | 147,151 | 151 | MI |
| 23 | AVAG161 | $(CT)_{12}$ | TGCAAGTACCTGTAAAAAAGAG | CTACAACTCTAACTACCTTAAC | 55 | 120 | 116,120 | PI |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 24 | AVAG1611 | | | | | 100,106 | 100 | MI |
| 25 | AVAG17 | $(CT)_{19}$ | GATAAGCATGAAAAAATCCATA | ACTCGGAGCTGGGGATGGTGGC | 45 | 119,149 | 119,149 | BI |
| 26 | AVA01 | $(T)_{20}$ | CTAGCATTGCTGAGTGTTAGTC | TTAGGTTGTTCGTTGTTGCCAT | 50 | 125 | 125,132 | PI |
| 27 | AVAG18 | $(GA)_{10}$ | AAAAGGGTTGCAGGTATGGC | GTAGTTCAGGAAAGTGTCACT | 60 | 121 | 121,128 | PI |
| 28 | AVAG19 | $(GA)_{16}$ | ATCCACTTACCACTCCCCTATT | AGTCTGCTCACATTGGTGGG | 50 | 178,186 | 164,178 | FI |
| 29 | AVAG20 | $(CT)_{10}$,$(GA)_{14}$ | GACAAGGTTCTTATCGAGATGC | CGATGAAACCGCTCCTC | 50 | 159,172 | 159,172 | BI |
| 30 | AVAG21 | $(CT)_{22}$ | TGTAAGTTTTAACCCCACAA | AATCACTATTAGAGTTTTCAGTCG | 50 | 192,200 | 172,184 | FI |
| 31 | AVAG22 | $(GA)_{15}$ | GATCATCAAGTCCTCCTTGG | GATCTCATAGTCCAAATAATGC | 55 | 117 | 103,117 | PI |
| 32 | AVAG23 | $(TC)_{15}$ | CATCTAAAACAGGCATCGAGC | CTTGTCTCCAGTGAGCGTGA | 50 | 211,224 | 224 | MI |
| 33 | AVAG24 | $(TC)_{13}$ | AGAAACCCCCTCACCAACTC | TAGAGAGCCCAGATGAGGGA | 45 | 98,100 | 98,100 | BI |
| 34 | AVA02 | $(A)_{20}$,$(T)_{15}$ | ATCATAATAGCCAGCAACGTC | GGGGTGGGACTCCTCATTTT | 45 | 96 | 100,124 | PI |
| 35 | AVGGGGAT01 | $(GGGGAT)_{7}$ | CTTTTAATCCCCACAGATGG | GAAAACCTCCACCACCACC | 45 | 94 | 94,112 | PI |
| 36 | AVA03 | $(A)_{16}$ | CCCTACATCACCATACAAAGCA | TGGAAATCTCTATCATCTCCTC | 50 | 95,99 | 93,99 | FI |
| 37 | AVA04 | $(T)_{25}$ | TCTTCTCCATTGAACAATGTGG | AACCATTGATAAATTGGTTTGG | 50 | 114,120 | 111,118 | FI |
| 38 | AVA05 | $(T)_{18}$ | TGTCATCGACGCTTACTTCA | GATCTTGTGATGGGTCTTGAC | 50 | 87,90 | 80 | MI |
| 39 | AVMIX05 | $(ATG)_{4}$, $(T)_{15}$ | TCAAGTAGACGACCCCAACA | CATCATCATCACCTTCCCCC | 50 | 99,101 | 101 | MI |
| 40 | AVA06 | $(A)_{25}$ | TCATTAGCTCCCGTGCGTTC | TCTCCAACGGGTTCCACGT | 50 | 200 | 151,200 | PI |
| 41 | AVA07 | $(T)_{15}$ | CTCAGCCCCTACAGGTGCG | TTAGTCTGAACTCTGTTTTGGC | 50 | 153,155 | 153,157 | FI |
| 42 | AVA08 | $(A)_{23}$ | ATGTAAGTCAGCAAGTGGCG | AGACCCAGAATCTCCACCG | 50 | 88 | 84,88 | PI |
| 43 | AVMIX06 | $(T)_{24}$, $(CT)_{4}$ | TTTCTTTTTTGTGTTGGCCC | GGGTTTGTCCATCTTTGCC | 50 | 115 | 109,115 | PI |
| 44 | AVA09 | $(T)_{29}$ | GAGAGGTAGTTGGGTCGTAAGA | AAACATTGAGGCCTTAAATTGG | 50 | 68,85 | 70 | MI |
| 45 | AVA10 | $(T)_{18}$ | ACCGCAGATTTGGGTAACAC | ACAATCTGAACCGCTCATCA | 50 | 145,151 | 145,151 | BI |
| 46 | AVA11 | $(A)_{19}$ | CTCCCAGAAAGAAACCAGCA | AACCAAATGTAAACATGGGAC | 50 | 131,139 | 131,139 | BI |
| 47 | AVA12 | $(A)_{18}$ | CCACCATTGCCCTACATG | CGATTGATAATTGGGGCTGT | 50 | 136,165 | 136 | MI |
| 48 | AVA13 | $(A)_{15}$ | TGATGGGTGGCATCTGATTA | AGGTCAAGGAAAGAGGACAGC | 50 | 104,107 | 104,107 | BI |
| 49 | AVAG25 | $(TC)_{14}$ | ATGGTTTTTTCCTGCCCTTT | AACAAGCCCCCTAAAAGAA | 50 | 106,110 | 100 | MI |
| 50 | AVAG26 | $(TC)_{29}$ | AAGCTTAGAAGCAAACCAAAGAT | TGTTTTTCATTTTCACCCAAG | 45 | 119,123 | 119 | MI |

[a] The optimal annealing temperature (°C) in the PCR reaction

[b] Each locus was classified to one of four groups (Groover et al. 1994):

MI (Maternal Informative) – loci which are informative for the maternal gametes only (the paternal genotype is homozygous);

PI (Paternal Informative) – loci which are informative for the paternal gametes only (the maternal genotype is homozygous);

BI (Both Informative) – loci which are heterozygote in both parents with the same two alleles;

FI (Fully Informative) – both parents are heterozygous and the number of alleles is three or four

on the linkage groups did not reveal any deviation from the Poisson distribution, indicating no significant deviation from randomness. The total map length is 352.6 cM.

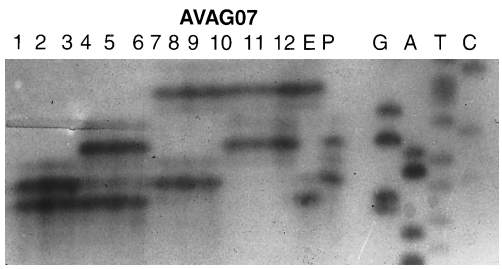


**Fig. 1** Segregation of the alleles of the AVA607 locus. *Lanes 1–12* offspring, *E* 'Ettinger', *P* 'Pinkerton', *G,A,T,C* M13 sequence as size standard

## Discussion

### SSR abundance in avocado

While in the human genome (Hudson et al. 1995), $(GT/CA)_n$ repeats are the most frequent, $(TC/AG)_n$ and $(AT/TA)_n$ repeats seem to be the most frequent in plants genomes (Cregan 1992; Langercrantz et al. 1993; Wang et al. 1994). In this study, it was found that the two most frequent repeats in avocado are $(A/T)_n$ and $(TC/AG)_n$. The number of CA repeats in the genome of forest trees was estimated to be 5,000, while the estimate for AG repeats was 300,000 (Condit and Hubble 1991). Roder et al. (1995) estimated the number of $(GA)_n$ and $(GT)_n$ repeats in the haploid wheat genome to be $3.6 \times 10^4$ and $2.3 \times 10^4$, respectively. A screening of GenBank barley sequences showed that $(A/T)_n$ is the

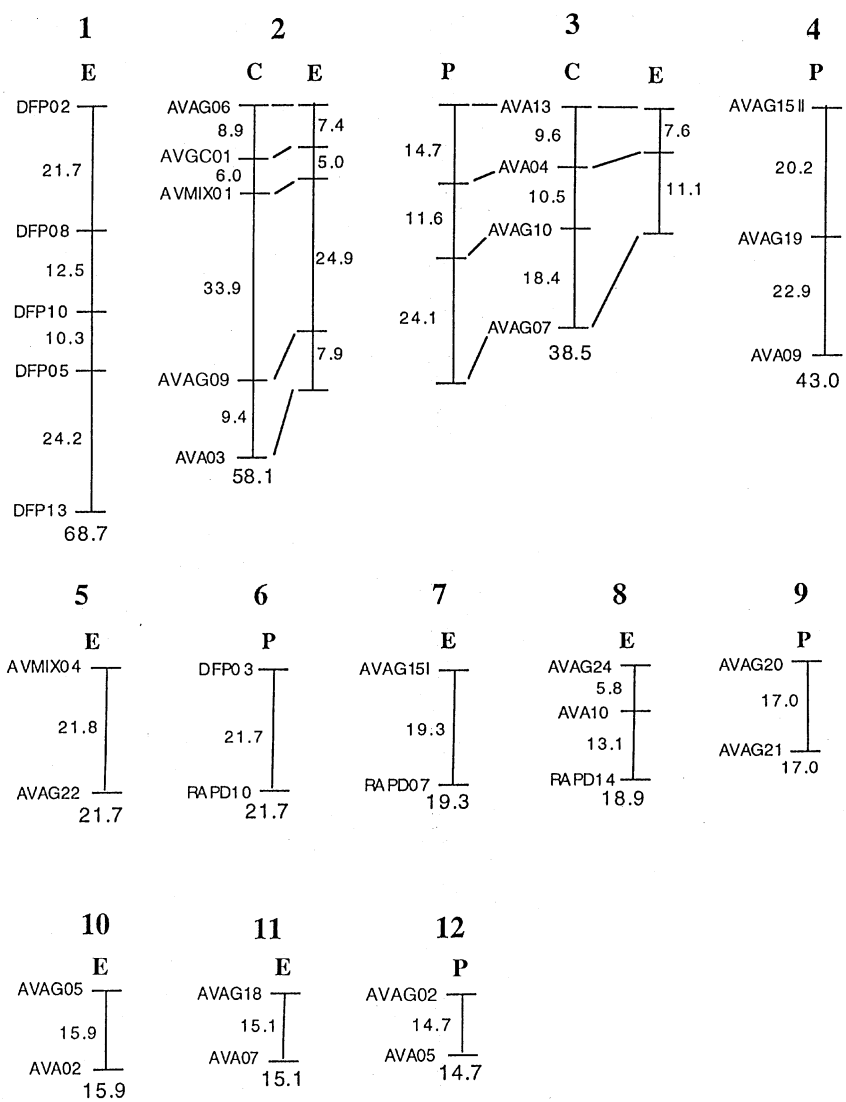


**Fig. 2** Avocado linkage groups. *E* 'Ettinger' map, *P* 'Pinkerton' map, *C* combined map. Distances between markers are in centiMorgans

most frequent microsatellite in that species (Becker and Heun 1995).

Due to the lack of prior information on polymorphism of short repeats in avocado, we decided to analyze SSR loci with smaller numbers of repeats than is commonly considered to be the minimum number that defines a microsatellite (Cregan 1992; Beckman and Weber 1992; Bell and Ecker 1994; Langercrantz et al. 1993; Morgante and Olivieri 1993; Wang et al. 1994). In retrospect, no difference in polymorphism level of $(A/T)_n$ markers was found between those having a high number of repeats (more than 20) or a smaller number of repeats (less than 20). Only one $(TC/AG)_n$ marker with $n < 10$ repeats was used and thus no conclusion can be drawn about the polymorphism level of SSRs with a small number of $(TC/AG)_n$ repeats. The high number of predicted microsatellites in the avocado genome (approx. 45,200, Table 1) indicates that an SSR-based map with high resolution can be obtained.

## From positive clones to SSR primers

Of the positive clones 76% gave reliable sequence data, and primers were designed for 28% of the positive clones (similar to Lavi et al. 1994b). In humans, Weissenbach et al. (1992) reported that from 12,014 clones that putatively contained a $(CA)_n$ microsatellite, 23% gave unreliable sequence data, 12% had short repeats, and primers were designed from 25% of the sequenced clones. This figure is close to the 28% obtained in the current study. It is quite possible to improve our results by better selection of the positive clones, by improving the sequencing procedure and by a better selection of primers (P. Cregan, unpublished results).

## Mapping the avocado genome

Of the 90 SSR, RAPD and DFP markers that were analyzed, 34 were mapped in linkage groups and 56 markers remain unlinked. The proportion of unlinked loci (62%) is higher than in other maps. In soybean, 40 SSRs were integrand into a known map using 60 $F_2$ plants (Akkaya et al. 1995). When only the 40 SSRs were analyzed, 18 of them (45%) remained unlinked. Poisson distribution analysis indicated that the SSR markers are randomly dispersed in the avocado genome. Marker distribution analyses in other species have shown similar results (Akkaya et al. 1995).

Linkage was found between a RAPD marker and a DFP band (linkage group 6) but not between SSR markers and DFP bands. The lack of linkage between SSR and DFP markers could result if DFP loci are not randomly distributed in the avocado genome but are clustered in centromeric and telomeric regions as has been reported to occur in humans (Royle et al. 1988). The differences in recombination frequencies between the paternal and maternal genomes (Fig. 2 linkage groups 2 and 3) are in agreement with De Vicente and Tanksley (1991). They detected differences in recombination frequency in tomato using the same cultivar as either the female or the male parent.

The mapping approach which we used to map the avocado genome consists of transforming the avocado mapping data to $BC_1$ data which meet the MAPMAKER requirements. The results of this analysis were verified using other computer programs. None of the computer programs we used could efficiently map the avocado data. The genome of the loblolly pine, which is an out-crossing species, was mapped using the JOINMAP computer program (Groover et al. 1994). This program is suitable for analyzing the avocado mapping data but could not efficiently map markers that deviate from Mendelian expectations. The proportion of non-Mendelian markers in avocado is estimated to be 15% (Mhameed et al. 1996).

The majority of the DNA markers used in this study to map the avocado genome are SSRs, which are now the marker of choice in human genetics linkage analysis. SSR markers provide a polymorphic, abundant, locus-specific, and reliable marker system for use in avocado. The DFP markers used here are also highly polymorphic but are not PCR-based or locus-specific. The use of such a multilocus marker for linkage analysis can be problematic. RAPD markers, the third type of DNA marker used in this study, are PCR-based, but as a dominant marker, they have a low level of informativeness. A further difficulty with RAPD markers is their poor reproducibility. Thus, we conclude that for avocado and other sub-tropical fruit trees SSR markers appear to be the most desirable marker system for the purpose of linkage map development. The linkage map reported here is currently being used for the detection of QTL controlling important fruit quality traits in avocado.

**Acknowledgements** We wish to acknowledge the expert assistance of Dr. J. Van Ooijen who did the linkage analysis with the JOINMAP program and of Dr. T. Cox-Matise who helped us with the analyses of the MULTIMAP program. We highly appreciate the technical assistance of D. Saada and A. Chapniq.

## References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci 35:1439–1445

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plan Mol Biol 27:835–845

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

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137–144

Bernatzky RB, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Botstein D, White RL, Skolnick M, Davus RW (1980) Construction of genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Cass LG, Kirven KA, Christofferson RE (1990) Isolation and characterization of a cellulase gene family member expressed during avocado fruit ripening. Mol Gen Genet 223:76–86

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

Cregan PB (1992) Simple sequence repeat DNA length polymorphisms. Probe 2:18–22

Cregan PB, Bhagwat AA, Akkaya MS, Rongwen J (1994) Microsatellite fingerprinting and mapping of soybean. Methods Mol Cell Biol 5:49–61

De Vicente MC, Tanksley SD (1991) Genome-wide reduction in recombination of backcross progeny derived from male versus female gametes in an interspecific cross of tomato. Theor Appl Genet 83:173–178

Degani C, Goldring A, Adato I, El-Batsri R, Gazit S (1990) Pollen parent effect on outcrossing rate, yield, and fruit characteristics of 'Fuerte' avocado. HortScience 25:471–473

Dietrich W, Kats H, Lincoln SE, Shin HS, Friedman J, Dracopoli NC, Lander ES (1992) A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423–447

Durham RE, Liou PC, Gmitter FG Jr, Moore GA (1992) Linkage of restriction fragment length polymorphism and isozymes in *Citrus*. Theor Appl Genet 84:39–48

Feinberg AP, Vogelstein B (1983) A technique for ratio-labeling DNA restriction endonuclease fragments to high specific activity. Ann Biochem 132:6–13

Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theor Appl Genet 91:262–269

Fritsch EP, Sambrook J, Maniatis T (1989) Molecular cloning: a laboratory manual 2nd edn. Cold Spring Harbor University Press, New York

Furnier GH, Cummings MP, Clegg MT (1990) Evolution of the avocados as revealed by DNA restriction site variation. J Hered 81:183–188

Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchel-Olds T, Sherman B, Vujcic S, Williams C, Neale D (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138:1293–1300

Haberfeld A, Cahaner A, Yoffe O, Plotzky Y, Hillel J (1991) DNA fingerprint of farm animals generated by micro and minisatellite DNA probes. Anim Genet 22:299–305

Helentjaris T, Slocum M, Wright S, Schffer A, Nienhuis J (1986) Construction of linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor Appl Genet 72:761–769

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

Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH, Hu X, Colbert AME, Rosenberg C, Reeve-Daly MP, Rozen S, Hui L, Wu X, Vestergaard C, Wilson KM, Bae JS, Marita S, Ganiatsas S, Evans CA, DeAngelis MM, Ingalls KA, Nahf RW, Horton LT, Anderson MO, Collymore AJ, Ye W, Kouyoumjian V, Zemsteva IS, Tam J, Devine R, Courtney DF, Renaud MT, Nguyen H, O'Connor TJ, Fizames C, Faure S, Gyapay G, Dib C, Morissette J, Orlin JB, Birren BW, Goodman N, Weissenbach J, Hawkins TL, Foote S, Page DC, Lander ES (1995) An STS-based map of the human genome. Science 270:1945–1954

Kijas JMH, Fowler JCS, Thomas MR (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. Genome 38:349–355

Lander ES, Green P (1987) Construction of multi-locus genetic linkage maps in humans. Proc Natl Acad Sci USA 84:2363–2367

Lander ES, Green P, Abrahamson J, Barlow A, Daley M, Lincoln S, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

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

Lathrop GM, Lalouel JM, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81:3443:3446

Lavi U, Hillel J, Vainstein A, Lahav E, Sharon D (1991) Application of DNA fingerprints for identification and analysis of avocado. J Am Soc Hortic Sci 116:1078–1081

Lavi U, Akkaya M, Bhagwat A, Lahav E, Cregan PB (1994a) Methodology of generation and characteristics of simple sequence repeat DNA markers in avocado (*Persea americana* M.). Euphytica 80:171–177

Lavi U, Cregan PB, Schaap T, Hillel J (1994b) Application of DNA markers for identification and breeding of perennial fruit crops. Plant Breed Rev 7:195–226

Matise TC, Perlin M, Chakravarti A (1994) Automated construction of genetic linkage maps using an expert system (MULTIMAP): application to a human linkage map. Nat Genet 6:384–390

Mhameed S, Hillel J, Lahav E, Sharon D, Lavi U (1995) Genetic association between DNA fingerprint fragments and loci controlling agriculturally important traits in avocado (*Persea americana* Mill.). Ephytica 81:81–87

Mhameed S, Sharon D, Hillel J, Lahav E, Kaufman D, Lavi U (1996) Level of heterozygosity and mode of inheritance of variable number of tandem repeat loci in avocado. J Am Soc Hortic Sci 121:778–782

Mhameed S, Sharon D, Kaufman D, Lahav E, Hillel J, Degani C, Lavi U (1997) Genetic relationships within avocado (*Persea americana*) cultivars and between *Persea* species. Theor Appl Genet 94:279–286

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

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

Roder MS, Plaschke J, Konig SU, Borner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90:43–48

Royle NJ, Clarkson RE, Wong Z, Jeffreys AJ (1988) Clustering of hypervariable minisatellites in the proterminal region of human autosomes. Genomics 3:352–360

Senior ML, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of targeted repeats using a CT primer. Genome 36:884–889

Senior ML, Chin ECL, Lee M, Smith JSC, Stuber CW (1996) Simple sequence repeat markers developed from maize sequences found in the Genbank database: map construction. Crop Sci 36:1676–1683

Stam P, Van Ooijen JW (1995) JOINMAP (tm) version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463–6471

Thomas MR, Scott NS (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor Appl Genet 86:985–990

Veilleux RE, Shen LY, Paz MM (1995) Analysis of the genetic composition of anther-derived potato by randomly amplified polymorphic DNA and simple sequence repeats. Genome 38:1153–1162

Vowden CJ, Ridout MS, Tobutt KR (1995) LINKEM: a program for genetic linkage analysis. J Hered 86:249–250

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88:1–6

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801