

UC Riverside

UC Riverside Previously Published Works

Title

AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*

Permalink

<https://escholarship.org/uc/item/6dh1s48s>

Journal

Tree Genetics & Genomes, 6(1)

ISSN

1614-2950

Authors

Kepiro, J. L.
Roose, M. L.

Publication Date

2009-12-01

DOI

10.1007/s11295-009-0223-z

Peer reviewed

AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*

J. L. Kepiro · M. L. Roose

Received: 27 July 2007 / Revised: 29 August 2008 / Accepted: 24 June 2009 / Published online: 28 July 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Some citrus varieties express a form of apomixis termed nucellar embryony in which the adventive embryos develop from nucellus tissue surrounding the embryo sac. This trait results in many seeds containing multiple embryos (polyembryony). Inheritance of the frequency of polyembryony was studied in 88 progeny from a cross of *Citrus maxima* (monoembryonic) × *Poncirus trifoliata* (polyembryonic). The frequency of polyembryonic seed produced by each progeny was determined by scoring 100–500 seeds for the number of seedlings to emerge from each seed. Two groups of eight individuals from each extreme of the population were chosen for bulked segregant analysis with amplified fragment length polymorphism markers amplified with 256 primer combinations. Candidate markers identified in the bulks as linked to the trait were tested on the 32 individuals used to create the bulks and then on the remaining plants in the population. Five candidate markers tightly linked to polyembryony in *P. trifoliata* were identified. Specific marker alleles were present in nearly all progeny that produced polyembryonic seed, and alternate alleles were present in nearly all progeny that produced only monoembryonic seed. The region defined by

these markers very likely contains a gene that is essential for the production of polyembryonic seeds by apomixis, but also shows segregation distortion. The proportion of polyembryonic seeds varied widely among the hybrid progeny, probably due to other genes. Scoring 119 progeny of a *P. trifoliata* selfed population for the closely linked markers and the proportion of polyembryonic seeds confirmed close linkage between these markers and polyembryony.

Keywords Nucellar embryony · Facultative apomixis · Citrus

Introduction

Many citrus cultivars produce apomictic seeds by a mechanism generally termed nucellar embryony because adventive embryogenesis occurs in cells of the somatic nucellar tissue that surrounds the embryo sac (Kobayashi et al. 1979). Such embryos are genetically identical to the maternal parent. This characteristic is used to propagate genetically uniform seedlings for use as rootstocks for citrus, but cultivars having the trait are difficult to use as female parents in citrus breeding. Nucellar embryos do not mature to a germinable stage unless the endosperm develops following pollination and fertilization (Esen and Soost 1977). In many varieties, a sexual (zygotic) embryo may also develop within the same seed. Developmental competition between the zygotic and nucellar embryos as well as the genotype of the zygote will affect maturation, and in seeds with mature nucellar embryos, immature zygotic and nucellar embryos are often present but do not germinate (Ueno et al. 1967; Esen and Soost 1977;

Communicated by F. GMITTER

Electronic supplementary material The online version of this article (doi:10.1007/s11295-009-0223-z) contains supplementary material, which is available to authorized users.

J. L. Kepiro · M. L. Roose (✉)
Department of Botany and Plant Sciences,
University of California,
Riverside, CA 92521, USA
e-mail: mikeal.roose@ucr.edu

Koltunow et al. 1995; Wakana and Uemoto 1988). Thus, a single citrus seed may contain one or more mature embryos and additional immature embryos, one (mature or immature) of which may be a sexually produced embryo. Such seeds are termed polyembryonic, and the cultivar that produces them is said to have polyembryony. Even a genotype with polyembryony may produce many seeds from which only one nucellar or zygotic seedling germinates. Citrus genotypes very occasionally produce seeds containing two zygotic embryos (fraternal or identical twins), but this is rare and appears to be independent of nucellar embryony (Frost 1926).

Citrus cultivars vary considerably in the percentage of seeds that contain multiple mature embryos and in the proportion of seedlings that are of nucellar vs sexual origin. In general, cultivars that produce a high proportion of seeds from which multiple seedlings emerge also produce a high percentage of nucellar seedlings (Soost and Roose 1996). All known *Citrus maxima* [Burm.] Merrill (pummelo) cultivars produce only sexual (zygotic) seedlings, while many other taxa including orange [*Citrus sinensis* (L.) Osbeck], grapefruit (*C. x paradisi* Macf.), lemon [*Citrus limon* (L.) Burm. f.], some mandarins (*Citrus reticulata* Blanco), and most trifoliolate orange [*Poncirus trifoliata* (L.) Raf.] cultivars produce polyembryonic seeds that contain nucellar embryos. Among cultivars having nucellar embryony, phenotypic data from crosses and molecular marker studies show that the proportion of seedlings that are of nucellar origin varies widely (Moore and Castle 1988; Ruiz et al. 2000; Xiang and Roose 1988). This genetic variability has allowed several studies of the inheritance of nucellar embryony in *Citrus* and its sexually compatible relative, *P. trifoliata*. Parlevliet and Cameron (1959) suggested that nucellar embryony is controlled by a single major dominant gene that is heterozygous in trifoliolate orange and absent in “Chandler” pummelo. They also suggest that minor genes may control the level of expression. Later, Cameron and Soost (1979) found that segregation ratios in some populations were not consistent with the single dominant gene model. Other work suggests that several genes control nucellar embryony and that polyembryony is an independent trait (Garcia et al. 1999; Asins et al. 2002).

Bulked segregant analysis is a powerful approach for detecting markers linked to genes with a large effects on a trait, but variants such as “selective genotyping” (Lander and Botstein 1989) can be used to identify markers for more quantitative traits. The purpose of this study was to examine the frequency of nucellar embryony in progeny of a cross of “Chandler” pummelo (*C. maxima*) with *P. trifoliata* and to use bulked segregant analysis (Michelmore et al. 1991) to identify amplified fragment length polymorphism (AFLP) markers (Vos et al. 1995) linked to genes for this trait.

Materials and methods

Plant materials and description of populations

Two populations were evaluated for nucellar embryony by seed germination and scoring for single or multiple embryos. Data for the number of monoembryonic and polyembryonic seeds per seed lot was collected from individuals in the population for a minimum of 2 years from 1999 to 2002. Some individuals that produced few seeds or seeds with a very low number of polyembryonic seeds warranted collection of data in additional years.

Population 1 This population results from two different crosses made in 1986. The maternal parent was the monoembryonic *C. maxima*, “Chandler” pummelo. The pollen sources were two closely related polyembryonic *P. trifoliata* cultivars, “Rubidoux” and “Webber Fawcett.” Pollen for this population was initially collected from “Rubidoux,” but the peak of its bloom is at least 4 weeks earlier than that of “Chandler.” “Webber Fawcett’s” bloom period coincided better with “Chandler’s,” and it was substituted as a pollen donor to produce additional hybrids. Unfortunately, there are no records of which fruit resulted from pollination by “Rubidoux” or “Webber Fawcett.” Therefore, the fruit and seed from the two crosses were harvested in bulk, resulting in a population with a maternal parent of “Chandler” and the majority of the progeny with the male parent “Rubidoux” and others with the male parent “Webber Fawcett.” Molecular markers can distinguish the pollen donor of the progeny, although for the purposes of this study, “Webber Fawcett” and “Rubidoux” can be considered nearly identical genetically (Fang and Roose 1997; Fang et al. 1997) and both have the traits of polyembryony and nucellar embryony. For 88 progeny in this population, at least five fruits per tree were collected for at least two seasons from 1999 to 2002, and scoring data for monoembryony and polyembryony was collected.

Population 2 This population was established from seed collected from open-pollinated *P. trifoliata* “Flying Dragon”, “Pomeroy,” and “Rubidoux” trees. The population was previously characterized by Khan and Roose (1988). Nucellar progeny genetically identical to the maternal parent were culled from this population by molecular marker analysis. The marker data indicated that nearly all sexual seedlings from open pollination of *P. trifoliata* originated by selfing, but a few were identified as hybrids between *P. trifoliata* cultivars. All open-pollinated progeny of *P. trifoliata* were expected to result from selfing or intermating between *P. trifoliata* cultivars because they bloom several weeks before *Citrus*. One

hundred nineteen progeny were in this population: 71 from “Flying Dragon”, 41 from “Pomeroy,” and 7 from “Rubidoux”. Germination data were collected for two or more years from 1998 to 2002.

Seed collection and germinations

Fruits were collected at maturity in the fall and seed were extracted, cleaned, treated (Klotz 1978), and stored. Before germination, seeds were surface-sterilized in 10% bleach (v/v) for 10 min, then rinsed twice in autoclaved sterile water for 5 min per rinse. Thirty to 60 seeds were evenly distributed into sterile Petri dishes lined with wetted filter paper and germinated in a growth chamber at 31.1°C, and the paper was rewetted as necessary to keep seeds moist. After 10 to 30 days, seeds that had germinated and opened the seed coat far enough that the contents of each seed could be seen were removed for scoring. The number of monoembryonic and/or polyembryonic seeds were counted and recorded.

DNA purification

Young leaves were collected in spring and leaf DNA was extracted using the cetyltrimethyl ammonium bromide protocols of Webb and Knapp (1990) as modified by Fang et al. (1997). DNA concentration was determined using a fluorometric method (Labarca and Paigen 1980) with a TKO 100 fluorometer (Hoeffer Scientific Instruments, San Francisco, CA, USA). DNA from rare multiple seedlings produced by individuals with low and very low polyembryony was extracted using DNeasy® plant mini kit (Qiagen Inc., Valencia, CA, USA) from either fresh leaf material or N₂ quick-frozen shoots, roots, or cotyledons of seedlings. DNA concentration was standardized at 32 ng/μl.

AFLP technique for restriction/ligation and PCR

An AFLP system II kit for small genomes (Invitrogen LifeTechnologies Corporation, Carlsbad, CA, USA) and AFLP template preparation kits (Li-Cor Inc., Lincoln, NE, USA) were used for AFLP reactions. The protocols of the Invitrogen AFLP system II manual were followed as modified by Myburg and Remington (2000). The reaction volumes were reduced for the restriction/ligation (R/L), pre-amplifications, and selective amplifications. All R/Ls and polymerase chain reactions (PCRs) were in 200-μl PCR eight-tube strips (Stratagene, La Jolla, CA, USA) using a PTC-100 thermocycler (MJ Research, Waltham, MA, USA).

The R/L step of the protocol was performed using 125 ng of genomic DNA to yield 12.5 μl of ligated

product. The R/L product was diluted 1:10 and used as template in the pre-amplification. Invitrogen pre-amplification primer mix II for the pre-amplification reactions contained primers E+0, M+C, and dNTPs. “E = *EcoRI*” and “M = *MseI*” denote the core primer sequence, +0 (plus zero) indicates no selective bases, whereas +C indicates the addition of one selective cytosine nucleotide on the core primer sequence. The pre-amplification PCR mix contained pre-amp primer mix (10.15 μl), 10X PCR buffer with 15 mM MgCl₂ (1.25 μl), and *Taq* DNA polymerase (Invitrogen, 0.5 U/reaction). An 11.50 μl aliquot of this master mix was combined with 1.0 μl template (R/L product diluted 1:10) for a total reaction volume of 12.5 μl. PCR cycling consists of 30 cycles of 94°C 30 s, 56°C 30 s + 1 s/cycle, 72°C 60 s + 1 s/cycle, 72°C 3 min, 4°C hold. The pre-amplification product was verified on 1% agarose gel. Based on the product yield, the pre-amplified template was diluted between 20- and 50-fold with TE buffer (pH8.0) to equalize the concentration of the template before selective amplification.

For selective amplification, the PCR conditions were essentially as described by Invitrogen protocols except as noted below. We used all 16 possible E+2 primers, each containing an infrared fluorescence dye (IRD) tag and mixed to the concentration instructed by the manufacturer, 1.0 μM (Li-Cor Biosciences). The *MseI*+C+2 (hereafter MC+2) primers included the eight primers mixed with dNTPs provided in the kit and eight additional MC+2 selective primers (Invitrogen) prepared at 6.7 ng/μl and containing 0.9 mM dNTPs. The 16 E+2 IRD-labeled and 16 MC+2 unlabeled primers made 256 primer combinations available for screening. Selective amplification was performed using multiplex PCR (Myburg and Remington 2000) with two different E+2 primer sequences (one IRD-700-labeled and the other IRD-800-labeled) which were combined with one MC+2 unlabeled primer in a single tube PCR mix.

Some primer combinations required the volume of each *EcoRI* primer used in the master mix to be adjusted by (±)0.1-μl increments to achieve near-equal intensity of bands on the IRD-700 and IRD-800 images. Selective amplifications were initially run with a final (1.5 mM) MgCl₂, but many primer combinations produced sharper, more intense, and consistent bands with a final (3.0 mM) MgCl₂, and the master mix (MgCl₂) was modified as necessary. A selective amplification reaction mix contained: 0.4 pM IRD-700 *EcoRI* primer (0.40 μl at 1.0 μM), 0.4 pM IRD-800 *EcoRI* primer, Invitrogen *MseI* primer mix containing dNTPs (2.33 μl), distilled water (5.13 μl), 10X PCR buffer with 15 mM MgCl₂ (1.23 μl), and *Taq* DNA polymerase (Invitrogen, 0.7 U/reaction). The volume of master mix per reaction was 10 μl. Addition of 2.3 μl of pre-amplified template (diluted

1:50) completed the reaction mix. The AFLP selective PCR program used was 94°C 15 s, 65°C 30 s –0.7°C/cycle, 72°C for 1 min, 13 cycles; 94°C 15 s, 56°C 30 s + 1 s/cycle, 72°C 1 min + 1 s/cycle, 30 cycles; 72°C for 3 min, 4°C hold. The completed reactions were denatured by addition of 5.0 µl Li-Cor stop buffer to each tube. Tubes were then heated to 94°C for 3 min and transferred to ice for 10 min before loading.

Bulked segregant analysis

To create the bulks, equal volumes (10 µl) of each of the standardized selective (1:50 diluted) templates were combined for each class, nucellar/polyembryonic, and sexual/monoembryonic (templates were tested individually for performance in selective amplification before combining). The amount of pre-amplified DNA added from each individual was estimated to be ~4 ng. Each bulk contained eight individuals of a specific class. The bulks were screened with 256 different primer combinations in a high-throughput strategy using multiplexed PCR and IRD-labeled primers.

Gels and images

The selective amplification PCR products were separated by electrophoresis using a Li-Cor 4200 Global IR2 system and 6.5–8% Long Ranger (Cambrex Bio Science, Rockland, ME, USA) or 6.5% Kplus (Li-Cor) polyacrylamide gel matrix. Electrophoresis parameters were 1,500 V, 40 amp, 40 W, 45°C, scan speed of 3, and collection time of 3 h and 30 min. A 50- to 700-bp size standard labeled with IRD-700 and/or IRD-800 (Li-Cor) was loaded in the center and outside lanes. When visualizing multiplexed PCR products, the IRD-700/800-labeled markers were mixed equally prior to loading. Image files were cropped, optimized for intensity and contrast, then printed and scored manually. The size (in base pairs) of each scored marker was estimated using the Li-Cor 50- to 700-bp ladder. AFLP markers associated with nucellar embryony were designated EMB-1 to EMB-7 with an additional letter (M or P) indicating whether the marker band was associated with monoembryony or polyembryony, respectively. Using the standard nomenclature for AFLP markers described by Cho et al. (1998), the marker name was constructed from the primer combination (described by KeyGene NV, Wageningen, The Netherlands), the relative molecular weight, and the linkage phase (M or P) from which the marker was amplified; EMB-2 to EMB-6 correspond to E22M47.095-P, (E11M49.290-P and E11M49.292-M), (E11M60.315-P and E11M60.317-M), E11M48.082-M, and E16M54.070-P (apparently co-dominant alleles are listed in brackets).

Results

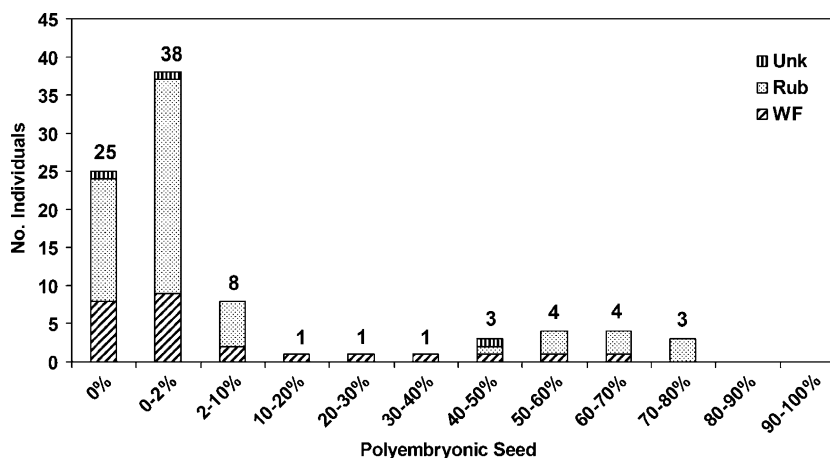
Trait scoring

The hybrid population composed of 88 progeny from monoembryonic *C. maxima* “Chandler” pummelo × polyembryonic *P. trifoliata* (“Webber Fawcett” or “Rubidoux”) was chosen for detailed analysis. For those progeny having more than 20% polyembryonic seeds, the number of seeds scored ranged from 72 to 191, with a mean of 131. More seeds were sampled from progeny having less than 20% polyembryony, the mean number being 316 and the range 143 to 547 (see Electronic supplementary material, Table S1 for details). The proportion of polyembryonic seed was fairly consistent over years, so data were summed over all years of seed collection. Seedlings classified as originating from pollination by “Webber Fawcett” and “Rubidoux” using AFLP markers (Kepiro 2003) had similar frequency distributions for percentage of polyembryonic seeds (Fig. 1). In combined data, the frequency distribution of percentage of polyembryonic seeds was bimodal and skewed to the right, with 25 progeny having no polyembryonic seeds, 46 having 0–10%, one having 10–20%, and 16 having more than 20% (Fig. 1).

For bulked segregant analysis, the polyembryonic bulks were composed of 15 hybrids having at least 10% and one hybrid having 3.7% polyembryonic seeds after 1 year of scoring. The monoembryonic bulk included only plants with 0.5% or less polyembryonic seeds. These were classified as likely to have only sexual reproduction (including rare zygotic twins). Although we eventually identified 17 progeny that produced more than 10% polyembryonic seeds, only 15 of these produced fruit during the first year of scoring when the bulks were formed, and therefore, we included one plant with a low level of polyembryony in one of the polyembryonic bulks. Similarly, although we eventually identified 25 progeny having no polyembryonic seeds, one progeny tree that had no polyembryonic seeds among those germinated during the first year and was included in the monoembryonic bulk produced two polyembryonic seeds in a later year. Classification of individuals as having nucellar embryony was generally based on the criteria used by Iwamasa et al. (1967) and Cameron and Soost (1979) in which trees expressing nucellar embryony have a high frequency of polyembryonic seed.

The population of open-pollinated trifoliolate orange seedlings was also scored for the percentage of polyembryonic seeds. A mean of 589 and a minimum of 227 seeds were scored for each progeny with less than 10% polyembryony, and at least 66 seeds were scored for each progeny with between 10% and 50% polyembryony. Individuals with high and low levels of polyembryony

Fig. 1 Frequency distribution of percent polyembryonic seeds in 88 progeny from “Chandler” pummelo × trifoliolate orange. An average of 283 seeds per individual were scored as monoembryonic or polyembryonic. Bars show cumulative totals for seedlings resulting from pollination by “Rubidoux” (stippled), “Webber Fawcett” (diagonal stripes), and unknown (vertical stripes) trifoliolate orange parents



were observed among progeny of all three cultivars (Fig. 2), and the proportion of progeny with less than 1% polyembryonic seed was fairly similar between the cultivars with larger sample sizes (0.21 for “Pomeroy” and 0.34 for “Flying Dragon”). In the data for all progeny, the frequency distribution was bimodal, with eight progeny having 0% polyembryony, 41 having 0–10%, 15 having 10–20%, and 55 having more than 20% polyembryonic seeds. There was a transgressive segregation, with 45 progeny having more than 24.4% polyembryonic seeds, while the three seed parent cultivars (“Flying Dragon”, Rubidoux, and “Webber Fawcett”) had 24.4%, 17.0%, and 19.3%, respectively).

Linked markers identified using BSA

The bulk templates S1, S2, (S = sexually reproducing, each bulk containing eight monoembryonic individuals) and N1, N2 (N = nucellar reproductivity, each bulk containing eight

polyembryonic individuals) were screened with 256 AFLP primer combinations. Some primer combinations amplified more bands or produced clearer images than others. PCR conditions were not optimized for specific primer combinations in the screening because the PCR protocol previously listed generated fair to excellent images for most primer combinations. The images of the bulk products were visually screened for detectable polymorphisms between the two sexually reproducing bulks S1 and S2 and the two bulks with characteristics of nucellar embryony, N1 and N2.

Fifty-six primer combinations produced markers with some observable difference in intensity level or polymorphism between S and N bulks. In those cases, the original 1:50 diluted pre-amplified template of each individual in the bulks (individuals M1–M16 being monoembryonic and reproducing sexually and P1–P16 being polyembryonic and capable of reproducing by nucellar embryony) were

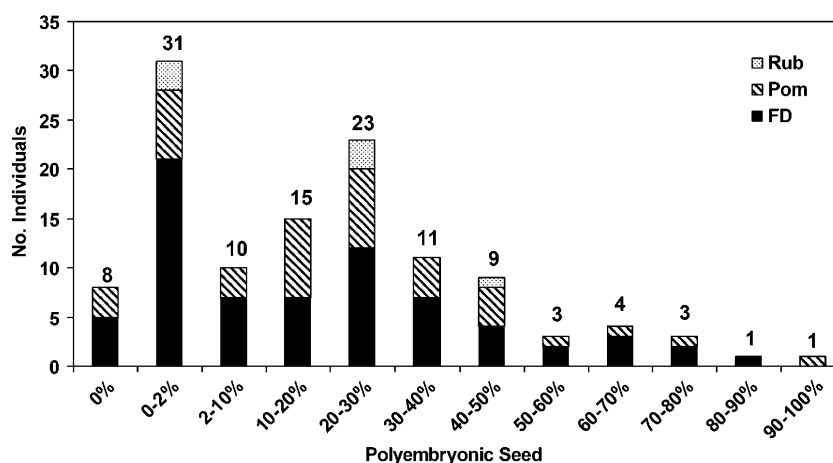


Fig. 2 Frequency distribution of percent polyembryonic seeds in 119 progeny resulting from open pollination (mainly selfing) of three selections of trifoliolate orange, “Pomeroy”, “Flying Dragon,” and “Rubidoux.” At least 227 seeds were scored for all progeny with less than 10% polyembryony, and at least 66 seeds were

scored for each progeny with between 10% and 50% polyembryony. Bars show cumulative totals for seedlings from “Rubidoux” (stippled), “Pomeroy” (diagonal stripes), and “Flying Dragon” (solid) trifoliolate orange parents

amplified separately and visualized on the Li-Cor with parental standards.

Bulked segregant analysis (BSA) identified seven candidate markers probably corresponding to five loci tightly linked to polyembryony and segregating in the bulks and individuals comprising the bulks in agreement with the classification scheme (Fig. 3). Markers EMB-2P, EMB-3P, EMB-4P, and EMB-6P were detected in the polyembryonic individuals chosen to most likely have nucellar embryony. EMB-3M, EMB-4M, and EMB-5M were detected in the monoembryonic individuals most likely to be strictly sexual. Markers EMB-3M/P and EMB-4M/P are co-dominant markers (two fragments of similar but not identical size amplified with the same primer set, both of which were present in the parent genotype and either, but not both, of which was present in each progeny individual). Markers EMB-2P, EMB-4P, EMB-5M, and EMB-6P segregated in complete agreement with the classification of the trait based on percentage of polyembryonic seed (0:16 or 16:0), while for EMB-3 M/P, two progeny had apparent crossovers between the marker and trait gene. All of these detectable polymorphisms are inherited from the *P. trifoliata* parent. The seven markers found to be most tightly linked to the trait in the analysis of individuals M1–M16 and P1–P16 were then tested on the remainder of the population. In the test population of 88 progeny, there was no recombination between markers EMB-2P, EMB-4P, EMB-4M, EMB-5M, and EMB-6P,

and recombination events separated EMB3-M/P from the other markers.

Marker EMB-6P was chosen to represent the inheritance of linked markers EMB-2P, EMB-4P, EMB-5M, and EMB-6P. When EMB-6P was present, markers EMB-2P and EMB-4P were also present and EMB-5M was absent. If EMB-6P, EMB-4P, and EMB-2P were absent, then EMB-5M was present. All individuals in the population with 1.9% or more polyembryonic seed had markers EMB-2P, EMB-4P, and EMB-6P. Five of the nine individuals producing between 1.0% and 1.4% polyembryonic seed had these markers. Only one hybrid with less than 1% polyembryonic seed had markers EMB-2P, EMB-4P, and EMB-6P, which are linked to polyembryony. The cluster of linked markers had significant segregation distortion: 32 of the 88 hybrids inherited the most tightly linked markers, and chi-square analysis was significant at $p < 0.025$. A linkage analysis using JoinMap 3.0 (Van Ooijen and Voorrips 2001) showed that these seven markers occurred at three positions in a 6.4-cM interval, EMB-3M/P being 6.3 cM from a single locus that included EMB-2P, EMB-4P, EMB-5M, and EMB-6P. Marker EMB-4M mapped 0.1 cM distal to this cluster because three individuals were scored as missing for EMB-4M, whereas none were missing for EMB-4P.

The AFLP primer combinations producing the most tightly linked markers in population one were used to test 119 individuals in the *P. trifoliata* population for markers

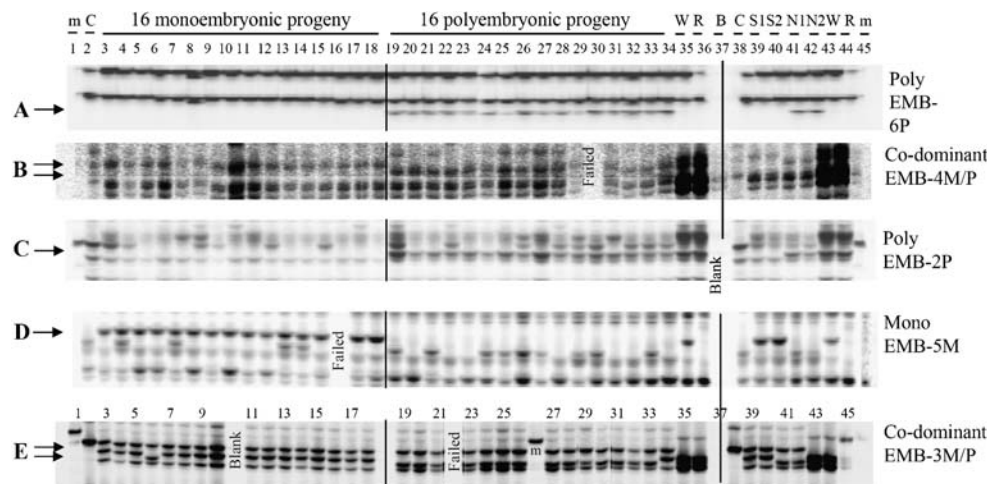


Fig. 3 Seven AFLP markers from five loci tightly linked to nucellar embryony. Lanes 1 and 45 are size markers (m) which are mostly not visible in these cropped images. Lanes 2 and 38, “Chandler” (C), lanes 35 and 43 “Webber Fawcett” (W), lanes 36 and 44 “Rubidoux” (R). Lanes 3–10 and 11–18 are monoembryonic individuals that comprised the strictly sexual bulks, S1 and S2, respectively. Lanes 19–27 and 28–34 are the individuals that comprised the polyembryonic bulks, N1 and N2, respectively. Lane 37 is blank, as is lane 11 in E. Lanes 38–44 are the original bulks flanked by the parents. Arrows indicate segregating markers. Several AFLP reactions failed; these lanes and blank lanes are indicated. Marker EMB-3 M/P (section E)

contains additional blank, and marker lanes and samples have been renumbered to match the other images. The amplification of markers EMB-6P and EMB-5 M was weak in trifoliolate parental templates made from DNA isolated from leaves collected in the fall; however, with enhancements (not shown), the origin of these markers was determined to be *P. trifoliata*. Markers EMB-2 to EMB-6 correspond to E22M47.095-P, (E11M49.290-P and E11M49.292-M), (E11M60.315-P and E11M60.317-M), E11M48.082-M, and E16M54.070-P where marker size in base pairs is given after the “dot” at the end of the marker name, and apparently, co-dominant alleles are grouped using parentheses

EMB-2P, EMB-3P, EMB-3M, EMB-4P, EMB-4M, EMB-5M, and EMB-6P which were expected to segregate 3:1 for dominant markers or 1:2:1 for co-dominant markers. Based on the close linkage of EMB-6P with polyembryony in population one, this marker was also tracked in the individuals of the *P. trifoliata* population for linkage to polyembryony. All individuals in the *P. trifoliata* population that produced 1.94% or more polyembryonic seed had marker EMB-6P. Marker EMB-6P was faint in one of the 84 plants that produced more than 1.94% polyembryonic seed. In this plant, EMB-4P was scored as present, confirming the presence of this chromosome segment. Marker EMB-6P was present in only four of the 38 plants scored for markers and having less than 1.94% polyembryonic seed, and EMB-5 M was present in all plants in which EMB-6P was absent. The other markers showed the expected co-segregation with EMB-6P and EMB-5M.

AFLP analyses of seedling from hybrids with low frequencies of polyembryonic seed

During the germination and scoring work, seedlings from seed with multiple seedlings were collected if the tree producing the seed had a low frequency of polyembryony. Polyembryonic seeds of two such trees were analyzed with four AFLP primer combinations (E19M59, E18M59, E21M62, and E12M62), each detecting 18 or more loci that segregated in population 1. We expect about 50% of these loci (at least 36 total) to be heterozygous in each parent tree, a number likely adequate to distinguish nearly all nucellar and zygotic seedlings. Siblings from a polyembryonic seed were loaded in adjacent lanes, and their banding patterns were compared to each other and to that of the tree bearing the fruit (maternal parent). Marker phenotypes of most of the fingerprinted seedlings from the low polyembryonic hybrids were consistent with an origin by selfing. Each heterozygous marker should segregate 3:1, so 25% of such markers are expected to be absent in each seedling. Some of these seedlings produced fingerprints nearly identical to the maternal parent, but one or more markers were absent in most seedlings when compared to the parent, consistent with segregation. Seedlings from four polyembryonic seeds of hybrid plant 35 (1.2% polyembryonic seed and EMB-6P+) were analyzed with 27 high-quality markers heterozygous in plant 35. Three of the four polyembryonic seeds from hybrid plant 35 produced a total of 5 nucellar seedlings and allowed plant 35 to be classified as having the trait of nucellar embryony. Among the six zygotic seedlings from these four seeds, both fraternal and identical twins were detected, and from one seed, we obtained a pair of identical twins and an additional fraternal twin. Analysis of ten seedlings from four polyembryonic seeds from hybrid plant 42 (1.9%

polyembryonic seed and EMB-6P+) with 32 markers heterozygous in the maternal parent showed that all were of zygotic origin, with some seeds containing fraternal twins, some identical twins, and one containing two pairs of identical twins. Thus, the presence of nucellar embryony in this genotype was not confirmed by analysis of these few seedlings. We also analyzed 12 seedlings from five polyembryonic seeds of three highly polyembryonic genotypes. All seedlings had banding patterns identical to that of their maternal parent and were of nucellar origin.

Discussion

The trait of polyembryony above 1.4% in the Chandler pummelo × trifoliolate orange population apparently requires a dominant allele inherited from *P. trifoliata* that was in coupling with markers EMB-2P, EMB-4P, and EMB-6P. A strong association between marker EMB-6P and production of polyembryonic seed was also observed in the open-pollinated *P. trifoliata* population. Allelic variation in minor genes could explain the variability in the level of polyembryony among polyembryonic genotypes.

The three markers segregating with the extreme phenotypes of population 1 were tested on the individuals in the population not initially classified. These data showed that all hybrids of this population with more than 1.4% polyembryony have the markers EMB-2P, EMB-4P, and EMB-6P, which are linked to high (>20%) polyembryony. With only one exception, marker EMB-5M was present and markers EMB-2P, EMB-4P, and EMB-6P were absent in all individuals with less than 1% polyembryony.

Nine hybrids in population 1 produced polyembryonic seed at frequencies of 1.0–1.4%. Five of these have EMB-6P, and AFLP analysis of open-pollinated seedlings showed that three of these had nucellar reproduction. Availability of material prevented extensive testing of a large number of seedlings from seeds with multiple embryos when the frequency of polyembryony was low. Therefore, three hybrid plants having marker EMB-6P and producing 0.2%, 1.1%, and 1.4% polyembryonic seed cannot be classified as strictly sexual or capable of nucellar embryony. Hybrids with less than 1.4% polyembryonic seed and lacking EMB-6P may have the trait of nucellar embryony, but the marker may have been uncoupled by a crossover event. It is possible that the lower range of polyembryony associated with nucellar reproduction and the upper limit of polyembryony by zygotic twinning overlap in the range of 1.0–1.4% in this population. Frost (1926) reported ten cases among approximately 1,000 *Citrus* hybrid seedlings where apparent twinning resulted in two hybrid seedlings from one seed. His recorded observations of twinning were in the 1% range for monoembryonic *Citrus*.

Use of the percentage of polyembryonic seed scoring data to classify individuals as capable of nucellar embryogenesis was based on established methods from previous studies (Frost 1926; Iwamasa et al. 1967; Cameron and Soost 1979). The morphology of seedlings and progeny trees had been used by Frost (1926), as well as Cameron and Soost (1979), to classify progeny as zygotic or nucellar, but they admitted that the differences used to distinguish zygotic progeny from nucellar progeny were slight and sometimes difficult to distinguish. This may have resulted in reports of trees classified as strictly sexual monoembryonic parents producing up to 7% seed containing multiple zygotic embryos. Difficulty in classification also required the creation of such categories as “near-monoembryonic” to describe individuals producing between 7% and 10% polyembryonic seed. In the current study, germinating and scoring large numbers of seeds allowed capture of rare polyembryonic seeds. By scoring seedlings with molecular markers, some of the hybrids with low percentages of polyembryonic seed could be classified without ambiguity as having the ability to produce seedlings genetically identical to the maternal parent. AFLP fingerprinting using up to 27 marker loci confirmed that some seedlings, from trees producing as few as one in a hundred (1.0%) polyembryonic seeds, were genetically identical to the maternal parent.

Seedlings genetically identical to the mother tree were generated at a high frequency in polyembryonic seed produced by hybrids with high polyembryony. In the hybrid population, those individuals analyzed with AFLP and having >20% polyembryony expressed nucellar embryony in every seedling analyzed from polyembryonic seed, albeit a limited sample. However, hybrids with the linked marker EMB-6P and low polyembryony produced a smaller proportion of seedlings identical to the maternal parent when compared to the proportion of nucellar seeds produced by trees with high polyembryony. We did not use markers to confirm that apparently monoembryonic progeny trees produced only zygotic seedlings. Thus, an assumption of our analysis is that genotypes producing only monoembryonic seeds can produce only zygotic seedlings. Since we have shown that polyembryonic seeds of some genotypes that are nearly monoembryonic contain only zygotic seedlings, it seems reasonable that monoembryony is also generally associated with inability to produce nucellar seedlings. Our studies also emphasize the importance of studying large samples of seeds before inferring that a genotype is monoembryonic because some genotypes produce only about 2% polyembryonic seeds, yet have nucellar embryony. The abundance of nucellar progeny can be a major impediment when zygotic progeny are needed for new hybrid selection. As in previous studies (summarized in Soost and Roose 1996), the proportion of nucellar

seedlings declined with the percentage of polyembryonic seed produced. Understanding genetic control of the relationship between polyembryony and nucellar embryony will be important to breeders.

Nucellar embryos arise from somatic nucellus tissue and so are genetically identical to the maternal parent (Wilms et al. 1983). In hybrids with greater than 20% polyembryony, the nucellus cells are initialized to enter the embryony pathway resulting in nucellar progeny identical to the maternal parent in all seeds analyzed. In trees that produced only low percentages of polyembryonic seed, the genotypes of most seedlings were consistent with an origin by selfing and were not identical to that of the maternal parent.

The association between marker alleles and the proportion of polyembryonic seed is weakest for those hybrids with low but non-zero percentages (1–2%) of polyembryonic seed. This relationship would not be predicted if crossovers between the marker and the trait locus were the only factor reducing the perfect association between the marker and the trait. Production of rare zygotic twins is one factor that may lead to misclassification of progeny having very low percentages of polyembryonic seed. In addition, it seems possible that the genomic region tagged by this marker is not sufficient to produce highly polyembryonic seeds and nucellar progeny. This possibility is addressed by a quantitative trait loci (QTL) analysis (Kepiro and Roose, submitted).

Parlevliet and Cameron (1959) proposed a genetic model for nucellar embryony/polyembryony being controlled by a major dominant gene *P* and minor or modifier genes influencing the level of expression. Segregation ratios fit this model in some crosses, but not in all. However, this model is difficult to test because the minimum percentage of polyembryonic seed to classify progeny as having nucellar embryony was set at a somewhat arbitrary level. Iwamasa et al. (1967) excluded hybrids with between 6% and 20% polyembryony from analysis. In doing so, they were able to fit most segregation ratios in their crosses to a single gene model adding support to the postulation of Parlevliet and Cameron (1959). In crosses that varied widely from the expected segregation ratios, Iwamasa et al. (1967) suggested that duplicate genes or modifying genes complicated the ratios.

Cameron and Soost (1979) revised the single dominant gene model for crosses of *Citrus* × *Poncirus*. They proposed a model with two dominant genes, each being heterozygous in *Poncirus* *P1p1 P2p2*. Wide variation in expected segregation ratios could be explained if the sexually reproducing *Citrus* parent carried zero or one, but not both of the dominant alleles. Hong et al. (2001) proposed a model for apomixis (nucellar embryony) in *Citrus* × *Poncirus* involving two complementary genes, *A1* and *A2*. They suggest that genotypes *A1_A2_* produce

seeds through apomixis. As part of their model *Ala1* homozygous plants are lethal, which requires that apomictic plants have genotype *Ala1 Ala2_*. In their model, *Poncirus* is *Ala1Ala2A2*.

We have identified markers tightly linked to a dominant allele in the *Poncirus* parent that has segregation distortion. In population 11, marker EMB-6P segregates 32:56, which is significantly different ($p < 0.025$) from an expected segregation ratio of 1:1. In the open-pollinated *P. trifoliata* population, the frequency distribution of polyembryony was fairly similar among progeny of the two parents with larger sample sizes, suggesting similar genetic control and justifying analysis combined data. In this population, marker EMB-6P segregated 85:34. This ratio is also biased toward the monoembryony allele if we assume that all progeny originated by selfing so that the expected ratio would be 3:1 (or ~90:30). The observed ratio (~2.5:1) was not significantly different from 3:1 or from the 2:1 model proposed by Hong et al. (2001). However, in order for the lethal homozygous genotype *Ala1* to explain the segregation distortion of EMB-6P in population 1, the Chandler pummelo would be genotype *Ala1a2a2*, whereas our assumption under a two-gene model would be that Chandler is *al1a1a2a2*. The marker EMB-6P could not be amplified from the Chandler parent. Furthermore, if Chandler is *Ala1a2a2* and capable of introducing the lethal condition *Ala1*, we would expect that two of three of the viable progeny carry the *Al* allele and produce polyembryonic seed by nucellar embryony. The results presented here contradict that scenario. Slightly over one third of the hybrids in the Chandler \times trifoliata population carried what would be the *Al* allele proposed by Hong et al. (2001). Alternatively, if marker EMB-6P was linked to the second locus proposed by Hong et al. (2001) and *Poncirus*'s genotype was *Ala1Ala2A2*, segregation in the hybrid population would not have been observed. Therefore, it appears that segregation distortion of EMB-6P and its linkage to nucellar embryony does not fit the Hong et al. (2001) model for apomixis inheritance from *Poncirus*.

It is possible to explain the skewed segregation of EMB-6P by various two-locus models such as gametophytic lethals epistatic to the marked nucellar embryony allele or a linked lethal, but our data do not allow adequate tests of these models

The model for apomixis inheritance proposed by Garcia et al. (1999) and revised by Asins et al. (2002) is based on QTL analysis. Their cross was *Citrus volkameriana* ("Volkamer" lemon) \times *P. trifoliata* "Rubidoux." Both of these parents have nucellar embryony, which increases the complexity of the analysis, and only 38 progeny were phenotyped for polyembryony and nucellar embryony, a small population size for QTL analysis. Garcia et al. (1999) mapped two QTLs associated with apomixis in *Poncirus*,

Apo4 (TAA27) on *Poncirus* linkage group P5, and *Apo5* (OPD07060) on *Poncirus* linkage group P4. They also identified the RAPD marker (OPB05040) as associated with polyembryony in *Poncirus*, but the marker could not be mapped. After revising the original analysis using additional markers and recoding the quantitative trait values, Asins et al. (2002) reported that *Apo5* in *Poncirus* disappeared. The *Apo4* QTL (TAA27) is no longer listed as associated with apomixis in *Poncirus*, although *Apo3* (Egp47) on *Poncirus* linkage group P3 was added, but is only significant at $p = 0.1$ or 10%. Previously, Garcia et al. (1999) had reported *Apo1*, *Apo2*, *Apo3*, and *Apo6* in *C. volkameriana*. *Apo2*, a *C. volkameriana* QTL, has the strongest but a negative gene effect and is related to monoembryony/polyembryony. The revised analysis by Asins et al. (2002) reports two additional QTLs in *C. volkameriana*, naming them *Apo5* and *Apo7*.

The segregation ratios and marker linkages in our population from monoembryonic *C. maxima* "Chandler" pummelo \times polyembryonic *P. trifoliata* "Webber Fawcett" or "Rubidoux" were expected to differ from *C. volkameriana* ("Volkamer" lemon) \times *P. trifoliata* "Rubidoux" based on the parental genotypes and their embryony types. *C. maxima*, "Chandler" was chosen because it lacked nucellar embryony, as do its parents (Soost and Roose 1996). Therefore, Chandler was not expected to contribute genes sufficient for nucellar embryony, and no markers were found in Chandler linked to polyembryony or nucellar embryony (Kepiro and Roose, submitted). The markers tightly linked to nucellar embryony/polyembryony that were identified here were inherited from *Poncirus*. Because different types of populations and no common markers were used, our results are not easily compared with those of Garcia et al. (1999) and Asins et al. (2002). The assessment of Asins et al. (2002) is that "apomixis has to be considered as a continuous variable although not normally distributed, which imply that new statistical methodologies for QTL detection have to be developed at allow more powerful estimation of gene effects and individual contribution of QTLs detected." While this may be true in complex crosses between divergent, polyembryonic parents, the evidence presented here indicates that a specific chromosome segment of *Poncirus* is essential for nucellar embryony. Additional genes that control the percentage of polyembryonic seed produced are surely involved, since the proportion of polyembryonic seeds varies from 10% to 75% among genotypes inferred as having nucellar embryony, results essentially similar to the model proposed by Parlevliet and Cameron (1959).

We propose that the model of Parlevliet and Cameron (1959) be modified to specify the trait as "nucellar embryony" and that *P. trifoliata* "Webber Fawcett" and "Rubidoux" are *Ne1ne1* for the locus marked by EMB-6P

and other markers. As additional modifier gene(s) are identified, it is proposed they be named, for the purposes of model building, based on the aspect of the trait upon which they have the largest effect. Therefore, if a second locus is identified that affects the inheritance of nucellar embryony, that locus should be named *Ne2*. A locus that affects the number of embryos per seed should be named *Pe1* for polyembryony and subsequently identified loci affecting percentages of polyembryony be given consecutive numbers.

The inheritance of apomixis in *Poncirus* can be described by a model with a single dominant gene *Ne1*, conferring the trait of nucellar embryony, and heterozygous in *Poncirus Ne1ne1*, but showing segregation distortion. This model also predicts a (2 polyembryonic:1 monoembryonic) segregation in *Poncirus* selfed, a ratio that is consistent with the observed 85:34 segregation of EMB-6P. It is also proposed that the percentage of polyembryonic seed is influenced by yet to be identified loci that do not prevent the *Ne1* allele from initializing cells of the nucellus to enter the embryonic pathway of development. Assuming at least one heterozygous polyembryony modifier gene in *Poncirus*, the genotype of *P. trifoliata* “Rubidoux” would be *Ne1ne1Pelpe1*.

In summary, bulked segregant analysis with AFLP markers identified a single genomic region defined by several marker loci that is strongly associated with the percentage of polyembryonic seeds produced by trees in a *Citrus* × *Poncirus* cross. It is very likely that this region contains a gene that is essential for the production of polyembryonic seeds by apomixis. However, the proportion of polyembryonic seeds varies widely among hybrids, probably due to the influence of other genes. Furthermore, the linked markers show significant segregation distortion, which may have obscured this major gene effect in previous studies that included only phenotypic data. Marker analysis of seedlings produced by some hybrid trees suggests that while polyembryony is usually associated with apomixis, twinning and perhaps other processes may also contribute to a low frequency of polyembryony.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Asins MJ, Garcia MR, Ruiz C, Carbonell EA (2002) Molecular markers for the genetic analysis of apomixis. In: Jain SM et al (eds) Molecular techniques in crop improvement. Kluwer, The Netherlands, pp 265–281
- Cameron JW, Soost RK (1979) Sexual and nucellar embryony in F1 hybrids and advanced crosses of *Citrus* with *Poncirus*. J Am Soc Hortic Sci 104:408–410
- Cho YG, McCouch SR, Kuiper M, Kang MR, Pot J, Groenen JTM, Eun MY (1998) Integrated map of AFLP, SSLP and RFLP markers using a recombinant inbred population of rice (*Oryza sativa* L.). Theor Appl Genet 97:370–380
- Esen A, Soost RK (1977) Adventive embryogenesis in *Citrus* and its relation to pollination and fertilization. Am J Bot 64:607–614
- Fang DQ, Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theor Appl Genet 95:408–417
- Fang DQ, Roose ML, Krueger RR, Federici CT (1997) Fingerprinting trifoliolate orange germ plasm accessions with isozymes, RFLPs, and inter-simple sequence repeat markers. Theor Appl Genet 95:211–219
- Frost HB (1926) Polyembryony, heterozygosis and chimeras in *Citrus*. Hilgardia 1:365–402
- Garcia R, Asins MJ, Forner J, Carbonell EA (1999) Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. Theor Appl Genet 99:511–518
- Hong QB, Xiang SQ, Chen KL, Chen LG (2001) Two complementary dominant genes controlling apomixis in genus *Citrus* and *Poncirus*. Acta Genet Sin 28:1062–1067
- Iwamasa M, Ueno I, Nishiura M (1967) Inheritance of nucellar embryony in citrus. Bull Hort Res Sta Japan 7:1–10
- Kepiro JL (2003) Molecular genetic analysis of nucellar embryony (apomixis) in *Citrus maxima* × *Poncirus trifoliata*. PhD Dissertation, University of California, Riverside
- Khan IA, Roose ML (1988) Frequency and characteristics of nucellar and zygotic seedlings in three cultivars of trifoliolate orange. J Am Soc Hortic Sci 113:105–110
- Klotz LJ (1978) Fungal, bacterial, and nonparasitic diseases and injures originating in the seedbed, nursery, and orchard. In: Reuther W et al (eds) The citrus industry, vol IV. Crop protection. University of California Press, Berkeley, pp 1–66
- Kobayashi S, Ieda I, Nakantani M (1979) Studies on the nucellar embryogenesis in *Citrus* II. Formation of the primordium cell of the nucellar embryo in the ovule of the flower bud, and its meristematic activity. J Jpn Soc Hort Sci 48:179–185
- Koltunow AM, Soltys K, Nito N, McClure S (1995) Anther, ovule, seed, and nucellar embryo development in *Citrus sinensis* cv. Valencia. Can J Bot 73:1567–1582
- Labarca C, Paigen K (1980) A simple, rapid, and sensitive DNA assay procedure. Anal Biochem 102:344–352
- Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199
- Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci U S A 88:9828–9832
- Moore GA, Castle WS (1988) Morphological and isozymic analysis of open-pollinated *Citrus* rootstock populations. J Hered 79:59–63
- Myburg AA, Remington DL (2000) Protocol for high-throughput AFLP analysis using Li-Cor IR2 automated sequencers. Forest Biotechnology Group, Department of Forestry, Department of Genetics, North Carolina State University, Raleigh
- Parlevliet JE, Cameron JW (1959) Evidence on the inheritance of nucellar embryony in citrus. Proc Amer Soc Hort Sci 74:252–260
- Ruiz C, Breto MP, Asins MJ (2000) A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. Euphytica 112:89–94

- Soost RK, Roose ML (1996) Citrus, chapter 6. In: Janick J, Moore JN (eds) Fruit breeding vol I: tree and tropical fruits. Wiley, New York, pp 257–323
- Ueno I, Iwamasa M, Nishiura M (1967) Embryo number of various citrus varieties and their relatives. Bull Hort Res Sta Japan 7:11–21
- Van Ooijen JW, Voorrips RE (2001) JoinMap® version 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Vos P, Hogers R, Bleeker M, Reijans M, Tv-d L, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
- Wakana A, Uemoto S (1988) Adventive embryogenesis in citrus (Rutaceae) II. Postfertilization development. Am J Bot 75:1033–1047
- Webb DM, Knapp SJ (1990) DNA extraction from a previously recalcitrant plant genus. Plant Mol Biol Rep 8:180–185
- Wilms HJ, Van Went JL, Cresti M, Ciampolini F (1983) Adventive embryogenesis in *Citrus*. Caryologia 36:65–78
- Xiang C, Roose ML (1988) Frequency and characteristics of nucellar and zygotic seedlines in 12 citrus rootstocks. Sci Horticulturae 37:47–59