2012

Genomics of Tropical Fruit Tree Crops

R. S. Arias
U.S.D.A. National Peanut Res. Lab., renee.arias@ars.usda.gov

James W. Borrone
Oklahoma State University

Cecile L. Tondo
Subtropical Horticulture Research Station, USDA-ARS

David N. Kuhn
USDA-ARS

Brian M. Irish
Tropical Agriculture Research Station , USDA-ARS

See next page for additional authors

Chapter 7
Genomics of Tropical Fruit Tree Crops

Renée S. Arias, James W. Borrone, Cecile L. Tondo, David N. Kuhn, Brian M. Irish, and Raymond J. Schnell

Abstract The genetic improvement of tropical fruit trees is limited when compared to progress achieved in temperate fruit trees and annual crops. Tropical fruit tree breeding programs require significant resources to develop new cultivars that are adapted to modern shipping and storage requirements. The use of molecular markers in tropical fruit tree breeding is greatly assisting in solving a number of difficult challenges for breeders such as the development of complex family structures for recombination mapping and for recurrent selection. A review of the literature on molecular markers development and new techniques for increasing single-nucleotide polymorphic markers is discussed. The development of marker-assisted breeding for these tropical tree crops is also discussed.

Keywords SNP • Microsatellites • SSR • Molecular markers • Mango • Avocado • Lychee • Longan • Linkage mapping

R.S. Arias, Ph.D. (✉)
National Peanut Research Laboratory, USDA-ARS,
1011 Forrester Dr., S.E, 39842 Dawson, GA, USA
e-mail: renee.arias@ars.usda.gov

J.W. Borrone, MS, Ph.D.
Department of Entomology & Plant Pathology, Oklahoma State University,
Stillwater, OK, USA

C.L. Tondo, Ph.D. • D.N. Kuhn, Ph.D.
Subtropical Horticulture Research Station, USDA-ARS,
Miami, FL, USA

B.M. Irish, BS, MS, Ph.D.
Tropical Agriculture Research Station, USDA-ARS, Mayaguez, PR, USA

R.J. Schnell, Ph.D.
National Germplasm Repository for Tropical/Subtropical Fruit Crops,
USDA-ARS, Miami, FL, USA
e-mail: Ray.Schnell@effem.com

R.J. Schnell and P.M. Priyadarshan (eds.), *Genomics of Tree Crops*,
DOI 10.1007/978-1-4614-0920-5_7,

Introduction

Genetic improvement of tropical fruit trees[1] has lagged far behind the progress achieved in temperate fruit trees. The reasons for this are many and include the lack of resources, political stability, and infrastructure in many of the lesser-developed countries where these species are grown. Nevertheless, the value of these tree fruit crops in providing locally available quality nutrition is significant. The commercial potential of these crops is often limited because of logistics involving transportation and storage. The commercial value of these crops could be greatly enhanced if cultivars that were adapted to commercial production, storage, and shipping were developed.

Tropical fruit trees are a major source of carbohydrates and vitamins for much of the developing world population. The world production of the major tropical fruits from trees is expected to reach 65 million tons for the year 2010, with 98% of the global production by developing countries. The production in the year 2010 for the top three tropical fruit trees, mango, papaya, and avocado, is expected to be 30.7, 12.4, and 3.1 million tons, respectively. For the minor tropical fruit trees, such as lychee, durian, rambutan, and guava, the production is smaller, and trading occurs at regional markets; therefore, there are no global statistics for these species (FAO 2003).

Compounds with a vast number of applications are present in tropical fruit trees. To mention some examples, *Garcinia mangostana* Linn. can accumulate up to 56% of oil in seeds (Hawkins and Kridl 1998). Other compounds from *Garcinia* spp. include antioxidants (Obolskiy et al. 2009) used for medicinal purposes, as well as antifungal compounds (Geetha et al. 1997) both extracted from the fruits' pericarp. The fruits of *Crataeva* spp. are high in carotenoids (Englberger et al. 2009), and the fruits of *Mangifera indica* Linn., *Carica papaya* Linn., and *Psidium guajava* Linn. are a source of carotenoids and vitamin C (Oliveira et al. 2010). Antimicrobial compounds are produced by *Diospyros blancoi* A. DC. (Ragasa et al. 2009), *Nephelium longana* Cambess. (Ripa et al. 2010), *Irvingia gabonensis* (Aubrey-Lecomte ex O. Rorke) Baill. (Kuete et al. 2007), and *Inga fendleriana* Benth. (Pistelli et al. 2009). Anticancer activity has been described for several tropical fruit tree species, for example, *Irvingia malayana* Oliv. ex A. W. Benn. has antiangiogenic properties (Ng et al. 2010), compounds from *Annona* spp. arrest cancer cells at G1 (Yuan et al. 2003), and *Morinda citrifolia* Linn. is active against cancer cells (Liu et al. 2001). Other useful applications described for tropical fruit trees are the cardioprotective activity of *Garcinia mangostana* compounds (Devi-Sampath and Vijayaraghavan 2007), the effective gastroenteritis control by *Spondias purpurea* Linn. (Caceres et al. 1993), the boost for the immune system by *Morinda citrifolia* (Palu et al. 2008), the anti-snake venom properties of *Tamarindus indica* Linn. (Ushanandini et al. 2006), and the antifatigue activity of *Dimocarpus longan* Lour. polysaccharides

[1] The tropical fruit tree species discussed in this chapter will generally exclude tree species whose commercial value is derived from non-fruit parts, are predominantly temperate species, are monocots, or are discussed in another chapter in this edition.

(Zheng et al. 2010). The importance and diversity of this group of species warrant a greatly expanded effort in genetics and genomics for compound discovery and improvement of human health.

The use of genomics in tropical fruit crops has largely been confined to the development of isozyme and dominant PCR-based markers and their use for germplasm diversity analysis and clonal fingerprinting. Information from dominant markers is of limited use in genomics applications; nonetheless, it has aided in the creation of a few genetic recombination maps and in our understanding of genetic diversity in germplasm collections. More recently, codominant microsatellite markers, also known as simple sequence repeat (SSR) markers, have been developed for a number of tropical fruit crops and have been used for parentage analysis, clonal fingerprinting, genetic diversity analysis, and development of genetic linkage maps. With the increased sequencing capacity of second- and third-generation pyrosequencing, many transcriptomes (gene space) of these species will be sequenced over the next 510 years. The generation of this new information should lead to increased interest in tropical tree fruit crops and to new opportunities to increase the rate of genetic gain in breeding programs.

Breeding tropical fruit trees is complicated by their reproductive biology. For instance, avocado (*Persea americana* Mill.) has an unusual flower behavior with two complementary types of flowering patterns, called "A" and "B," that promotes outcrossing (Bergh 1969; Davenport 1986). Avocado does not contain any self-incompatibility system, and self-pollination is frequently observed. A mature tree can produce upwards of a million flowers with only 1% of these setting fruit. A single ovary develops into the seed from a single pollination, so the generation of large numbers of seedlings by hand pollination is not practical. In mango (*Mangifera indica*), a similar situation exists; flowering is strongly influenced by weather, and some genotypes flower very irregularly. The flowers of mango are small, and both perfect and staminate (male) flowers occur in the same inflorescence. Hand pollination is possible but difficult to perform on a large scale to generate large numbers of progeny (Sharma et al. 1971 in Singh et al. 1980; Pinto et al. 2004). In addition, polyembryony in mango complicates breeding schemes. In polyembryonic cultivars, seedlings arise from nucellar tissue or from a zygote, but distinguishing between the two can be complicated (Schnell et al. 1994). For many of the tropical fruit tree species in the Sapindaceae, flower polymorphisms (i.e., monoecy, dioecy, androdioecy, gynodioecy, etc.) add complexity to the breeding process. In lychee (*Litchi chinensis* Sonn.), there are three types of flowers appearing in irregular sequence or simultaneously on the same inflorescence. There are male flowers, hermaphrodite flowers that act as females, and hermaphrodite flowers that act as males (Morton 1987). Again, the flowers are small and not amenable to hand pollination (Stern and Gazit 2003). In longan (*Dimocarpus longan*), there are three flower types: staminate (functionally male), pistillate (functionally female), and hermaphroditic (bisexual). Flowering in each panicle occurs in progressive openings of staminate (male) flowers first, then pistillate flowers followed by hermaphroditic flowers functioning as females and then hermaphroditic flowers functioning as males. Pollination is mainly by small insects but also by wind (Blanche et al. 2006).

Lack of genetic diversity in mangosteen (*Garcinia mangostana*) is a consequence of its mode of reproduction as an obligate apomict. Only anecdotal reports of male trees have been made for this dioecious species (Sand et al. 2005). In addition, an extremely long juvenile stage, upwards of 10 years, has been described for this particular fruit tree species (Poerwanto 2002).

In most of these tropical fruit tree species, controlled pollinations are very difficult to achieve, so breeders and horticulturalists have relied on the use of open-pollinated progeny for selection. The use of maternal half-sib families is very inefficient, and for most tropical fruit trees, no recurrent parentage selection has occurred. The ability to generate large numbers of codominant markers quickly using next-generation sequencing has revolutionized what can be accomplished in tropical fruit tree breeding. Identifying full-sib families is relatively simple using SSR markers. By planting trees in a polycross design (Iyer and Schnell 2009) or harvesting from commercial orchards where only a few cultivars are being grown, a breeder can now have large full-sib families to perform selection. As reported later in this chapter, the use of these types of molecular markers has greatly accelerated mango and avocado breeding efforts. Furthermore, recent efforts have focused on the development of new SSR markers for a number of tropical fruit tree species including lychee, longan, rambutan, Spanish lime, sapodilla, mamey, *Annona* spp., mangosteen, *Artocarpus* spp., and star fruit which are available to breeders and geneticists working on these crops.

Applications of Molecular Genetics and Molecular Marker Development

The main genetic tools used for the identification and breeding of cultivars of domesticated species are morphological and molecular markers (Tanksley 1993). Morphological markers in tropical fruit trees take years to be usable, for example, until the plants overcome their juvenility stage besides being subject to epigenetics and environmental factors. Molecular markers, instead, can be used from any tissue at any time during the plant growth, expediting the process of variety identification and breeding, and overcoming the limitations of traditional methods (Azofeifa-Delgado 2006). Furthermore, molecular marker analysis of tropical fruit trees help understand the past domestication of taxa, determine effective population sizes, and assess the value of a decentralized approach for future domestication (Jamnadass et al. 2009). In tropical fruit tree species for which expressed sequence tags (ESTs) are available, these can be used to develop markers such as SSRs and single nucleotide polymorphisms (SNPs) in relation to flavor, color, fragrance, vitamins, fruit softening, and other traits of interest, for example, *Actinidia* spp. (Crowhurst et al. 2008). SNPs are becoming more popular than SSRs as genetic markers in linkage analysis because they are more abundant and suitable for automatic allele calling (Novelli et al. 2004; Selmer et al. 2009). SNPs have been developed only for a few tropical fruit tree species, one in relation to resistance to papaya ringspot virus, PRSV, in *Carica papaya* Linn. (Dillon et al. 2006), another for phylogenetic studies

in *Citrus* spp. (Novelli et al. 2004) and for genotyping and linkage mapping in *Theobroma cacao* Linn. (Livingstone et al. 2010). However, at present, the cost of developing and testing SNPs is still higher than for developing SSRs. Nonetheless, SNP discovery is actively being carried out in crops such as avocado and mango.

SSRs have been the most widely employed class of molecular markers used in genetic studies with applications in many fields of genetics including genetic resources conservation, population genetics, molecular breeding, and paternity testing (Ellegren 2004). This range of applications is due to the fact that SSR markers are codominant, multiallelic, and highly reproducible; have high resolution; are amenable to high throughput; and are based on polymerase chain reaction (PCR) (Oliveira et al. 2006). As a convention, SSRs are regions in the genome where a group of bases (1–8 bp long) are repeated in tandem (Richard et al. 2008). These regions can be isolated either by data mining of existing sequences or by generating SSR-enriched libraries (Kijas et al. 1994; Zane et al. 2002). With the exception of *Carica papaya* and *Theobroma cacao* for which the genomes have been sequenced (Ming et al. 2008; Argout et al. 2010), most tropical fruit tree species do not have enough DNA sequence information to use data mining for identifying potential markers. Alternatively, expressed sequence tag (EST) information (i.e., cDNA) may be used to develop markers. We have summarized the number of entries for nucleotides, ESTs, and SSRs listed in the National Center for Biotechnology Information (NCBI) database, GenBank, for some of the most important tropical fruit trees in Table 7.1.

The use of molecular markers in tropical fruit trees can sometimes be hindered by socioeconomic reasons. One limitation is the significant cost associated with the development of markers for each crop. Tropical fruit tree species are distributed in a large number of taxonomic groups (Muchugi et al. 2008), and though transferring markers from other species could be used to reduce costs (Viruel and Hormaza 2004), such transferability is not feasible among distant taxa (Ellis and Burke 2007). In general, funds to study each tropical fruit tree species are scarce resulting in insufficient information on molecular markers. Additionally, the practical implementation of the existing results on molecular markers is often limited by the absence of guidance on how to best apply them (Muchugi et al. 2008). For example, from the top 25 indigenous tropical fruit tree species identified as priority by the International Centre for Research in Agroforestry (ICRAF), only for eight has some work been done using molecular markers (Jamnadass et al. 2009). For most of these species, the amount of genetic information available is negligible or null. A list of the main tropical fruit tree species and the molecular markers that have been developed for each of them is shown in Table 7.2.

SSR Isolation Using Pyrosequencing

The isolation of SSRs from species for which little to no genetic information is available, such as most tropical fruit trees, can be difficult. At the USDA-ARS Mid South Area Genomics Laboratory (MSAGL), an effective pipeline has been created to isolate SSR markers from these species. The process first involved a slight

Table 7.1 Summary of number of entries in NCBI for nucleotides, ESTs, and SSRs for some of the most important tropical fruit trees

Species	Nucleotide	EST	SSR	Plant type
Actinidia deliciosa + *A. arguta* (*kiwi*)	230 + 78	57,751 + 7,257	30	Vine
Carica papaya	51,217	77,393	45+	
Musa spp. and hybrids	4,210	31,268	550	Monocot
Citrus spp.	2,592	549,188	106	Temperate
Persea americana	493	16,558	0	Tropical tree
Mangifera indica	401	68	462	Tropical tree
Cocos nucifera	382	6	0	monocot
Diospyros kaki	296	9,474	69	Tropical tree
Citrus grandis (*C. maxima*)	202	0	17	Tropical tree
Spondias purpurea	141	0	0	Tropical tree
Dimocarpus longan	137	66	0	Tropical tree
Annona cherimolia	120	0	97	Tropical tree
Inga edulis	95	0	5	Tropical tree
Litchi chinensis	88	0	27	Tropical tree
Psidium sp. (guava)	75	0	24	Tropical tree
Bactris gasipaes	71	0	46	Monocot
Durio spp.	66	0	7	Tropical tree
Garcinia mangostana	64	149	0	Tropical tree
Morinda citrifolia	62	0	0	Tropical tree
Adansonia spp.	55	0	18	Tropical tree
Passiflora edulis (passion fruit)	51	0	10	Vine
Manilkara zapota	34	0	0	Tropical tree
Averrhoa carambola	31	0	0	Tropical tree
Annona muricata	30	0	0	Tropical tree
Anacardium occidentale	30	0	21	Tropical tree
Tamarindus indica	25	0	0	Tropical tree
Irvingia spp.	22	0	0	Tropical tree
Nephelium lappaceum	14	0	0	Tropical tree
Pometia pinnata	14	0	0	Tropical tree
Artocarpus heterophyllus	10	1	0	Tropical tree
Annona squamosa	8	0	4	Tropical tree
Artocarpus altilis	7	1	0	Tropical tree
Nephelium ramboutan-ake	7	0	7	Tropical tree
Annona reticulata	6	0	4	Tropical tree
Melicoccus bijugatus	6	0	0	Tropical tree
Syzygium samarangense	5	0	0	Tropical tree
Premna serratifolia	3	0	0	Tropical tree
Sandoricum koetjape	2	0	0	Tropical tree
Garcinia portoricensis	1	0	0	Tropical tree
Pouteria sapota	1	0	0	Tropical tree
Spondias dulcis	1	0	0	Tropical tree
Dracontomelon vitiense	0	0	0	Tropical tree
Garcinia binucao	0	0	0	Tropical tree
Garcinia cochinchinensis	0	0	0	Tropical tree
Crataeva speciosa	0	0	0	Tropical tree

Species were sorted first by decreasing number of nucleotide entries and then by decreasing number of ESTs. The *orange highlight* corresponds to species considered minor fruit crops, underutilized or rare, depending on the region, by the International Tropical Fruits Network (www.itfnet.org). The *green highlight* corresponds to species economically important that are either called "trees" though they are monocots, that are dicots but not exactly trees, or that are temperate

Table 7.2 Summary table of publications on molecular markers per tropical fruit tree species

Species	Work done on molecular markers
Carica papaya L.	CAPS marker linked to PRSV-P resistance (45); 50 accessions analyzed using nine isozyme systems (Ocampo et al. 2006); high-density genetic map with 712 SSRs and 277 AFLPs (Blas et al. 2009); draft of complete genome (97)
Persea americana Mill.	25 SSRs used on 37 cultivars and wild relatives (11); SNP development based on resequencing of 4 nuclear loci of 21 wild accessions of *P. americana* (Chen et al. 2008); screening of 56 SSR markers on *P. americana* (162)
Mangifera indica L.	15 SSR markers tested on 59 *M. indica* cultivars (137); 28 SSRs tested on 15 *M. indica* cultivars (46); 16 SSRs tested on 28 mango cultivars (157); 19 SSRs used on 307 *M. indica* accessions (Duval et al. 2006); six populations of *M. indica* analyzed by RAPD (Diaz-Matallana et al. 2009); 11 SSRs used in the phylogeographical analysis of *M. indica* (Hirano et al. 2010)
Diospyros kaki L. f.	Six SSRs used to screen 12 genotypes (Guo and Luo et al. 2008)
Citrus maxima L. Osbeck	370 accessions, 76 *C. maxima*, tested with 24 SSR markers (Barkley et al. 2006); genetic map of *Citrus* based on 256 RAPD markers used on 94 hybrids (de-Oliveira et al. 2005); putative SNPs annotated for *Citrus* (Martinez-Godoy et al. 2008)
Spondias purpurea, S. dulcis	Chloroplast spacer *trn*G-*trn*S identified five haplotype varieties under cultivation that were not present and probably were lost to extinction in the wild (Miller and Schaal et al. 2005); 216 individuals of 34 populations with two primer pair RFLPs (Miller and Schaal et al. 2006)
Dimocarpus longan Lour.	28 RAPD markers used to uniquely identify each of 22 accessions (163); partial rbcL 66 AFLP markers used on 41 *D. longan* accessions (82)
Annona cherimolia Mill., A. muricata L., *A. squamosa* L.	Development of monomorphic SSRs as phylogenetic markers of *Annona* species (Chatrou et al. 2009); 13 isozyme loci used to study segregation in 14 self-fertilized cultivars (Perfectti and Pascual et al. 1996); 206 cultivars analyzed with 23 isozyme loci in a germplasm bank (Perfecti and Pascual et al. 2005); 94 new SSRs, 58 of them polymorphic, were tested on 23 cherimoya cultivars (Escribano et al. 2008); Nine accessions discriminated by 14 polymorphic RAPD fragments (Brown et al. 2003); 5 *Anonna* species analyzed with 11 allozymes indicated *A. muricata* more genetically distant from the rest (Samuel et al. 1991)
Inga edulis	ITS sequencing and phylogenetic study of 32 species of the seven sections of Inga (Richardson et al. 2001); five SSR loci to study 189 trees for possible genetic erosion (Hollingsworth et al. 2005); chloroplast trnL-F region and five SSR loci used to study five locations in the Peruvian Amazon (Dawson et al. 2008)
Litchi chinensis Sonn.	12 SSRs tested on 21 *Litchi* cultivars (159); 16 SSR markers tested on 58 *Litchi* cultivars (81)
Psidium spp.	23 SSRs tested on 16 accessions of *Psidium guajava* and three other species (Risterucci et al. 2005); RAPD markers associated to quercetin (Feria-romero et al. 2009); AFLP on 48 guava cultivars (Hernández-Delgado et al. 2007); revision of molecular tools applied to *Psidium* (Rai et al. 2010)
Durio spp. Adans.	Ten *Durio* species were analyzed for phylogenetic relationships using RFLP on two chloroplast genes ndhC-trnV and rbcL (Santoso et al. 2005); phylogenetic relationships inferred from analysis of ndhF and ITS sequences (Nyffeler & Baum 2000)

(continued)

Table 7.2 (continued)

Species	Work done on molecular markers
Garcinia mangostana, G. cochinchinensis, G. portoricensis, G. binucao	Random amplified DNA fingerprinting (RAF) used on 37 accessions of *G. mangostana* (Ramage et al. 2004); 21 trees analyzed with 5 primers for RAPD analysis (Sobir et al. 2007)
Morinda citrifolia	11 individuals of 3 varieties of *M. citrifolia* analyzed to determine geographical origin by using nrETS, nrlTS, rps16 and trnT-F sequence data (Razafimandimbison et al. 2010)
Adansonia spp. L.	AFLP analysis of 137 individuals from six populations (Assogbadjo et al. 2006); 18 SSR markers to analyze 214 individuals of *Adansonia digitata* and 30 individuals of other *Adansonia* species (Larsen et al. 2009); 11 populations of 4 countries in West Africa analyzed by AFLP (Kyndt et al. 2009); phylogeography of 344 individuals from 74 populations analyzed by chloroplast DNA (Tsy et al. 2009)
Manilkara zapota (L.) van Royen	Analysis of four populations of *M. zapota* from Mexico using 4 polymorphic RAPD markers (Heaton et al. 1999); 12 polymorphic SSR loci developed for *M. huberi* (Azevedo et al. 2005)
Averrhoa carambola L.	None found
Anacardium occidentale	Fingerprinting of 19 *A. occidentale* accessions using 50 RAPD primers, 12 ISSRs, and 6 AFLPs (Archak et al. 2003); 91 individuals of *A. occidentale* analyzed by AFLP (Archak et al. 2009); 21 polymorphic SSRs from *A. occidentale* tested on other *Anacardium* species (Croxford et al. 2006)
Tamarindus indica L.	RAPDs used to analyze ten populations of *T. indica* (Diallo et al. 2007)
Irvingia spp.	Eight CAPS primers used to analyze *I. gabonenensis* and *I. wombolu* (Lowe et al. 1998); 130 individuals of *I. gabonensis* and *I. wombolu* from West Africa were analyzed by RAPDs (Lowe et al. 2000); AFLP analysis of 15 accessions of *I. gabonensis* (Ude et al. 2006)
Nephelium lappaceum and *N. ramboutan-ake*	Seven SSR markers from *Litchi chinensis* had amplification in *N. ramboutan-ake* (Sim et al. 2005)
Pometia pinnata	None found
Artocarpus altilis (Parkinson), *A. heterophyllus* Lam.	26 *A. heterophyllus* accessions analyzed by AFLP using 12 primer pairs (136); Six populations of *A. altilis* analyzed by AFLP and 15 morphological traits (Sreekumar et al. 2007); Eight primer pairs used for AFLP analysis of 50 *A. heterophyllus* accessions (Shyamalamma et al. 2008); AFLP analysis of 200 breadfruit samples of *A. camansi* and *A. mariannensis* (166)
Melicoccus bijugatus Jacq.	None found
Syzygium samarangense	Isozymes used on a related rainforest species *S. nervosum* (Shapcott et al. 1999); 8 SSRs developed for *S. sayeri* (Hillyer et al. 2007)
Premna serratifolia L.	None found
Sandoricum koetjape	None found
Pouteria sapota	20 cultivars of *M. zapota* analyzed by RAPD markers (Meghala et al. 2005)
Dracontomelon vitiense	None found
Crataeva speciosa Volkens	None found

The order of species is the same as in Table 7.1, in which NCBI entries are summarized

modification to the DNA extraction method, given the presence of copious latex and phenolics in the vegetative tissues. Second, for generating SSR-enriched libraries, the method previously developed by Techen et al. (2010) was modified to adapt to high-throughput pyrosequencing with a Roche 454 GS-FLX (F. Hoffmann-La Roche Ltd., Basel, Switzerland). One modification employs two adapters (Techen et al. 2010) that allow simultaneous loading of pairs of samples in the same region of picotiter plates; the adapters act as bar coding to separate the samples via bioinformatics. Another modification is reducing the number of PCR cycles during library preparation to minimize redundant sequences. SSRs were isolated from the following species/crop groups: *Nephelium lappaceum* Linn. (rambutan), *Manilkara zapota* Linn. (sapodilla), *Pouteria sapota* Jacq. (sapote), *Litchi chinensis* (lychee), *Melicoccus bijugatus* Jacq. (Spanish lime*), Annona squamosa* Linn. (sugar apple), *Dimocarpus longan* (longan), *Averrhoa carambola* Linn. (star fruit), *Artocarpus altilis* (Parkinson) Fosberg (breadfruit), and *Garcinia mangostana* (mangosteen).

For the combined species, a total of 2,510,291 reads were assembled into 224,815 contigs in which a total of 49,898 SSR repeats (not including mononucleotides) were detected. A total of 10,310 primer pairs were designed under stringent conditions (Tm 65°C and 3′ GC clamp). Only 384 primer sets were tested per species. This was done to identify and select the SSR markers that were most readily amplifiable, reproducible, and which detected the highest allelic diversity for variety/cultivar identification and plant breeding. The SSR development process was accomplished in a short amount of time (less than 6 months), and greater than 96% of the markers resulted in SSR amplification. A detailed protocol for the SSR development procedure is available in Methods in Molecular Biology, Humana Press (Arias et al. 2010).

Specific Examples of Genomics in Tree Fruits

Avocado

The avocado (*Persea americana*) is an evergreen subtropical tree that is native from Mexico to northern South America and produces a fruit that is unique and nutritious. This fruit was known by the Aztecs as “ahuacacuauhitl,” which was later shortened by the Spaniards to “aguacate.” In the United States, avocado was introduced to Florida in 1833, California in 1848, and Hawaii by 1855 (Nakasone and Paull 1998). Major commercial production of avocado in the United States is limited to California and Florida. In 2000, global production exceeded 2.4 MMT, and the major producers were Mexico, Indonesia, South Africa, and the USA (Anonymous 2002).

Persea americana has been subdivided into three horticultural groups: Mexican (*P. americana* var. *drymifolia* (Schect. & Cham.) Blake), Guatemalan (*P. americana* var. *guatemalensis* Wms.), and West Indian (*P. americana* var. *americana* Mill.)

races. The West Indian race is known to be from the lowland areas of the Pacific coast of Central America and not the West Indies, while the Guatemalan and Mexican races are native to specific highland areas within each country (Scora and Bergh 1992). The three racial groups can be distinguished by the percentage of oil content in the fruit, with the West Indian cultivars ranging from 2.5% to 8.0%, Guatemalan accessions from 10% to 13%, and Mexican accessions ranging from 15% to 20% (Knight 2002). The racial classes also vary phenotypically for characters such as fruit size and shape, skin thickness, skin color, seed size, and fruit ripening (Lahav and Lavi 2002). Sterility barriers do not exist between or among the three racial types (Lahav and Lavi 2002). Avocado has a distinct flowering habit known as protogynous, diurnally synchronous dichogamy (Bergh 1969). This type of reproductive behavior promotes outcrossing; however, significant amounts of self-pollination are known to occur in commercial plantings (Davenport et al. 1994; Borrone et al. 2008; Schnell et al. 2009). Named cultivars often originate from open-pollinated seedlings. The unknown pollen parent has often been estimated based on the flower types of available donor trees. Morphological characters have been used to infer parentage, although these are influenced by environmental factors and may not unambiguously distinguish closely related genotypes or interracial hybrids. Many of the cultivars grown in Florida are interracial hybrids between Guatemalan and West Indian types, while those grown in California are hybrids between Mexican and Guatemalan types.

The haploid genome size of avocado has been estimated to be 8.83×10^8 bp (Arumuganathan and Earle 1991a). Avocado contains 24 chromosomes with bivalent pairing at meiosis indicating that $n = 12$ (Darlington and Wylie 1945). Avocado has been proposed to be the result of an ancient polyploid event (Chanderbali et al. 2008). However, evidence from analyses of germplasm collections and seedling populations using SSR markers has demonstrated diploid inheritance for most markers, although some primer pairs have amplified more than one locus (Schnell et al. 2003; Borrone et al. 2007, 2008, 2009).

Gene Discovery

Recently, a large number of sequences have been generated from the transcriptome of *P. americana* from two projects, the Floral Genome Project (Albert et al. 2005) and the Ancestral Angiosperm Genome Project (Wall et al. 2008). The Floral Genome Project, initiated to investigate the evolutionary development of floral patterning (Chanderbali et al. 2009), produced two directionally cloned, non-normalized, Sanger-sequenced cDNA libraries and are deposited in the EST database of GenBank (16,558 sequences). Both libraries, Pam01 (8,735 sequences) and Pam01b (7,823 sequences), were developed from premeiotic flower buds. The Pam01 library was the source of the EST-SSRs used to evaluate outcrossing rates (Borrone et al.

2007, 2008; Schnell et al. 2009) and to develop a moderately dense genetic map of avocado (Borrone et al. 2009).

The Ancestral Angiosperm Genome Project continued where the Floral Genome Project left off using pyrosequencing (Wall et al. 2008). A total of 12 libraries, developed from various tissue types and stages, generated 1,698,670 "passing" sequences covering 490.7 Mb. The unprocessed, unfiltered, raw data is deposited in GenBank as 14 Sequence Read Archives (SRAs). The processed (filtered, trimmed, and quality checked) sequence information is readily obtained from the Ancestral Angiosperm Genome Project website (http://ancangio.uga.edu/content/aagp-home). The website includes a detailed description of each library, a number of tools allowing comparison of the libraries with one another, and retrieval of individual reads and contigs. The entire dataset derived from pyrosequencing has been assembled with the cDNA libraries (Pam01 and Pam01b, developed from the Floral Genome Project) in a "Sanger-454 hybrid" assembly of 115,151 contigs. Polymorphic regions, SNPs and SSRs, have been identified in the assembled sequences as part of the pipeline process. The entire assembly is available for download at the site. Additional cDNA libraries have been developed from *Persea americana* var. *drymifolia* but are not yet publically available (López-Gómez et al. 2007).

An additional resource, not yet deposited in the NCBI database, is sequences of small RNAs (Blake Myers and Pamela Green, Delaware Biotech Institute, University of Delaware; NSF Plant Genome Program Award #0638525). Small RNAs, 21–24 nucleotides long, consist of two major types known as small interfering RNA (siRNA) and microRNA (miRNA) and are implicated in regulation of gene transcription/translation. These sequences were generated by pyrosequencing (Accerbi et al. 2010). A total of 12,689,456 sequences, 5,918,358 identified as being distinct, have been generated from three separate libraries: leaves (3,800,961 total; 1,263,610 distinct), flower (5,325,501 total; 2,870,885 distinct), and fruit (3,562,994 total; 1,783,863 distinct). The sequences are available through the Comparative Sequencing of Plant Small RNAs website (http://smallrna.udel.edu/).

Marker Development

Previously, there was a limited set of SSR markers for *P. americana* (Sharon et al. 1997; Ashworth et al. 2004), some of which do not consistently amplify in all varieties (Ashworth and Clegg 2003; Ashworth et al. 2004). For example, 14 of 39 SSR markers tested were suitable to fingerprint diverse collections of *P. americana* (Schnell et al. 2003). To increase the number of informative SSR markers, publicly available *P. americana* ESTs were screened (Borrone et al. 2007, 2009) bringing the number of markers available to upwards of 300. Data-mining SSRs from expressed sequence tags (ESTs) has proven effective for generating markers for fingerprinting, genetic mapping, and comparative mapping among species (Varshney et al. 2005).

Linkage Mapping

The first linkage map of avocado was produced by Sharon et al. (1997) from 50 progeny of a cross between 'Pinkerton' and 'Ettinger' using 50 SSR markers, 17 random amplified polymorphic DNA (RAPD) markers, and 23 minisatellite DNA fingerprint (DFP) markers. Twelve linkage groups with 34 mapped loci covering 352.6 centimorgans (cM) were identified. Seven linkage groups contained two markers, two linkage groups contained three markers, one linkage group contained four markers, and two linkage groups contained five markers each. A larger population and an increased number of genetic markers were needed to produce a linkage map useful for quantitative trait loci (QTL) discovery. The development of over 300 SSR markers (Sharon et al. 1997; Ashworth et al. 2004; Borrone et al. 2007, 2009) enabled the development of a partially saturated genetic linkage map and the potential identification of QTLs controlling horticultural traits of interest in avocado. Mapping populations have been developed in California (Ashworth et al. 2007; Chen et al. 2007) and in Spain (Viruel et al. 2007) by producing full-sib families that are similar in size to those used by Sharon et al. (1997) and are focused upon Guatemalan-Mexican hybrids.

In Florida, a large population of seedlings from a commercial grove interplanted with two cultivars of opposite flowering types, 'Tonnage' (type B) and 'Simmonds' (type A), in approximately equal numbers was screened to determine the outcrossing rate in avocado under south Florida conditions (Borrone et al. 2008). Eight fully informative SSR markers identified 870 seedlings as progeny of a reciprocal cross between 'Tonnage' and 'Simmonds'. Using these seedlings, the first large mapping population and linkage map for QTL discovery were developed in avocado, focusing upon West Indian-Guatemalan hybrids (Borrone et al. 2009). The final linkage map (Fig. 7.1, Table 7.3) was constructed from 163 markers generated by 135 primer pairs, 112 designed from EST-SSRs and 23 SSR primers previously developed by Sharon et al. (1997). Twenty primer pairs amplified more than one locus, with 15 amplifying two loci, four amplifying three loci, and one amplifying six loci. 'Tonnage' was heterozygous for 92.0% of the total markers used in its sub-composite map, whereas 'Simmonds' was heterozygous for 41.0% of the total markers in its sub-composite map, comparable with the previous estimate for heterozygosity for each cultivar using 12 SSR markers (Borrone et al. 2008). Eighteen markers were fully informative with four polymorphic alleles between the two parents, and the other 145 markers were partially informative. One hundred nineteen markers were scored on both sub-composite maps. An additional 43 markers were scored only on the 'Tonnage' × 'Simmonds' progeny, and one additional marker was scored only on the 'Simmonds' × 'Tonnage' progeny.

Twelve linkage groups (LG) representing the haploid set of 12 chromosomes of *P. americana* were generated from 163 markers, at a minimum logarithm (base 10) of odds (LOD) score of 4.0 (Fig. 7.1, Table 7.3). Linkage groups ranged in size, longest to shortest, from 157.3 (LG2) to 2.4 cM (LG12), and the number of markers mapped per group ranged from 29 (LG1) to 2 (LG12). The total length of the Florida

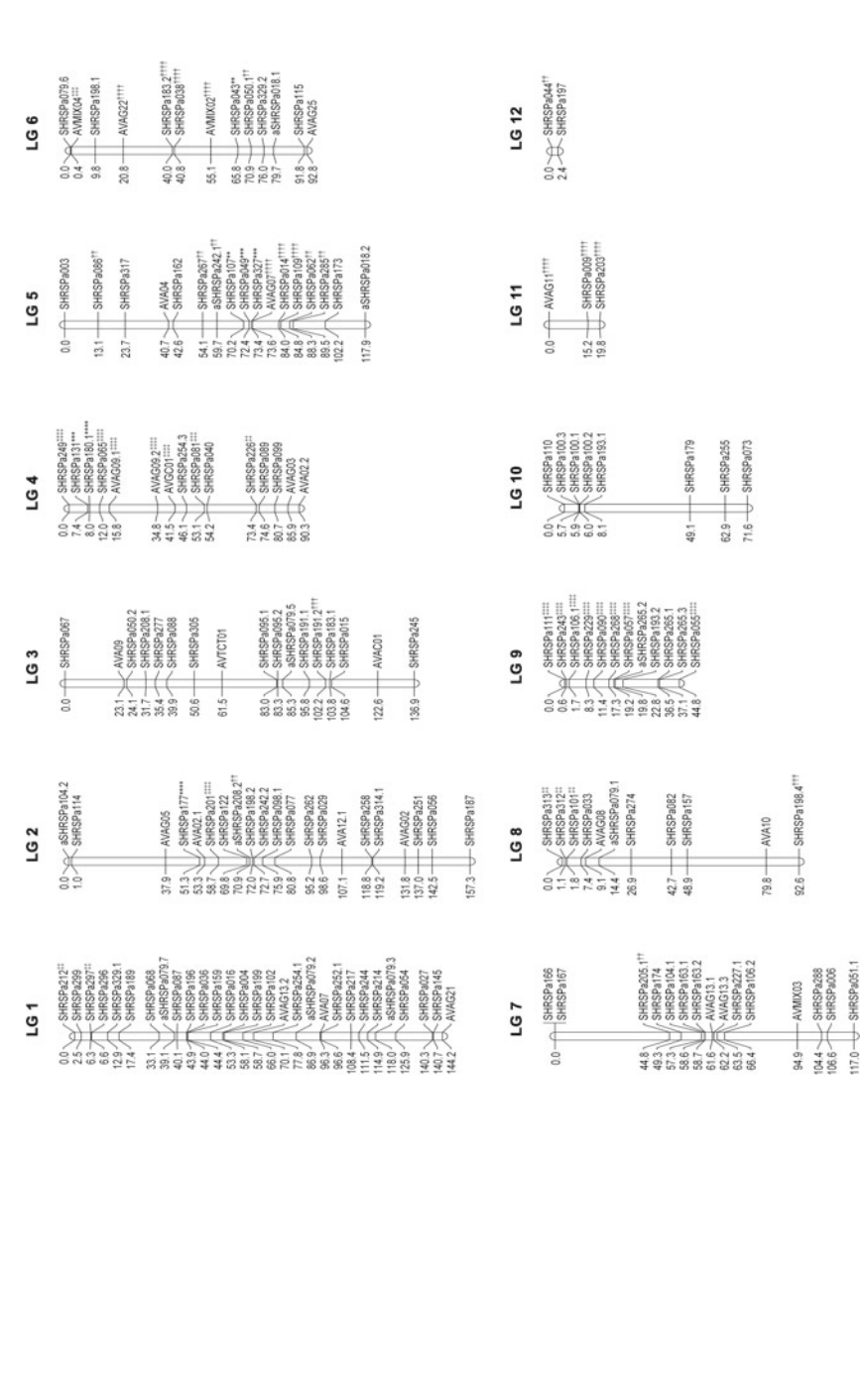

Fig. 7.1 Avocado linkage map constructed from 163 markers generated by 135 primer pairs

Table 7.3 *Persea americana* linkage mapping results for the Florida mapping population. Mapping results of the reciprocal cross (sub-composite maps) and parental maps

Map	LG	1		2	3	4	5			6	7	8	9	10	11	12	Total
'Tonnage' × 'Simmonds'							**5a**[a]	**5b**	**5c**								
456 individuals	No. loci	29		21	16	15	12	3	2	13	15	11	12	8	3	2	162
	Length (cM)	143.9		170.1	138.7	95.7	117.3	17.6	29.3	93.7	119.6	95.3	46.6	71.3	23.7	3.1	1,119.0
	cM/locus	5.0		8.1	8.7	6.4	9.8	5.9	14.7	7.2	8.0	8.7	3.9	8.9	7.9	1.6	7.5
'Tonnage' maternal	No. loci	27		18	16	15	15			12	13	11	11	8	3	2	149
	Length (cM)	146.5		154.7	138.6	96.3	116.5			92.1	121.1	95.4	47.0	70.9	23.8	3.1	1,106
	cM/locus	5.4		8.6	8.7	6.4	7.8			7.7	9.3	8.7	4.3	8.9	7.9	1.6	7.1
'Simmonds' paternal							**5a**	**5b**									
	No. loci	16		11	3	4	3	3		7	5	0	6	2	0	0	60
	Length (cM)	144.8		113.6	38.5	35	19.9	41.9		100.8	68.6	–	48.4	46.9	–	–	658.4
	cM/locus	9.1		10.3	12.8	8.8	6.6	14.0		14.4	13.7	–	8.1	23.5	–	–	12.1
'Simmonds' × 'Tonnage'							**5a**	**5b**									
259 individuals	No. loci	20		16	11	13	10	2		9	12	8	8	6	3	2	120
	Length (cM)	141.1		132.2	115.5	79.6	84.4	14.8		89.2	99.1	74.6	39.8	68.9	13.0	1.2	953.4
	cM/locus	7.1		8.3	10.5	6.1	8.4	7.4		9.9	8.3	9.3	5	11.5	4.3	0.6	7.4
'Simmonds' maternal		**1a**	**1b**														
	No. loci	11	2	8	0	4	6			6	3	0	5	0	0	0	45
	Length (cM)	69.1	16.5	92.6	–	41.3	103.6			92.3	42.6	–	44.3	–	–	–	502.3
	cM/locus	6.3	8.3	11.6	–	10.3	17.3			15.4	14.2	–	8.9	–	–	–	11.5
'Tonnage' paternal	No. loci	18		14	11	13	10			9	11	8	8	6	3	2	113
	Length (cM)	141.1		124.6	115.4	77.5	106.3			85.6	99.5	74.6	37.1	68.8	13	1.2	944.7
	cM/locus	7.8		8.9	10.5	6.0	10.6			9.5	9.0	9.3	4.6	11.5	4.3	0.6	7.7

[a]LG 5a represents the entire length of the LG in the 'Tonnage' × 'Simmonds' map and was used to calculate the overall length of the linkage map. For all others, LG5 are smaller-sized fragments of the entire LG

F_1 map was 1,087.4 cM. This Florida F_1 map is three times the size of the map reported by Sharon et al. (1997) and contains five times as many markers. Comparison of markers shared between the Florida F_1 map and the F_1 map of Sharon et al. (1997) allowed the identification of several analogous linkage groups. Ten linkage groups of the 12 described by Sharon et al. (1997) contained SSR markers. Nine of the ten corresponded to seven linkage groups in the Florida F_1 map. Distances between SSR markers within LGs reported by Sharon et al. (1997) and distances between these same markers placed in the sub-composite maps and the Florida F_1 map corresponded well, which is remarkable given the disparity in the population sizes and the numbers of markers between the two maps. The 'Tonnage' × 'Simmonds' linkage groups were longer than the 'Simmonds' × 'Tonnage' linkage groups (Table 7.3). This can be attributed to the larger number of markers mapped in the 'Tonnage' × 'Simmonds' sub-composite map versus its reciprocal (162 vs. 120) and to the larger number of individuals in this cross (456), slightly less than double that of the 'Simmonds' × 'Tonnage' cross (259). Also, LG5 formed three linkage groups in the 'Tonnage' × 'Simmonds' map and two linkage groups in the 'Simmonds' × 'Tonnage' map. The parts of this linkage group reformed a single linkage group in the composite map. Five linkage groups (LGs 3, 4, 8, 11, and 12) were composed almost entirely from markers polymorphic in the 'Tonnage' parent. The formation of linkage groups for all maps was straightforward except for a small number of markers due to the high homozygosity of 'Simmonds'.

The number of SSR markers now available for avocado allowed the development of this first moderate-density map composed of 163 loci. Linkage maps developed using molecular markers, such as this Florida F1 map, enable the detection and use of QTLs affecting traits of economic importance. The two parental cultivars 'Simmonds' and 'Tonnage' differ for many phenotypic traits useful in humid lowland environments like south Florida. 'Simmonds' is of the West Indian race and believed to be a seedling of 'Pollock'. It was selected in south Florida, propagated commercially in 1921, and is still a major commercial cultivar some 90 years later. 'Simmonds' has a light green, oblong-oval to pear-shaped large fruit with medium-sized seed and low oil content (3–6%), imparts tolerance to *Phytophthora* root rot (PRR) in its progeny, has the "A" flowering type, and is a high-yielding, early-season (June–July) cultivar (Campbell and Malo 1978; Ploetz et al. 2002). West Indian selections were the only important commercial cultivars in Florida until the 1920s when competition from Cuba depressed the market for Florida avocados. A number of Guatemalan-West Indian hybrids had since been selected which ripened in the fall and winter, extending the season and plantings shifted to include these hybrids. One of these was the cultivar 'Tonnage' which is now considered a minor cultivar in south Florida. 'Tonnage', a seedling of 'Taylor', was first propagated commercially in 1930. It has a dark green, pear-shaped fruit with medium seed size and moderate oil content (8–15%), does not impart tolerance to PRR in its progeny, is a "B" flowering type, and is a late-season (August–September) cultivar (Ploetz et al. 2002).

Recently, the USDA-ARS Subtropical Horticulture Research Station (SHRS) in Miami, FL, has produced another large mapping population of 'Hass' × 'Bacon' (~1,000 individual seedlings). 'Hass' is the most important avocado cultivar worldwide (Ashworth et al. 2007), and growers receive a premium price for 'Hass' fruit.

A commercial California ‘Hass’ avocado orchard was identified with adjacent ‘Bacon’ pollinizer rows. ‘Hass’ is an interracial hybrid of mostly Guatemalan ancestry, and ‘Bacon’ is of the Mexican landrace; they are of opposite flowering types (A and B, respectively) and, thus, expected to outcross. The full-sib family members from this ‘Hass’ × ‘Bacon’ cross were identified by genotyping with five SSR markers, and seedling trees, product of self-pollination, were excluded. However, a map has not yet been produced. Presently, the limiting factor is the availability of sufficient molecular markers to make a saturated, high-resolution map for QTL mapping of important agronomic traits such as oil content and composition, flowering type, flowering time, and cold tolerance. Once the field data for agronomic traits is collected on this population, this map will be useful for detection and use of QTLs.

SNP Discovery

SNP discovery from transcriptome and genomic sequence data is needed for obtaining sufficient molecular markers for complete linkage map saturation. The following strategy has been used at SHRS to generate and validate new SNP and SSR markers. The strategy takes advantage of two next-generation sequencing platforms, Illumina GAII (Illumina, Inc., San Diego, CA) and Roche 454, and uses both transcriptome and genomic sequence data to identify and validate SSRs and SNPs. Once sufficient SNPs have been identified (5–10 K), the production of an Illumina Infinium oligonucleotide array for genotyping of mapping populations may be employed. ‘Hass’, the most important commercially grown avocado cultivar in the world, has been selected as the reference cultivar for the transcriptome sequencing.

Transcriptome Sequencing

RNA was isolated from leaves, unopened flowers, female flowers, and male flowers from ‘Hass’, ‘Bacon’, ‘Simmonds’, and ‘Tonnage’, from tissue collected from genotyped clones at the SHRS. The RNA from the ‘Hass’ tissue was pooled for Roche 454 sequencing to generate the reference transcriptome. In addition, a large EST library was produced by Roche 454 sequencing of developing mesocarp tissue of ‘Hass’ and that data was used to expand the ‘Hass’ reference transcriptome (John Ohlrogge, Michigan State Univ., personal communication 2010). The RNA from ‘Bacon’, ‘Simmonds’, and ‘Tonnage’ leaves and flowers was pooled for each clone and sequenced by Illumina GAII. Illumina reads were aligned to the ‘Hass’ reference transcriptome, and a variant report with single-nucleotide polymorphisms and indels generated. The variant report was filtered for heterozygous SNP loci, as a locus that is heterozygous in any parent of the mapping populations can be mapped. SSR motifs are identified from the ‘Hass’ transcriptome sequences as described

earlier in this chapter. The SNP and SSR discovery will yield a sufficient number of markers to completely saturate the genetic maps for both populations.

Reduced Representational Sequencing of the 'Hass' Genome

Reduced representational sequencing of the genome of 'Hass' from hypomethylated DNA is being carried out using Roche 454 pyrosequencing to obtain a ~2–4X coverage of the avocado gene space. Hypomethylated DNA is actively transcribed, and this should enrich for gene-containing regions of the genome. In the cacao SNP and SSR discovery project (http://www.cacaogenomedb.org), the genomic sequence has proved essential for identifying intron-exon junctions in the transcriptome sequences prior to SNP oligonucleotide design or SSR primer design. In addition, the assembled gene space of 'Hass' will be made available to the avocado genome sequencing group, producing a hypomethylated library of the Mexican landrace cultivar by an alternative method. Such a comparison will be of great interest for future genome sequencing projects targeting only the gene space.

Genome Sequencing Project

The avocado genome sequencing project has been under way for over a year in Mexico at the Plant Biotechnology Unit, Centre for Research and Advanced Studies (CINVESTAV; Luis Herrera-Estrella and Gustavo Hernandez), as well as by the biofuels research on non-seed oils group in the Plant Lipid Metabolism lab at Michigan State University (John Ohlrogge). The Mexican avocado genome sequencing project is using a highly homozygous individual from the Mexican landrace (*P. americana* var. *drymifolia*). Using flanking sequences of the 163 SSR markers mapped in the Florida F1 population, CINVESTAV has designed overgo probes to anchor their physical map to the genetic recombination map. The current version of the avocado physical map has 5,008 contigs, and the estimated coverage is 5X. Using the 163 markers will certainly help in the anchoring process, but thousands of markers would allow anchoring and correct ordering of the entire physical map of avocado. The Illumina Infinium chip being designed from the SNP development project will saturate the avocado genetic recombination map and anchor physical map contigs.

Mango

Mango (*Mangifera indica*) is a significant tree fruit crop grown commercially in tropical and subtropical areas of many countries. It has been under cultivation in India for at least 4,000 years, and over 1,000 varieties are known to exist in that country

(Mukherjee 1953). India is the largest producer of mangos with 10.8 million MT accounting for 41% of the world's mango production (Gunjate 2009, Proc VIII int. mango sym.). Most commercial mango cultivation is in tropical and subtropical regions; however, mangos are also cultivated in a wide range of marginal climatic areas throughout the world. Mango genotypes are divided into two distinct categories or types based on their origin: monoembryonic mangos, which are mostly subtropical (Indian types), and polyembryonic mangos, which are mostly tropical (Southeast Asian types). The seeds of Indian types characteristically contain a zygote embryo, and the fruit skin is highly colored (mixes of red, purple, and yellow), while the seeds of Indo-Chinese types contain several nucellar embryos, and the skin is soft or pale in color (green to light green to yellow) (Iyer and Degani 1997; Viruel et al. 2005).

Information on the cytology of mango is quite limited. Only *Mangifera* species *M. indica, M. caloneura, M. sylvatica, M. foetida, M. caesia, M. odorata,* and *M. zeylanica* have been studied, and these were found to have chromosome numbers of $2n = 40$ and $n = 20$ (Mukherjee 1950, 1957, 1963). Chromosome numbers and ploidy status of other *Mangifera* species have yet to be studied (Bompard and Schnell 1997). Mango has a small haploid genome size (0.91 pg), which is three times larger than *Arabidopsis thaliana* (L.) Heynh. and comparable to that of rice (Arumuganathan and Earle 1991b). Mango has been referred to as an allopolyploid based on the conclusions drawn by Mukherjee (1950). Due to the presence of secondary associations at metaphase of meiosis, he suggested that the basic chromosome number of *Mangifera* is $n = 8$. In addition, the high number of somatic chromosomes and the correspondingly high number of nucleolar chromosomes led him to conclude that mango is an allopolyploid. However, the evidence used to arrive at this conclusion is not unequivocal. In fact, the molecular marker evidence is antithetical to this conclusion. Results from Duval et al. (2005), Schnell et al. (2005, 2006), and Viruel et al. (2005) all resulted in normal diploid segregation for SSR markers.

Applications of Molecular Genetics

Isozymes were the first markers to be used for fingerprinting mango cultivars to determine self- vs. cross-pollination and to estimate genetic relationships (Degani et al. 1990; Knight and Schnell 1994). RAPD markers were also used to fingerprint cultivars and estimate genetic relationships in mango (Schnell et al. 1995). In that research, a group of 'Haden' seedlings and a random group of seedlings were evaluated using 11 RAPD primers. This study supported the 'Haden' parentage of 'Eldon', 'Lippens', 'Tommy-Atkins', and 'Zill'; however, the parentage of 'Glenn' and 'Osteen' was questioned. Adato et al. (1995) used DNA fingerprinting (DFP) to evaluate genetic relationships between 26 mango cultivars and 14 rootstocks. They provided a pedigree that further confirmed the relationship between many of the 'Haden' seedlings. Lopez-Valenzuela et al. (1997) used RAPD markers to estimate genetic diversity among 15 rootstock cultivars using 13 markers and identified a specific RAPD band associated only with the polyembryonic types. Eiadthong et al.

(1999) utilized anchored simple sequence repeat markers to analyze 22 mango cultivars; they were able to distinguish genotypes. However, the authors were unable to find markers unique to either monoembryonic or polyembryonic types or for the Thai cultivars selected for green harvest (crispy mango) from the cultivars selected for ripe fruit production. Pandit et al. (2007) also used inter-simple sequence repeat (ISSR) markers to evaluate 60 elite Indian mango cultivars and 10 non-Indian cultivars. They were not able to distinguish Indian cultivars from northern and southern India, and they concluded that ISSR markers could not be considered a comprehensive marker system for mango. Kashkush et al. (2001) utilized amplified fragment length polymorphisms (AFLP) to estimate genetic relationships between 16 cultivars and 7 rootstock cultivars. They also analyzed 29 progeny from a cross of 'Tommy-Atkins' and 'Keitt' and produced a crude linkage map that identified 13 of the 20 linkage groups. The start codon targeted (SCoT) markers have also proved useful in evaluating the relationship among the Xiang Ya Mango type cultivars in China (Luo et al. 2011). SCoT markers are useful for the confirmation of identical genotypes and clones with different names. SCoT and other dominant markers are useful for germplasm collection management, and they are technically simple and inexpensive to implement.

SSR Markers

Viruel et al. (2005) developed the first reported set of 16 SSR markers for mango, of which 14 produced the expected one or two amplification products per genotype. These 14 SSRs were used to evaluate 28 mango genotypes that included 14 Florida cultivars. Discrimination of all 28 genotypes was possible, and the average number of alleles per locus was 5.3. Previously known pedigree information for the 'Haden' family of mangos was confirmed and was in agreement with published RAPD and DFP analyses (Adato et al. 1995; Schnell et al. 1995) with one exception. Viruel's clone of 'Zill' was not resolved as a seedling of 'Haden'. Schnell et al. (2005) developed a second set of 15 SSR markers and analyzed 59 Florida cultivars and four related species. Two of the SSRs were monomorphic among the Florida cultivars; the other 13 had an average number of alleles per locus of 4.2 with polymorphism information content (PIC) values varying from 0.21 to 0.63.

Schnell et al. (2006) used 25 SSR loci to estimate genetic diversity among 203 unique mangos (*M. indica*), two *M. griffithii* Hook. f. & Thomson, and three *M. odorata* Griff. accessions maintained at the National Germplasm Repository (NGR) and at the Fairchild Tropical Garden (FTG) in Miami, Florida. The 25 SSR loci had an average of 6.96 alleles per locus and an average PIC value of 0.552 for the *M. indica* population. The total propagation error in the collection (i.e., plants that had been incorrectly labeled or grafted) was estimated to be 6.13%. When compared by origin, the Florida cultivars were more closely related to Indian than to Southeast Asian cultivars. Unbiased gene diversity (H_{nb}) of 0.600 and 0.582 was found for Indian and Southeast Asian cultivars, respectively, and both were higher than H_{nb} among Florida

cultivars (0.538). When compared by horticultural type, H_{nb} was higher among the polyembryonic types (0.596) than in the monoembryonic types (0.571).

Until recently, a total of 62 SSR markers had been developed for mango (Duval et al. 2005; Honsho et al. 2005; Schnell et al. 2005; Viruel et al. 2005). This number is more than adequate for genetic diversity studies and for parentage analysis as has been demonstrated by Schnell et al. (2006); however, these numbers are not sufficient to develop a saturated linkage map for the 20 linkage groups of mango. Over 300 new SSR markers have been developed at the USDA-ARS Mid South Area Genomics Laboratory (MSAGL) using pyrosequencing and verified using a mango diversity panel of 11 cultivars, using the method described earlier in this chapter. Another additional 1,800 SSR markers have been produced in the Australian mango genomics program (Ian Bally, personal communication 2010).

Parentage Analysis

The SSR markers that have been developed for mango are easily used to verify parentage using a software package like CERVUS (Marshall et al. 1998). When caging trees to exclude foreign pollen or using the polycross mating design, it is now simple to identify the male parent from a set of potential male parents. This has been very useful in cacao breeding where mistakes in pollination have led to the estimation of unreliable breeding values for parental clones.

Among the 64 Florida cultivars evaluated in the parentage analysis by Schnell et al. (2006), the genetic background was found to be based on as few as four Indian cultivars and the polyembryonic cultivar ‘Turpentine’. Two Indian cultivars, ‘Mulgoba’ and ‘Sandersha’, are in the background of most Florida types with ‘Amini’, ‘Bombay’, ‘Cambodiana’, ‘Long’, ‘Julie’, ‘Turpentine’, and ‘Nam doc Mai’ making lesser contributions. The seedling races of Cuba and Florida were considered the same by Popenoe (1920) who called them the West Indian race commonly known as ‘Turpentine’ in Florida. In the parentage analysis, ‘Turpentine 10’ was identified as a most probable paternal parent for ‘Haden’. ‘Haden’ was reported as the maternal parent for 10 cultivars included in the analysis, but based on the parentage analysis, 31 cultivars were found to have ‘Haden’ as one of the most likely parents. Likewise, the other important early Florida selection ‘Brooks’ is the parent of seven cultivars. ‘Haden’, ‘Brooks’, and seedlings of ‘Haden’ and ‘Brooks’ have contributed disproportionately to the genetic background for many of the cultivars in the Florida group.

Linkage Mapping

The first genetic linkage map in mango was reported by Kashkush et al. (2001), utilizing AFLP markers and 29 progeny from a cross of ‘Tommy-Atkins’ × ‘Keitt’ in Israel. They were able to map 34 AFLP loci and produced a crude linkage map that

identified 13 of the 20 linkage groups covering 160 cM. A second map has been produced using 60 progeny from a cross of 'Keitt' × 'Tommy-Atkins' in China using AFLP markers. Eighty-one markers with the correct segregation ratios were identified, and 39 of these were used to identify 15 linkage groups. The average distance between two adjacent markers was 14.74 cM. Improvement of the mango recombination map requires the development of more codominant molecular markers. Using Roche 454 sequencing and the SSR discovery pipeline discussed earlier in this chapter, new SSR markers have been developed and verified and are now being validated at SHRS on mapping populations. A suitable number of SSR markers are being identified to develop a moderately saturated recombination map for mango.

Florida Mapping Populations

Two experimental populations have been developed and planted in the field at SHRS as mapping populations. The first population is an F_2 population derived from self-pollination of 'Tommy-Atkins' consisting of 168 seedlings that were planted in the field in 1995. The second population is an F_2 population derived from self-pollination of 'Haden'. A total of 224 seedlings from a single isolated 'Haden' tree have been in the field for 3 years. Phenotypic data collection is in progress for both of these populations. The development of a saturated linkage map and the identification of QTL for important traits are objectives for the USDA-ARS program at SHRS in Miami for the next 5 years.

Australian Mapping Populations

In the Australian program, supported by Agri-Science Queensland, three mapping populations have been developed using controlled pollinations (Ian Bally, Queensland DEEDI, personal communication). The largest populations are 'Irwin' × 'Kensington Pride', 'Tommy-Atkins' × 'Kensington Pride', and 'Creeper' × 'Kensington Pride'. 'Kensington Pride' is the primary commercial mango cultivar in Australia and has a fine flavor and good productivity. A consensus linkage map has been produced using 700 AFLP and 10 SSR markers. Phenotypic data is currently being collected on the mapping populations for QTL analysis, and additional SSR markers are being added to the maps. These maps have not been published.

Marker-Assisted Selection

To develop a marker-assisted selection (MAS) program for mango, more extensive linkage maps need to be developed and mapping populations field evaluated. The progeny size of both of the Florida mapping populations is small (168 and 224). In

the Australian program, sizable populations have been developed from controlled pollination. These populations are the best candidates for QTL identification in mango. The populations developed from isolated groves by the USDA-ARS SHRS in Florida together with the Australian populations will be used in the next few years to produce a comprehensive linkage map and to identify QTL for disease resistance and for important horticultural traits.

Holton (2010) reported that the Queensland Primary Industries and Fisheries has invested in a gene discovery project for mango with the goal of discovering genes controlling consumer and grower traits, such as fruit quality and tree architecture. Using a multidisciplinary approach, they are sequencing expressed genes via ESTs and serial analysis of gene expression (SAGE), using next-generation sequencing producing low-pass genome converge, identifying candidate genes from fruit quality and tree architecture, and identifying aroma volatiles from fruit.

Lychee

The lychee, *Litchi chinensis*, is the sole member of the genus *Litchi* in the soapberry family Sapindaceae. It is a tropical fruit tree native to southern China, which is also the largest lychee-producing country and the center of diversity for this genus. It is a diploid with $x=15$ and $2n=30$. A number of manuscripts have been published using RAPD and AFLP markers to estimate genetic diversity in lychee including, Tongpamanak et al. (2002) working with Thai cultivars, Kumar et al. (2006) with Indian cultivars, and Jones et al. (2006) working with the USDA-ARS germplasm collection at the Pacific Area Basin Research Center (PBARC) in Hilo, Hawaii. Substantial genetic diversity has been found among cultivars and a high level of heterozygosity detected in these studies. A limited number of SSR markers were developed for lychee (Viruel and Hormaza 2004; Li et al. 2006) and used to evaluate germplasm collections. In contrast to the RAPD and AFLP marker studies, Viruel and Hormaza (2004) analyzed 21 lychee cultivars using SSR markers, and their results suggested a very narrow genetic base. Li et al. (2006) also found limited genetic diversity when evaluating 58 lychee cultivars from Hainan Island using SSR markers. Using the SSR-enrichment method of Techen et al. (2010) in combination with pyrosequencing, 384 markers were developed and are being verified for this species.

Linkage Map

A high-density recombination map was constructed using AFLP and RAPD molecular markers using 73 individuals from a cross of ‘Maguili’ × ‘Jiaohesanyuehong’ (Liu et al. 2010). These individuals are highly divergent, as ‘Maguili’ is a very late ripening cultivar, while the male ‘Jiaohesanyuehong’ is a very early ripening cultivar.

Using AFLP and RAPD markers that segregated in the double pseudo-testcross configuration, 425 maternal specific markers, 357 paternal specific markers, and 257 that were segregating in both parents were identified. Using JoinMap 3.0, a map was developed with 312 markers, covering a total genetic distance of 1,040 cM and identifying 16 linkage groups. Among the linkage groups, MS1 and MS2 had a large number of markers relative to the other linkage groups, 102 and 76, respectively. The remaining 14 linkage groups contained between 12 and 4 markers (Liu et al. 2010). The progeny of this cross were also evaluated for the length of the juvenile phase and for fruit maturation date and fruit growth duration. Twenty markers closely related to fruit maturation were discovered; however, QTLs were not identified for any of the traits (Fu et al. 2008).

Longan

Longan, *Dimocarpus longan*, is native to southern China in Yunnan province and is an important tree fruit crop in the Sapindaceae family. It is a diploid with $x = 15$ and $2n = 30$. The fruit has a bright black seed, and the literal meaning of the word longan is "eye of the dragon." Isozymes were first used to identify longan cultivars in the 1980s (Chen and Ke 1989); however, this method was not very useful in distinguishing cultivars. RAPDs were the first PCR-based marker developed for longan, allowing the separation of different groups and the identification of cultivars (Chen and Liu 2001; Lin et al. 2005; Zhong et al. 2007). Yonemoto et al. (2006), using RAPD markers, was able to distinguish three different groups of germplasm from Taiwan, China, and Malaysia. AFLPs were used by Yi et al. (2003), and they were able to distinguish 11 clusters among the 46 cultivars studied. ISSR markers (Hong 2007) have also been developed and used for cultivar identification and diversity analysis. Using the SSR-enrichment method of Techen et al. (2010) in combination with pyrosequencing, 384 markers were developed and are being verified for this species.

Linkage Map

Guo et al. (2010) constructed a linkage map using 94 progeny from a cross of 'Fengliduo' and 'Dawuyuan'. 'Fengliduo' is a high-quality cultivar, while 'Dawuyuan' has a large fruit size. Using RAPD, ISSR, sequence-related amplified polymorphism (SRAP), and AFLP markers in and a pseudo-testcross design, JoinMap 3.0 was used to construct the linkage maps. For the cultivar 'Fengliduo', 183 loci were mapped and 21 linkage groups were identified covering 965 cM. For the cultivar 'Dawuyuan', 251 markers were mapped into 22 linkage groups covering 1,064 cM. This is the first reported map on longan which needs to be enhanced with codominant markers.

Summary

The ease of development of molecular markers, especially those amenable to high throughput in a MAS program will greatly enhance the rate of genetic gain for many of these tropical fruit species. The development of second- and third-generation pyrosequencing combined with the SSR and SNP discovery pipelines has removed many of the constraints on the use of molecular markers in these tropical fruit species. Thousands of SSR markers and hundreds of thousands of SNP markers can now be discovered with just a few Roche 454 or Illumina runs. With the removal of the limitation on numbers of markers, much of the difficulty in applying these techniques has been greatly reduced. The more challenging problem is the development of large mapping populations and the accumulation of phenotypic data to identify QTL regions associated with traits of interest. Once these populations are made and evaluated, the application of MAS to tropical fruit crops is straightforward.

Complete genome sequencing for most of these species is still a distant possibility. *De novo* assembly of second-generation sequence data into pseudo-molecules representing complete chromosomes requires a significant investment of time and funds. Considering the limited commercial value of many of these tropical species, complete genome sequencing and assembly are not likely using current technology. However, the ability to generate thousands of markers and to apply MAS will greatly accelerate tropical fruit breeding programs.

References

Accerbi M, Schmidt SA, De Paoli E, Park S, Jeong D-H, Green PJ (2010) Methods for isolation of total RNA to recover miRNAs and other small RNAs from diverse species. In: Meyers BC, Green PJ (eds) Plant microRNAs. Methods in molecular biology, vol 592. Humana Press. doi: 10.1007/978-1-60327-005-2_3

Adato A, Sharon D, Lavi U (1995) Application of DNA fingerprints for identification and genetic analyses of mango (*Mangifera indica*) genotypes. J Am Soc Hortic Sci 120:259–264

Albert VA, Soltis DE, Carlson JE, Farmerie WG, Wall PK, Ilut DC, Solow TM, Mueller LA, Landherr LL, Hu Y, Buzgo M, Kim S, Yoo M-J, Frohlich MW, Perl-Treves R, Schlarbaum SE, Bliss BJ, Zhang X, Tanksley SD, Oppenheimer DG, Soltis PS, Ma H, dePamphilis CW, Leebens-Mack JH (2005) Floral gene resources from basal angiosperms for comparative genomics research. BMC Plant Biol 5(1):5

Anonymous (2002) FAOSTAT online database at http://www.fao.org/

Archak S, Gaikwad AB, Gautam D, Rao EV, Swamy KR, Karihaloo JL (2003) Comparative assessment of DNA fingerprinting techniques (RAPD, ISSR and AFLP) for genetic analysis of cashew (*Anacardium occidentale* L.) accessions of India. Genome 46:362–369

Archak S, Gaikwad AB, Swamy KR, Karihaloo JL (2009) Genetic analysis and historical perspective of cashew (*Anacardium occidentale* L.) introduction into India. Genome 52:222–230

Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzey J, Allegre M, Chaparro C, Legavre T, Maxomova SN, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Fernandes Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju JSS, Schuster SC, Carlson JE, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Berard A, Viot C, Boccara M, Risterucci AM, Guignon V, Sabau X, Axtell MJ, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza JM, Pitollat B, Gramacho K, D'Hont A, Brunel D,

Infante D, Kebe I, Costet P, Wing R, McCombie WR, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C (2010) The genome of *Theobroma cacao*. Nat Genet. doi:10.1038/ng 736

Arias RS, Ballard LL, Duke MV, Simpson SA, Liu XF, Irish BM, Scheffler BE (2010) Batch isolation of microsatellites from tropical species using pyrosequencing. In: Methods in molecular biology (in press)

Arumuganathan K, Earle ED (1991a) Estimation of nuclear DNA content of plants by flow cytometry. Plant Mol Biol Rep 9:229–233

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

Ashworth VETM, Clegg MT (2003) Microsatellite markers in avocado (*Persea americana* Mill.): genealogical relationships among cultivated avocado genotypes. J Hered 94:407–415

Ashworth VETM, Kobayashi MC, De La Cruz M, Clegg MT (2004) Microsatellite markers in avocado (*Persea americana* Mill.): development of dinucleotide and trinucleotide markers. Scientia Hortic 101:255–267

Ashworth VETM, Chen H, Clegg MT (2007) 17 Avocado. In: Kole C (ed) Genome mapping and molecular breeding in plants, fruits and nuts, vol 4. Springer, Berlin, pp 325–329

Assogbadjo AE, Kyndt T, Sinsin B, Gheysen G, van Damme P (2006) Patterns of genetic and morphometric diversity in baobab (*Adansonia digitata*) populations across different climatic zones of Benin (West Africa). Ann Bot 97:819–830

Azevedo VCR, Vinson CC, Ciampi AY (2005) Twelve microsatellite loci in *Manilkara huberi* (Ducke) Standl (Sapotaceae), an Amazonian timber species. Mol Ecol Notes 5:13–15

Azofeifa-Delgado A (2006) Uso de marcadores moleculares en plantas; aplicaciones en frutales del tropico. Agronomia Mesoamericana 17:221–242

Barkley NA, Roose ML, Krueger RR, Federici CT (2006) Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). Theor Appl Genet 112:1519–1531

Bergh BO (1969) Avocado (*Persea americana* Mill.) In: Ferwerda FP, Wit F (eds) Outlines of perennial crop breeding in the tropics. Misc. Paper 4. Landbouwhogeschool, Wageningen, pp 23–51

Blanche KR, Ludwig JA, Cunningham SA (2006) Proximity rainforest enhances pollination and fruit set in orchards. J Appl Ecol 43:1182–1187

Blas AL, Yu Q, Chen C, Veatch O, Moore PH et al (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome 52:716–725

Bompard JM, Schnell RJ (1997) Taxomomy and systematics. In: Litz RE (ed) The mango, botany, production and uses. CABI International, New York, pp 21–48

Borrone JW, Schnell RJ, Violi HA, Ploetz RC (2007) Seventy microsatellite markers from *Persea americana* Miller (avocado) expressed sequence tags. Mol Ecol Notes 7(3):439–444

Borrone JW, Olano CT, Kuhn DN, Brown JS, Schnell RJ, Violi HA (2008) Outcrossing in Florida avocados as measured using microsatellite markers. J Am Soc Hortic Sci 133:255–261

Borrone J, Brown JS, Tondo CT, Mauro-Herrera M, Kuhn DN, Violi HA, Sautter RT, Schnell RJ (2009) An EST-SSR-based linkage map for *Persea americana* Mill (avocado). Tree Genet Gen 5:553–560. doi:10.1007/s11295-009-0208-y

Brown J, Laurentín H, Dávila M (2003) Genetic relationships between nine *Annona muricata* L. accessions using RAPD markers. Fruits 58:255–259

Caceres A, Fletes L, Aguilar L, Ramirez O, Figueroa L et al (1993) Plants used in Guatemala for the treatment of gastrointestinal disorders. 3. Confirmation of activity against enterobacteria of 16 plants. J Ethnopharmacol 38:31–38

Campbell CW, Malo SE (1978) Review of methods for measuring avocado maturity in Florida. Proc Am Soc Hortic Sci Trop Reg 22:16–26

Chanderbali AS, Albert VA, Ashworth VETM, Clegg MT, Litz RE, Soltis DE, Soltis PS (2008) *Persea americana* (avocado): bringing ancient flowers to fruit in the genomics era. Bioessays 30(4):386–396

Chanderbali AS, Albert VA, Leebens-Mack J, Altman NS, Soltis DE, Soltis PS (2009) Transcriptional signatures of ancient floral developmental genetics in avocado (*Persea americana*; Lauraceae). Proc Natl Acad Sci USA 106(22):8929–8934

Chatrou LW, Escribano MP, Viruel MA, Maas JW, Richardson JE, Hormaza JI (2009) Flanking regions of monomorphic microsatellite loci provide a new source of data for plant species-level phylogenetics. Mol Phylogenet Evol 53:726–733

Chen X, Ke GW (1989) A preliminary study on group classification of Fujian longan varieties examined by the analytical technique of peroxidase isozymes. J China Fruit 1:18–21

Chen YZ, Liu CM (2001) Analysis and identification of cultivated varieties of longan by RAPD molecular markers. J China Fruit 4:28029

Chen H, Ashworth VETM, Xu S, Clegg MT (2007) Quantitative analysis of growth rate in avocado. J Am Soc Hortic Sci 132(5):691–696

Chen H, Morrell PL, de la Cruz M, Clegg MT (2008) Nucleotide diversity and linkage disequilibrium in wild avocado (*Persea americana* Mill.). J Hered 99:382–389

Crowhurst RN, Gleave AP, MacRae EA, Ampomah-Dwamena C, Atkinson RG et al (2008) Analysis of expressed sequence tags from *Actinidia*: applications of a cross species EST database for gene discovery in the areas of flavor, health, color and ripening. BMC Genomics 9:351

Croxford AE, Robson M, Wilkinson MJ (2006) Characterization and PCR multiplexing of polymorphic microsatellite loci in cashew (*Anacardium occidentale* L.) and their cross-species utilization. Mol Ecol Notes 6:249–251

Darlington CD, Wylie AP (1945) Chromosome atlas of flowering plants. George Allen & Unwin Ltd., London, p 16

Davenport TL (1986) Avocado flowering. Hort. Rev. 8:257–289

Davenport TL, Parnitzki P, Fricke S, Hughes MS (1994) Evidence and significance of self-pollination of avocados in Florida. J Am Soc Hortic Sci 119:1200–1207

Dawson IK, Hollingsworth PM, Doyle JJ, Kresovich S, Weber JC et al (2008) Origins and genetic conservation of tropical trees in agroforestry systems: a case study from the Peruvian Amazon. Conserv Genet 9:361–372

de Oliveira RP, Cristofani M, Machado MA (2005) Integrated genetic map of citrus based on RAPD markers. Fruits 60:187–193

Degani C, El-Batsri R, Gazit S (1990) Enzyme polymorphism in mango. J Am Soc Hortic Sci 115:844–847

Devi-Sampath P, Vijayaraghavan K (2007) Cardioprotective effect of alpha-mangostin, a xanthone derivative from mangosteen on tissue defense system against isoproterenol-induced myocardial infarction in rats. J Biochem Mol Toxicol 21:336–339

Diallo BO, Joly HI, McKey D, Hossaert-McKey M, Chevallier MH (2007) Genetic diversity of *Tamarindus indica* populations: any clues on the origin from its current distribution? Afr J Biotechnol 6:853–860

Diaz-Matallana M, Schuler-Garcia I, Ruiz-Garcia M, de Jaramillo EH (2009) Analysis of diversity among six populations of Colombian mango (*Mangifera indica* L. cvar. Hilacha) using RAPDs markers. Elect J Biotechnol 12:1–8

Dillon S, Ramage C, Ashmore S, Drew RA (2006) Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. Theor Appl Genet 113:1159–1169

Duval MF, Bunel J, Sitbon C, Risterucci AM (2005) Development of microsatellite markers for mango (*Mangifera indica* L.). Mol Ecol Notes 5:824–826

Duval MF, Bunel J, Sitbon C, Risterucci AM, Calabre C, Le Bellec F (2006) Genetic diversity of Caribbean mangoes (*Mangifera indica L.*) using microsatellite markers. Presented at proceedings of the eighth international mango symposium, Sun City, South Africa

Eiadthong W, Yonemori K, Sugiura A, Utsunomiya N, Subhadrabandhu S (1999) Identification of mango cultivars of Thailand and evaluation of their genetic variation using the amplified fragments by simple sequence repeat-(SSR-) anchored primers. Scientia Hortic 82:57–66

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435–445

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99:125–132

Englberger L, Schierle J, Hofmann P, Lorens A, Albert K et al (2009) Carotenoid and vitamin content of Micronesian atoll foods: Pandanus (*Pandanus tectorius*) and garlic pear (*Crataeva speciosa*) fruit. J Food Compos Anal 22:1–8

Escribano P, Viruel MA, Hormaza JI (2008) Development of 52 new polymorphic SSR markers from cherimoya (*Annona cherimola* Mill.): transferability to related taxa and selection of a reduced set for DNA fingerprinting and diversity studies. Mol Ecol Resour 8:317–321

FAO (2003) Medium-term prospects for agricultural commodities. Projections to the year 2010. Rome, Italy, p 193

Feria-Romero I, Astudillo-de la Vega H, Chavez-Soto MA, Rivera-Arce E, López M et al (2009) RAPD markers associated with quercetin accumulation in *Psidium guajava*. Biol Plant 53:125–128

Fu JX, Guo YS, Zhato HY, Xie JH, Liu CM, Hu YL, Ouyang R, Zhao YH, Zhou J, Huang SS (2008) A study on the genetic effect and MAS exploration of juvenile phase and fruit maturation period in a mapping population of litchi. In: Qui Dongliang (ed) 3rd IS on longan, lychee and other fruit trees in Spindaceae family. Acta Hortic 863:195–198

Geetha G, Banumathi B, Suresh G (1997) Evaluation of the antifungal activity of natural xanthones from *Garcinia mangostana* and their synthetic derivatives. J Nat Prod 60(5):519–524

Gunjate RT, Kumbhar AR, Thimaiah IM, Amin SM (2009). Acta Horticulturae No. 820 pp. 463–467

Guo DL, Luo ZR (2008) Microsatellite isolation and characterization in Japanese persimmon (*Diospyros kaki*). Biochem Genet 46:323–328

Guo YS, Liu CJ, Zhao YH, Xie JH, Ren PR, Huang TL, Fu JX, Lu BB, Liu CM (2010) Construction of a molecular genetic linkage map for longan based on RAPD, ISSR, SRAP and AFLP markers. Acta Hortic 863:141–147

Hawkins DJ, Kridl JC (1998) Characterization of acyl-ACP thioesterases of mangosteen (*Garcinia mangostana*) seed and high levels of stearate production in transgenic canola. Plant J 13:743–752

Heaton HJ, Whitkus R, Gomez-Pompa A (1999) Extreme ecological and phenotypic differences in the tropical tree chicozapote (*Manilkara zapota* (L.) P. Royen) are not matched by genetic divergence: a random amplified polymorphic DNA (RAPD) analysis. Mol Ecol 8:627–632

Hernández-Delgado S, Padilla-Ramírez JS, Nava-Cedillo A, Mayek-Pérez N (2007) Morphological and genetic diversity of Mexican guava germplasm. Plant Genet Res 5:131–141

Hillyer MJ, Boulter SL, Kitching RL, Hughes JM (2007) Isolation and characterization of eight polymorphic microsatellite loci in the rainforest canopy tree, *Syzygium sayeri* (Myrtaceae). Mol Ecol Notes 7:1199–1201

Hirano R, Oo TH, Watanabe KN (2010) Myanmar mango landraces reveal genetic uniqueness over common cultivars from Florida, India and Southeast Asia. Genome 53:321–330

Hollingsworth PM, Dawson IK, Goodall-Copestake WP, Richardson JE, Weber JC et al (2005) Do farmers reduce genetic diversity when they domesticate tropical trees? A case study from Amazonia. Mol Ecol 14:497–501

Holton T (2010) Mango genomics. Abs PAG XVIII, Jan 2010, San Diego

Hong ZT (2007) Inter-simple sequence repeat (SSR) analysis of different cultivars and seed-abortive character in longan (*Dimocarpus longan* Lour) D. Fuzhou: Fujian Agr For Univ 56

Honsho C, Nishiyama K et al (2005) Isolation and characterization of new microsatellite markers in mango (*Mangifera indica*). Mol Ecol Notes 5:152–154

Iyer CPA, Degani C (1997) Classical breeding and genetics. In: Litz RE (ed) The mango, botany, production and uses. CAB International, New York, pp 49–68

Iyer CPA, Schnell RJ (2009) Classical breeding and genetics. In: Litz RE (ed) The mango, botany, production and uses, vol 2. CAB International, New York, pp 67–96, Chapter 4

Jamnadass R, Lowe A, Dawson IK (2009) Molecular markers and the management of tropical trees: the case of indigenous fruits. Trop Plant Biol 2:1–12

Jones M, Zee FT, Moore PH, Kim MS, Ming R (2006) Genetic diversity of litchi germplasm assessed by AFLP marker. Plant Animal and Microbe Genomes Conference XIV P489, p. 223

Kashkush K, Jinggui F, Tomer E, Hillel J, Lavi U (2001) Cultivar identification and genetic map of mango (*Mangifera indica)*. Euphytica 122:129–136

Kijas JM, Fowler JC, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16:656–660

Knight RJ (2002) History distribution and uses. In: Whiley AW, Schaffer B, Wolstenholme BN (eds) The avocado: botany production and uses. CAB International, Wallingford

Knight RJ, Schnell RJ (1994) Mango introduction in Florida and the Haden cultivar's significance to the modern industry. Econ Bot 48(2):139–145

Kuete V, Wabo GF, Ngameni B, Mbaveng AT, Metuno R et al (2007) Antimicrobial activity of the methanolic extract, fractions and compounds from the stem bark of *Irvingia gabonensis* (Ixonanthaceae). J Ethnopharmacol 114:54–60

Kumar M, Gupta M, Shrivastava D, Prasad M, Passar US, Sarin NB (2006) Genetic relatedness among Indian litchi accessions (*Litchi chinensis* Sonn.) by RAPD markers. Int J Agri Res 1:390–400

Kyndt T, Assogbadjo AE, Hardy OJ, Glele Kakai R, Sinsin B et al (2009) Spatial genetic structuring of baobab (*Adansonia digitata*, Malvaceae) in the traditional agroforestry systems of West Africa. Am J Bot 96:950–957

Lahav E, Lavi U (2002) Genetics and classical breeding. In: Whiley A, Schaffer B, Wolstenholome B (eds) The avocado, botany, production, and uses. CAB International, New York, pp 39–69

Larsen A, Vaillant A, Verhaegen D, Kjær E (2009) Eighteen SSR-primers for tetraploid *Adansonia digitata* and its relatives. Conserv Genet Resour 1:325–328

Li MF, Zheng XQ, Zhu YQ, Wang XS, Liang SY et al (2006) Development and characterization of SSR markers in lychee (*Litchi chinensis*). Mol Ecol Notes 6:1205–1207

Lin T, Lin Y, Ishiki K (2005) Genetic diversity of *Dimocarpus longan* in China revealed by AFLP markers and partial rbcL gene sequences. Sci Hortic 103:489–498

Liu G, Bode A, Ma WY, Sang S, Ho CT, Dong Z (2001) Two novel glycosides from the fruits of *Morinda citrifolia* (noni) inhibit AP-1 transactivation and cell transformation in the mouse epidermal JB6 cell line. Cancer Res 61:5749–5756

Liu CM, Guo YS, Liu R, Zhang B, Xie JH, Hu YL, Ouyang R, Zhao YH, Fu JC (2010) Construction of a high density molecular linkage map for lychee based on AFLP and RAPD markers. Acta Hortic 863:87–94

Livingstone DS, Motamayor JC, Schnell RJ, Cariaga K, Freeman B, Meerow AW, Brown JS, Kuhn DN (2010) Mol Breed. doi:10.1007/s11032-010-9416-2

López-Gómez R, Torres-Cárdenas Y, Chávez-Moctezuma M, Salgado-Garciglia R, Jiménez-Moraila B, Corona-Armenta G, Herrera-Estrella L (2007) Genomics of Avocado Criollo Fruit (*Persea americana* var. *drymifolia*). In: Proceedings VI World Avocado Congress (Actas VI Congreso Mundial del Aguacate) 2007. Viña Del Mar, Chile, 12–16 Nov 2007

Lopez-Valenzuela JA, Martinez O, Paredes-Lopez O (1997) Geographic differentiation and embryo type identification in *Mangifera indica* L. cultivars using RAPD markers. Hortscience 32:1105–1108

Lowe AJ, Russell JR, Powell W, Dawson IK (1998) Identification and characterization of nuclear, cleaved amplified polymorphic sequence (CAPS) loci in *Irvingia gabonensis* and *I. wombolu*, indigenous fruit trees of west and central Africa. Mol Ecol 7:1786–1788

Lowe AJ, Gillies AC, Wilson J, Dawson IK (2000) Conservation genetics of bush mango from central/west Africa: implications from random amplified polymorphic DNA analysis. Mol Ecol 9:831–841

Luo C, He XH, Chen H, Ou SJ, Gao MP, Brown JS, Tondo CT, Schnell RJ (2011) Genetic diversity of mango cultivars estimated using SCot and ISSR markers. Current Science (in press)

Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655

Martinez-Godoy MA, Mauri N, Juarez J, Marques MC, Santiago J et al (2008) A genome-wide 20 K citrus microarray for gene expression analysis. BMC Genomics 9:318

Meghala R, Ravishankar KV, Anand L, Rekha A (2005) Genetic diversity of Indian sapota (*Manilkara zapota*) cultivars characterized by RAPD markers. PGR Newsletter 142:43–46: FAO-Bioversity

Miller A, Schaal B (2005) Domestication of a Mesoamerican cultivated fruit tree, *Spondias purpurea*. Proc Natl Acad Sci USA 102:12801–12806

Miller AJ, Schaal BA (2006) Domestication and the distribution of genetic variation in wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* L. (Anacardiaceae). Mol Ecol 15:1467–1480

Ming R, Hou SB, Feng Y, Yu QY, Dionne-Laporte A et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–U997

Morton J (1987) Lychee. In: Fruits of warm climates. Miami, pp 249–259. http://www.hort.purdue.edu/newcrop/morton/lychee.html

Muchugi A, Kadu C, Kindt R, Kipruto H, Lemurt S et al (2008) Molecular markers for tropical trees. A practical guide to principles and procedures. World Agroforestry Centre, Nairobi

Mukherjee SK (1950) Mango: its allopolyploid nature. Nature 166:196–197

Mukherjee SK (1953) The mango – its botany, cultivation, uses and future improvement, especially as observed in India. Econ. Bot. 7:130–62

Mukherjee SK (1957) Cytology of some Malayan species of *Mangifera*. Cytologia 22:239–241

Mukherjee SK (1963) Cytology and breeding of mango. Punjab Hortic J 3:107–115 (observed in India. *Economic Botany* 7)

Nakasone HY, Paull RE (1998) Avocado. In: Tropical fruits. CAB International, New York, pp 76–102

Ng KW, Salhimi SM, Majid AM, Chan KL (2010) Anti-angiogenic and cytotoxicity studies of some medicinal plants. Planta Med 76:935–940

Novelli VM, Takita MA, Machado MA (2004) Identification and analysis of single nucleotide polymorphisms (SNPs) in citrus. Euphytica 138:227–237

Nyffeler R, Baum DA (2000) Phylogenetic relationships of the durians (Bombacaceae-Durioneae or/Malvaceae/Helicteroideae/Durioneae) based on chloroplast and nuclear ribosomal DNA sequences. Plant Systemat Evol 224:55–82

Obolskiy D, Pischel I, Siriwantanametanon N, Heinrich M (2009) *Garcinia mangostana* L. a phytochemical and pharmacological review. Phytother Res 23:1047–1065

Ocampo J, dEeckenbrugge GC, Bruyère S, de Lapeyre de Bellaire L, Ollitrault P (2006) Organization of morphological and genetic diversity of Caribbean and Venezuelan papaya germplasm. Fruits 61:25–37

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC (2006) Origin, evolution and genome distribution of microsatellites. Genet Mol Biol 29:294–307

Oliveira DD, Lobato AL, Ribeiro SM, Santana AM, Chaves JB, Pinheiro-Sant'ana HM (2010) Carotenoids and vitamin C during handling and distribution of guava (*Psidium guajava* L.), mango (*Mangifera indica* L.), and papaya (*Carica papaya* L.) at commercial restaurants. J Agric Food Chem 2010:4

Palu AK, Kim AH, West BJ, Deng S, Jensen J, White L (2008) The effects of *Morinda citrifolia* L. (noni) on the immune system: its molecular mechanisms of action. J Ethnopharmacol 115:502–506

Pandit SS, Mitra S, Giri AP, Pujari KH, Bhimarao PP, Jambhale ND, Gupta VS (2007) Genetic diversity analysis of mango cultivars using inter simple sequence repeat markers. Curr Sci 93(8):1135–1141

Perfectti F, Pascual L (1996) Segregation distortion of isozyme loci in cherimoya (*Annona cherimola* Mill). Theor Appl Genet 93:440–446

Perfectti F, Pascual L (2005) Genetic diversity in a worldwide collection of cherimoya cultivars. Genet Resour Crop Evol 52:959–966

Pinto ACQ, Andrade SRM, Ramos VHV, Cordeiro MCR (2004) Intervarietal hybridization in mango (*Mangifera indica* L.): techniques, main results and their limitations. Acta Hortic (ISHS) 645:327

Pistelli L, Bertoli A, Noccioli C, Mendez J, Musmanno RA et al (2009) Antimicrobial activity of *Inga fendleriana* extracts and isolated flavonoids. Nat Prod Commun 4:1679–1683

Ploetz R, Schnell RJ, Haynes J (2002) Variable response of open-pollinated seedlings progeny of avocado to Phytophthora root rot. Phytoparasitica 30(3):262–268

Poerwanto R (2002) Nurse stock plant – a new technique to enhance mangosteen (*Garcinia mangostana*) growth. Acta Hortic 575:751–756

Popenoe W (1920) Manual of tropical and subtropical fruits. Macmillan, New York

Ragasa CY, Puno MR, Sengson JM, Shen CC, Rideout JA, Raga DD (2009) Bioactive triterpenes from *Diospyros blancoi*. Nat Prod Res 23:1252–1258

Rai M, Asthana P, Jaiswal V, Jaiswal U (2010) Biotechnological advances in guava (*Psidium guajava* L.): recent developments and prospects for further research. Trees – Struct Func 24:1–12
Ramage CM, Sando L, Peace CP, Carroll BJ, Drew RA (2004) Genetic diversity revealed in the apomictic fruit species *Garcinia mangostana* L. (mangosteen). Euphytica 136:1–10
Razafimandimbison SG, McDowell TD, Halford DA, Bremer B (2010) Origin of the pantropical and nutriceutical *Morinda citrifolia* L. (Rubiaceae): comments on its distribution range and circumscription. J Biogeogr 37:520–529
Richard GF, Kerrest A, Dujon B (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol Mol Biol Rev 72:686–727
Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293:2242–2245
Ripa FA, Haque M, Bulbul IJ (2010) In vitro antibacterial, cytotoxic and antioxidant activities of plant *Nephelium longan*. Pak J Biol Sci 13:22–27
Risterucci AM, Duval MF, Rohde W, Billotte N (2005) Isolation and characterization of microsatellite loci from *Psidium guajava* L. Mol Ecol Notes 5:745–748
Samuel R, Pinsker W, Balasubramaniam S, Morawetz W (1991) Allozyme diversity and systematics in Annonaceae. Plant Systemat Evol 178:125–134
Sand L, Peach C, Ramage C, Carroll BJ, Drew R (2005) Assessment of genetic diversity in Australian-grown mangosteen (*Garcinia mangostana* L.) and its wild relatives. Acta Hortic 692:143–151
Santoso PJ, Saleh GB, Saleh NM, Napis S (2005) Phylogenetic relationships amongst 10 Durio species based on PCR-RFLP analysis of two chloroplast genes. Indones J Agric Sci 6:20–27
Schnell RJ, Knight RJ, Harkins DM, Zill G (1994) Eliminating zygotic seedlings in 'Turpentine' mango rootstock populations by visual roguing. Hortscience 29(4):319–320
Schnell RJ, Ronning CM, Knight RJ (1995) Identification of cultivars and validation of genetic relationships in *Mangifera indica* L. using RAPD markers. Theor Appl Genet 90:269–271
Schnell RJ, Olano CT, Campbell RJ, Brown JS (2001) AFLP analysis of genetic diversity within a jackfruit germplasm collection. Sci Hortic 91:261–274
Schnell RJ, Brown JS, Olano CT, Power EJ, Krol CA, Kuhn DN, Motamayor JC (2003) Evaluation of avocado germplasm using microsatellite markers. J Am Soc Hortic Sci 128(6):881–889
Schnell RJ, Olano CT, Quintanilla WE, Meerow AW (2005) Isolation and characterization of 15 microsatellite loci from mango (*Mangifera indica* L.) and cross-species amplification in closely related taxa. Mol Ecol Notes 5:625–627
Schnell RJ, Brown JS, Olano CT, Meerow AW, Campbell RJ, Kuhn DN (2006) Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite markers. J Am Soc Hortic Sci 131(2):214–224
Schnell RJ, Tondo CL, Brown JS, Kuhn DN, Ayala-Silva T, Borrone JW, Davenport TL (2009) Outcrossing rates between 'Bacon' pollinators and adjacent 'Hass' trees in a commercial California avocado grove estimated using microsatellite markers and the identification of two new lethal mutants. Hortscience 44(6):1522–1526
Scora RW, Bergh BO (1992) Origin of and taxonomic relationship within the genus *Persea*. In: Proceedings of the 2nd World Avocado Congress, vol 2, Riverside, pp 505–514
Selmer KK, Brandal K, Olstad OK, Birkenes B, Undlien DE, Egeland T (2009) Genome-wide linkage analysis with clustered SNP markers. J Biomol Screen 14:92–96
Shapcott A (1999) Vagility and monsoon rainforest archipelago of Northern Australia: patterns of genetic diversity in *Syzygium nervosum* (Myrtaceae). Biotropica 31:579–590
Sharon D, Cregan PB, Mhameed S, Kusharska K, Hillel J, Lahav E, Lavi U (1997) An integrated genetic linkage map of avocado. Theor Appl Genet 95:911–921
Shyamalamma S, Chandra SBC, Hedge M, Naryanswamy P (2008) Evaluation of genetic diversity in Jackfruit (*Artocarpus heterophyllus* Lam.) based on amplified fragment length polymorphism markers. Genet Mol Res 7:645–656
Sim CH, Mahani MC, Choong CY, Salma I (2005) Transferability of SSR markers from lychee (*Litchi chinensis* Sonn.) to pulasan (*Nephelium ramboutan-ake* L.). Fruits 60:379–385

Singh RN, Sharma DK, Majumder PK (1980) An efficient technique of mango hybridization. Sci Hortic 12:299–301

Sobir PR (2007) Mangosteen genetics and improvement. Int J Plant Breed 1:105–111

Sreekumar V, Binoy A, George S (2007) Genetic and morphological variation in breadfruit (*Artocarpus altilis* (Park.) Fosberg) in the Western Ghats of India using AFLP markers. Genet Resour Crop Evol 54:1659–1665

Stern RA, Gazit S (2003) The reproductive biology of the lychee. Hortic Rev 28:393–453

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205–233

Techen N, Arias RS, Glynn NC, Pan Z, Khan I, Scheffler BE (2010) Optimized construction of microsatellite-enriched libraries. Mol Ecol Resour 10:508–515

Tongpamanak P, Asalaha P, Peerasak S, Chalongchai B (2002) Determination of genetic diversity and relationships among Thai litchi accessions by RAPD and AFLP markers. Kasetsart J (Nat Sci) 36:370–380

Tsy JMLP, Lumaret R, Mayne D, Vall AOM, Abutaba YIM et al (2009) Chloroplast DNA phylogeography suggests a West African centre of origin for the baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). Mol Ecol 18:1707–1715

Ude GN, Dimkpa CO, Anegbeh PO, Shaibu AA, Tenkouano A et al (2006) Analysis of genetic diversity in accessions of *Irvingia gabonensis* (Aubry-Lecomte ex O'Rorke) Baill. Afr J Biotechnol 5:219–223

Ushanandini S, Nagaraju S, Kumar KH, Vedavathi M, Machiah DK et al (2006) The anti-snake venom properties of *Tamarindus indica* (leguminosae) seed extract. Phytother Res 20:851–858

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23(1):48–55

Viruel MA, Hormaza JI (2004) Development, characterization and variability analysis of microsatellites in lychee (*Litchi chinensis* Sonn., Sapindaceae). Theor Appl Genet 108:896–902

Viruel MA, Escribano P, Barbieri M, Ferri M, Hormaza JI (2005) Fingerprinting, embryo type and geographic differentiation in mango (*Mangifera indica* L., Anacardiaceae) with microsatellites. Mol Breed 15:383–393

Viruel MA, Gross-Germann E, Barceló A (2007) Desarrollo de un mapa genético con marcadores SSRs y AFLPs en aguacate. VI World Avocado Congress, Viña del Mar, Chile, 12–16 Nov 2007

Wall PK, Leebens-Mack J, Muller K, Field D, Altman N, dePamphilis CW (2008) PlantTribes: a gene and gene family resource for comparative genomics in plants. Nucleic Acids Res 36:D970–D976, http://ancangio.uga.edu/content/persea-americana

Yi GJ, Tan WP, Hur HQ (2003) Studies on the genetic diversity and relationship of longan cultivars by AFLP analysis. Acta Hortic Sin 30:272–276

Ying ZT, Davenport TL, Zhang TF, Schnell RJ, Tondo CL (2009) Selection of highly informative microsatellite markers to identify pollen donors in 'Hass' avocado orchards. Plant Mol Biol Rep 27:374–380

Yonemoto Y, Chowdhury AK, Kato H, Macha MM (2006) Cultivars identification and their genetic relationships in *Dimocarpus longan* subspecies based on RAPD markers. Sci Hortic 109:147–152

Yuan SS, Chang HL, Chen HW, Yeh YT, Kao YH et al (2003) Annonacin, a mono-tetrahydrofuran acetogenin, arrests cancer cells at the G1 phase and causes cytotoxicity in a Bax- and caspase-3-related pathway. Life Sci 72:2853–2861

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1–16

Zerega NJC, Ragone D, Motley TJ (2004) Complex origins of breadfruit (*Artocarpus altilis*, Moraceae): implications for human migrations in Oceania. Am J Bot 91:760–766

Zheng SQ, Jiang F, Gao HY, Zheng JG (2010) Preliminary observations on the antifatigue effects of longan (*Dimocarpus longan* Lour.) seed polysaccharides. Phytother Res 24:622–624

Zhong FL, Pan DM, Guo Z (2007) RAPD analysis of longan germplasm resources. J Chin Agric Sci Bull 23:558–563