17 Avocado

Vanessa E. T. M. Ashworth, Haofeng Chen, and Michael T. Clegg

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525, USA
e-mail: vashwort@uci.edu

17.1
Introduction

17.1.1
Crop Production

Avocado (*Persea americana* Mill.) is a high-value specialty crop enjoying widespread cultivation in the far corners of the world. Mexico (Michoacán State) is the primary producer, with over 200,000 acres yielding almost 2.1 billion lbs of avocados, followed by the United States (68,000 acres, 531 million lbs; 90% from California) and Chile (65,000 acres, 388 million lbs), and significant contributions (cited in order of decreasing fruit production) also by Peru, South Africa, Israel, Spain, Australia, and New Zealand (figures for 2004/2005 season; *http://www.avocadosource.com*). By far the most widely grown cultivar is 'Hass', a black-skinned form adapted to a Mediterranean climate and combining high yields with excellent flavor. The exemplary nutritive composition of avocado – including lutein, vitamin E, monounsaturated ("good") fat and folate – has also contributed to its growing popularity.

17.1.2
Botanical Races

Three botanical races are often distinguished to circumscribe subgroups combining particular fruit characteristics and ecological adaptations: the Mexican, Guatemalan and West Indian races. The Mexican botanical race includes cultivars with thin, black skin and anise-scented leaves. Cultivars with thick, rough, green skin belong to the Guatemalan botanical race and have a level of cold tolerance somewhat inferior to the Mexican race. Cultivars of the West Indian race – a misnomer as these genotypes appear to have arisen in coastal Guatemala – are cold-sensitive, often relatively salt-tolerant, and have fruit with low oil content. Wild *Persea* relatives bear little resemblance to their cultivated counterparts, but cultigens have been selected by local people throughout Central America, Mexico, and farther afield, which has led to a plethora of forms that are assigned to one of the three botanical races based on the best morphological and ecological "fit". A discontinuity between wild and selected avocado forms suggests that any intermediate genotypes have long disappeared, and that human selection has been in progress for some considerable time. Indeed, the first archaeological evidence from the Mexican state of Puebla dates back to ca. 8000 BC (Smith 1966, 1969; Whiley et al. 2002).

17.1.3
Classical Breeding

The process of avocado breeding today continues to rely on open-pollination of promising varieties and – until very recently – without knowledge of the pollen source and unassisted by molecular techniques. This selection approach is the consequence of several obstacles to traditional breeding that are related to the biology of the avocado tree itself. The first and foremost among these is the inability to perform controlled pollinations (avocado trees produce well over a million tiny flowers that are abundantly shed; reviewed in Davenport 1986). The bulky size of the avocado tree makes large-scale experimental trials land- and labor-intensive. The widespread use of rootstocks introduces additional cost, conflicting breeding strategies, and can even confound data collection when the rootstock outgrows its scion. Finally, trees are slow to mature, and five or more years elapse before reliable estimates of fruit yield and quality are obtained.

Genome Mapping and Molecular Breeding in Plants, Volume 4
Fruits and Nuts
C. Kole (Ed.)

17.2
Molecular Beginnings

Faced with these substantial challenges, attempts have been made over the past 25 years to adopt molecular approaches. Various molecular studies have focused on the development of markers, including isozymes (Torres and Bergh 1980), RFLPs (Furnier et al. 1990; Davis et al. 1998), microsatellites (SSRs) (Lavi et al. 1994; Ashworth et al. 2004), VNTRs (Mhameed et al. 1996), and RAPDs (Kobayashi et al. 2000). Nonetheless, molecular technologies have been slow to take hold.

Several factors have so far conspired to delay the application of molecular tools to avocado breeding, in addition to those pertaining to the biology of the plant. Most importantly, the public sequence databases have been largely devoid of any suitable gene regions or ESTs. Moreover, avocado is a member of family Lauraceae that lacks major crop species and represents an ancestral lineage predating the origin of the Eudicots (Stevens 2001 onwards). The phylogenetically distant position of avocado decreases the utility of sequences from the major model organisms, e.g., *Arabidopsis,* which in turn has discouraged the pursuit of alternative molecular approaches, such as candidate genes and comparative gene expression. Similarly, linkage and genetic mapping studies (Lavi et al. 1991, 1992, 1993; Mhameed et al. 1996; Sharon et al. 1997, 1998), though providing preliminary insight into associations between such traits as skin color and fiber presence in the fruit flesh, are limited in scope due to a relative shortage of markers in relation to the number of chromosomes ($x = 12$), a lack of experimental replication, and limited baseline cytogenetic work. It therefore seems unlikely that genes will be localized to particular chromosomes in the near future.

Our laboratory is currently involved in a QTL study that, unlike previous research, attempts to control for genetic variation across sites and environments. Our aim is to work toward marker-assisted selection that will eliminate the delay between genotype selection and evaluation of maturity-dependent productivity traits. In parallel, we have been pursuing a comparative sequencing project to study haplotype diversity for several genes in various avocado cultivars and wild relatives/cultigens.

17.2.1
SSR Markers

The development of molecular marker technology (primarily SSR markers) has benefited the breeding process in the short-term by enabling the study of genealogical relationships among varieties and cultigens (e.g., Furnier et al. 1990; Davis et al. 1998; Ashworth and Clegg 2003; Schnell et al. 2003) and of parentage/pollen movement in avocado orchards (Goldring et al. 1987; Degani et al. 1989; Davenport et al. 1994; Kobayashi et al. 2000; Garner et al. 2006). SSRs are now the markers of choice for avocado in studies devoted to pollen flow and varietal differentiation, having replaced some of the more costly or less informative markers. These and other marker applications detailed below have fuelled the need for a greater number of markers. One of our research priorities has been to address this need by screening two genomic DNA libraries of ‘Hass’ for dinucleotide and trinucleotide markers, yielding some 150 useful marker loci. Of these, we routinely use some 30 loci (see Ashworth et al. 2004 for primer sequences), and the remainder has been screened for ease of interpretation and utility in distinguishing a panel of ca. 5–10 cultivars. In relation to the overall number of library clones sequenced (ca. 1,500), the yield of SSR markers has been exceedingly low, due primarily to clone redundancy, an apparent scarcity of loci, and poor interpretability of the banding profiles.

SSR Applications: Genealogical Studies These impediments notwithstanding, considerable information is beginning to accumulate on cultivar relationships, the delimitation of the three botanical races, and pollen movement within avocado groves.

Building on previous studies that used RFLP markers (Furnier et al. 1990; Davis et al. 1998), we analyzed genealogical relationships between 35 cultivars using 25 SSR markers (Ashworth and Clegg 2003). Consistent with prior research, heterozygosity was high, particularly for the SSR markers (73.4 and 52.6% for dinucleotide and trinucleotide markers, respectively). Low statistical support (bootstrap) for the assemblages denoting the three botanical races, even after (known) hybrid genotypes had been excluded from the analysis, suggested ancient hybridization or a more recent origin of the botanical races than previously thought. Indeed, the hybrid nature of most extant cultivars is probably responsible for homogenizing some of the racial differences.

One practical conclusion from our genealogical study is that botanical race assignments and geographic origin of a cultivar are often confused or applied inconsistently. For example, a "Guatemalan" cultivar may be so called because it has a rough thick green skin (characteristic of the Guatemalan botanical race) or because it was collected in Guatemala. Even when the geographic origin is disregarded, race assignment is based on a somewhat vague formula combining fruit skin thickness and color, presence or absence of anise scent, and origin (if known) of the maternal parent, with several other criteria added based on breeder experience. Clearly, this convention does not readily accommodate genotypes that have lost one or more of the key parental (maternal) traits (e.g., skin color, thickness; anise scent) through genetic segregation, for which ample evidence exists (e.g., Storey et al. 1984; Bergh 1966, 1967; Bergh and Whitsell 1974, 1975), or having a complex hybrid status (multiple backcrossing). Consequently, SSR markers are now also being used to eliminate or change the assignment of mislabeled material (Schnell et al. 2003) or to unravel parentage (Ashworth and Clegg 2003). These SSR studies show that phenotype tracks genotype in an unpredictable manner, and that greater rigor is needed in the delimitation of botanical races and in the description of avocado cultivars.

Another conclusion emerging from our work with SSR markers (Ashworth et al. 2004) is that not all markers developed using the 'Hass' genomic DNA library are useful for typing more distant genotypes, especially the more divergent rootstocks or the related species *P. schiedeana* and *P. steyermarkii*. Conversely, loci that show insufficient variation (one or two alleles) within a core group of cultivars sometimes harbor additional alleles when applied to the more distant genotypes. These observations have important implications for marker development strategies, particularly with respect to the cut-off applied to the number of repeat units that qualifies a sequenced clone for inclusion in the subsequent round of marker development.

SSR Applications: Pollination The co-dominance of SSRs means that allelic segregation can reveal paternity of a given progeny genotype, a major advance in a crop that defies controlled pollination and where paternal origin has hence always been subject to conjecture. The application of SSRs to pollination studies has therefore received considerable attention, not only from breeders, but particularly from avocado growers who are interested in the relationship between pollination and yields: 'Hass' fruit commands a far higher market price than alternative cultivars that are conventionally integrated into an orchard for the sole purpose of enabling cross-pollination. This planting practice stems from the observation that 'Hass' trees adjacent to certain "pollinizer" cultivars produce a profusion of fruit that often far exceeds that of 'Hass' trees located only one row farther away. However, this yield boost is not always observed, and growers are eager to embrace marker technology to better understand the relationship of selfing and cross-pollination on yield in order to calculate the trade-off between planting a 'Hass' monoculture versus interplanting 'Hass' with less profitable "pollinizer" cultivars. In a multi-location study, RAPD markers (Kobayashi et al. 2000) suggested that while progressive distance from a "pollinizer" cultivar strongly reduced the percentage of outcrossed fruit, the positive correlation between outcrossing and yield is only weak, with other factors accounting for most of the yield differential.

There continues to be considerable interest in deploying SSR markers to tease apart the roles of pollen source and competing environmental factors. However, as important insights into pollen flow within commercial avocado orchards are starting to accumulate, new conundrums are also revealed. Thus, recent SSR studies (Davenport et al. 2005; Garner et al. 2006) are corroborating earlier work based on RAPD markers and isozymes, showing that pollen (mediated by bees) does not move far from its source tree, yet consensus continues to be elusive with regard to yield differences between selfed and outcrossed fruit.

Sometimes, contradictory findings may have humble origins, and it is possible that differential sample collection methodologies are at least partially responsible. The prevailing wisdom stipulates that selfing reduces yields and that prematurely shed fruit is likely to be selfed whereas outcrossed fruit is more likely to be retained on the tree to maturity (Degani et al. 1997). Consequently, the sampling of marble-sized fruit – a common practice in experiments involving large-scale DNA extraction – may be skewing yield data on the assumption that it is likely to over-represent the fraction of selfed fruit relative to the more mature fruit harvested during commercial operations. Molecular studies would thus often be tracking a different (more highly selfed) pool of fruit than that collected during a typical harvest. For example, two studies in commercial orchards (Kobayashi et al. 2000; Garner et al. 2006) report a very low outcrossing percentage (based on sampling marble-sized fruit), yet our own work (with trees raised from fruit that had been picked at maturity) bespeaks much higher rates of outcrossing (see below). Further studies will be needed to resolve the relationship between sampling stage and the impact of selfing and outcrossing on fruit yields.

17.3
QTL Analysis

The traditional method of breeding avocado still relies on the selection of progeny from a maternal source interplanted with pollen donors having desirable properties, followed by prolonged field trials to ascertain agronomic value. With a view to placing the process of avocado improvement onto a molecular footing, we have begun a QTL study that will feed into a marker-assisted method of selection. SSR markers associated with desirable (quantitative) traits will be used to preselect among seedlings: those that carry the markers will be retained, while genotypes lacking the marker will be eliminated from the outset.

Our QTL study is centered on 200 'Gwen' progeny genotypes that are clonally replicated twice at each of two environments (coastal and inland Southern California; 800 trees overall) and grafted onto a uniform 'Duke 7' rootstock. Growth and yield-related traits (tree height, stem girth, canopy diameter and several fruit yield and quality traits) have been measured once or twice a year since 2001, and the collection of SSR marker data from over 100 loci is in progress. Preliminary marker data indicate that virtually all our 'Gwen' progeny genotypes are outcrossed. Paternity of about one quarter is unclear, but the remaining three quarters are sired by 'Fuerte', 'Bacon', and 'Zutano' in approximately equal proportions. This genetic composition will furnish valuable information on the influence of the pollen source against a uniform maternal genetic background under replicated experimental conditions.

17.4
Comparative Sequencing Study

In order to explore the avocado genome at multiple levels, our lab has sequenced four genes (cellulase, chalcone synthase, flavonol-3-hydroxylase, and serine-threonine kinase) in a panel of 54 *Persea* genotypes (33 cultivars and 21 wild relatives/cultigens from Mexico, Ecuador, Costa Rica, and the West Indies). The nucleotide diversity (θ_W) for the four genes ranges from 0.003 to 0.016. Nucleotide diversity is moderate compared with other plants such as maize and barley, averaging 0.0157 and 0.0134 in the cultivars and germplasm, respectively, for flavonol-3-hydroxylase. These diversity values suggest that, subject to a more prolific representation of avocado DNA sequences on public databases, SNP development will be feasible, resulting in a much enlarged pool of genetic markers and paving the way for a more diversified approach to molecular studies in avocado.

17.5
Prospects

Happily, ESTs and other DNA sequences are now starting to accumulate, as avocado has been selected as an exemplar representing the basal angiosperms in the Floral Genome Project (*http://fgp.bio.psu.edu/fgp/*). Already, some 10,000 ESTs have been generated. More elaborate studies are proposed that will focus additionally on comparative gene expression in fruit development and on genetic mapping (Douglas Soltis, personal communication). The synergy between a fortified sequencing effort, gene expression studies, and existing molecular tools will open up new opportunities and prospects for avocado breeding.

Acknowledgement. Funding for research conducted by the authors was provided by the California Avocado Commission

References

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. Sci Hort 101:255–267

Bergh BO (1966) A Hass open-pollinated progeny-set. California Avocado Society Yearbook 50:64–77

Bergh BO (1967) Some late-maturing avocado seedlings of various parentage. California Avocado Society Yearbook 51:131–158

Bergh BO, Whitsell RH (1974) Self-pollinated Hass seedlings. California Avocado Society Yearbook 57:118–126

Bergh BO, Whitsell RH (1975) Self-pollinated Fuerte seedlings. California Avocado Society Yearbook 58:128–134

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 Hort Sci 119:1200–1207

Davenport TL, Ying Z, Schnell RJ (2005) Use of simple sequence repeats (SSR) to determine incidence and effectiveness of self- and cross-pollinated avocado fruit in Southern California. In: Proc California Avocado Res Symp, Oct 29, 2005, Univ of California, Riverside, pp 62–71

Davis J, Henderson D, Kobayashi M, Clegg MT (1998) Genealogical relationships among cultivated avocado as revealed through RFLP analyses. J Hered 89:319–323

Degani C, Goldring Z, Gazit S (1989) Pollen parent effect on outcrossing rate in 'Hass' and 'Fuerte' avocado plots during fruit development. J Am Soc Hort Sci 114:106–111

Degani C, El-Batsri R, Gazit S (1997) Oucrossing rate, yield, and selective fruit abscission in 'Ettinger' and 'Ardith' avocado plots. J Am Soc Hort Sci 122:813–817

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

Garner LC, Ashworth VETM, Clegg MT, Lovatt CJ (2006) The impact of outcrossing on yields of 'Hass' avocado (*Persea americana* Mill.). J Am Soc Hort Sci (in press)

Goldring A, Gazit S, Degani C (1987) Isozyme analysis of mature avocado embryos to determine outcrossing rate in a 'Hass' plot. J Am Soc Hort Sci 112:389–392

Kobayashi M, Lin J-Z, Davis J, Francis L, Clegg MT (2000) Quantitative analysis of avocado outcrossing and yield in California using RAPD markers. Sci Hort 86:135–149

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

Lavi U, Lahav E, Degani C, Gazit S (1992) The genetics of the juvenile phase in avocado. J Am Soc Hort Sci 117:781–784

Lavi U, Lahav E, Degani C, Gazit S, Hillel J (1993) Genetic variance compositions and heritabilities of several avocado traits. J Am Soc Hort Sci 118:400–404

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

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

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 Hort Sci 128:881–889

Sharon D, Cregan PB, Mhameed S, Kusharska M, Hillel J, Lahav E, Lavi U (1997) An integrated genetic linkage map of avocado. Theor Appl Genet 95:911–921

Sharon D, Hillel J, Mhameed S, Cregan PB, Lahav E, Lavi U (1998) Association between markers and loci controlling avocado traits. J Am Soc Hort Sci 123:1016–1022

Smith CE Jr (1966) Archaeological evidence for selection in avocado. Econ Bot 20:169–175

Smith CE Jr (1969) Additional notes on pre-conquest avocados in Mexico. Econ Bot 23:135–140

Stevens PF (2001 onwards) Angiosperm phylogeny website. Version 6, May 2005: http://www.mobot.org/MOBOT/research/Apweb/ (January 2006)

Storey WB, Bergh BO, Platt RG, Miller M (1984) Observations on a second-generation progeny of a Mexican × West Indian cross. California Avocado Society Yearbook 68:161–165

Torres AM, Bergh BO (1980) Fruit and leaf isozymes as genetic markers in avocado. J Am Soc Hort Sci 105:614–619

Whiley AW, Schaffer B, Wolstenholme BN (eds) (2002) Avocado: Botany, Production and Uses. CABI Publ, New York, USA