California Avocado Society 2009 Yearbook 92:137-164

L. Tommasini
C.L. Calderon-Vazquez
V.E.T.M. Ashworth
M.L. Durbin
M.T. Clegg

Department of Ecology and Evolutionary Biology
University of California, Irvine, CA

Towards a Program of Marker-Assisted Selection on Valuable Avocado Traits

Introduction: an integrated approach to avocado improvement is needed to accelerate breeding progress and to reduce overall costs.

(Note: terms with a superscript [G] are found in the Glossary.)

Avocado (*Persea americana* Mill.) breeders are interested in high yield and yield stability, but also in a combination of diverse biotic stress resistances, especially resistance to phytophthora root rot (caused by *Phytophthora cinnamomi* Rands) and fruit diseases, like phytophthora crown rot, or collar rot (caused by *Phytophthora citricola* Sawada), and anthracnose (caused by *Colletotrichum gloeosporioides* Penz.), Dothiorella/Colletotrichum complex fruit rot (caused by *D. aromatica* and *C. gloeosporioides*), and sunblotch viroid ASBVd. Improvement for abiotic stress tolerances, such

as rootstock tolerance to salinity, cold, heat, and low soil oxygen content, also plays a major role in avocado breeding. In addition to biotic and abiotic factors, breeding for quality trait profiles, including taste, long shelf life and nutrient content are high priority improvement targets.

Avocado breeders face numerous challenges; for instance, controlled crosses are difficult to perform in avocado, a long immature stage limits the rate of breeding progress and extensive land and labor requirements make avocado improvement expensive. Crossing is complicated by a reproductive system where an individual tree can produce about a million flowers, yet only a small fraction of these flowers ever set fruit that reach maturity, and a fair amount of normal fruitlets tend to drop. Hand pollination is not always reliable, and the control of pollination by caging of trees has had limited success. Avocado is highly heterozygous[G] resulting in an unpredictable progeny. Seedlings produced by a single tree (or cultivar) are extremely variable. However, a compensating advantage of avocado is the ability to propagate desirable genotypes via bud grafting. Avocado breeding programs are further challenged by an extended juvenility period that can reach 15 years or more (Bergh et al., 1996). Varietal improvement relies on multi-year field trials during which large numbers of seedlings are grown to maturity and compared for desirable characteristics. Therefore conventional breeding approaches in avocado involve extensive time, land resources, water resources and expensive labor. As a consequence, introgression and pyramiding[G] of multiple traits[G] remain a great challenge for avocado breeding programs.

The main traits targeted by avocado breeders can be monogenic[G] or oligogenic[G], but are generally controlled by multiple genes (polygenic [G]), known as quantitative trait loci. Heritability [G] (H^2) is defined as the proportion of phenotypic variation in a population that is attributable to genetic variation among individuals. Heritability analyses estimate the relative contributions of differences in genetic and non-genetic factors to the total phenotypic variance in a population. The value of H^2 can range from 0.0 for no genetic influence on trait variation to 1.0 for complete genetic determination of trait variation. A high heritability estimate ($H^2 > 50\%$) means

that selection should be effective, because parental trait values are a good predictor of progeny trait values. H^2 measures both non-additive and additive sources of genetic variation. Mass selection is only effective on the additive variance, so H^2 is not a perfect predictor of selection response. Narrow sense heritability $(h^2)^G$ is the ratio of the additive component of genetic variance to total phenotypic variance. Lavi et al. (1993) showed that genetic variance (both additive and non-additive) is large for most avocado traits and a large non-additive (dominant) genetic variance was also estimated in avocado, possibly explained by high levels of heterozygosity (Lavi et al., 1991).

Previous studies have shown that selection using molecular markers linked to target traits (known as Marker-Assisted Selection or MAS) can increase the efficiency of conventional breeding programs for various traits (Tester and Langridge, 2010). However, due to a paucity of markers, minimal linkage maps, and inadequately characterized breeding populations, molecular markers have yet to be employed as selection tools in avocado improvement programs. Various types of molecular marker technologies have been developed since the emergence of the first markers in the 1980s (Phillips and Vasil, 2001). The most recent generation of molecular markers is based on direct analysis of sequence variation rather than indirect analyses based on cloned probes (so called RFLPs) or Polymerase Chain Reaction (PCR)-based markers, like microsatellite markers (SSRs). Single base changes in a target sequence, called Single Nucleotide Polymorphisms (SNPs, pronounced snips), are the most abundant source of variation in plant and animal genomes. Recent avocado resequencing work by Chen et al. (2008) predicts a SNP density of roughly one polymorphic site out of every 120–150 sites, based on sequence data from 4 nuclear loci in 21 wild avocado accessions. Moreover, direct sequence analysis is the most robust form for analyzing genomic variation. Another advantage of SNP marker analysis consists in the high probability of finding a marker within the gene of interest due to the high density of SNPs across the genome (Syvänen, 2005). This provides a considerable advantage MAS programs. MAS offers the potential to com-

bine target traits in the same genotype more precisely, with less unintentional losses and in fewer selection cycles[G]. MAS in avocado breeding can be especially beneficial for 1) targeting traits that are expensive or time-consuming to evaluate, have complex inheritance or low penetrance, 2) targeting traits whose selection depends on specific environments or developmental stages (e.g. fruit stage), 3) pyramiding multiple monogenic traits[G] or several QTLs for a single target trait with complex inheritance, and 4) for speeding up backcross breeding[G].

A brief review of heritability estimates for selected crops

In this article we focus on heritability of different traits of major interest for avocado breeders: specifically on the content of carotenoids, sitosterol, proanthocyanidin monomers in avocado fruits, % pulp weight/fruit weight, growth rate, flower abundance and fruit set. To set the stage, we first briefly review heritability estimates for these traits in crop plants. There are relatively few estimates of the heritability of growth rate for trees and especially for fruit trees. Studies include Scots pine (*Pinus sylvestris* L.; Ericsson and Fries, 2004; Jansson et al., 2003; Persson and Andersson, 2003), loblolly pine (*Pinus taeda* L., Gwaze et al., 2002), and white spruce (*Picea glauca* (Moench) Voss, Yu et al., 2003; Zhang et al., 2004). Narrow-sense heritability[G] from these species varies from 0.0 to 62%, but most values are less than 30%. Flowering intensity and fruit density were studied in avocado by Lavi et al. (1993) and they estimated a moderate heritability value (H^2=49%). However the additive component of genetic variance was non-significant. The existence of substantial non-additive genetic variance was also indicated by narrow-sense and heritability values estimated for other avocado traits (Lavi et al., 1993). Heritability studies on fruit weight have been carried out in various species, including mango (*Mangifera indica* L.; Brown et al., 2009), chilli (*Capsicum annuum* L.; Manju and Sreelathakumary, 2002; Sreelathakumary and Rajamony, 2004), watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai; Gusmini and

Wehner, 2007) and peanut (*Arachis hypogaea* L.; Chiow and Wynne, 1983). Heritability from these species was generally very high, ranging from 81 to 99%, except for peanut (36%) and watermelon (41 to 59%). No specific studies of heritability of % pulp weight / fruit weight (% pw/fw) are available in the literature. However, estimates of the heritability of seed weight in different species are generally high: in cotton (*Gossypium hirsutum* L.) heritability was 77% (Khan et al., 2010), in cowpea (*Vigna unguiculata* L. (Walp.)) heritability for 100-seed weight was 75% and showed positive significant correlation with yield (Kheradnama and Niknejada, 1974), and in Lima beans heritability was 98% and 100-seed weight was one of the main yield component (Akande and Balogun, 2007). Godoy and Norden (1981) suggested that fruit and seed size traits are controlled by different genes based on a study in segregating peanut populations.

With regard to biochemical traits, a single heritability study on sitosterol content showed high heritability (H^2=90%) in oilseed rape (*Brassica napus* L., Amar et al., 2008a). Amar et al. (2008a) found a mild effect of the environment on β -sitosterol and a significant genotype x environment interaction effect[G]. Heritability of total carotenoids content, including ζ -carotene, α-carotene, β-carotene, phytoene, lycopene, where each of these carotenoids were estimated individually, was carried out in carrots (*Daucus carota* L.; Santos and Simons, 2006). Estimates of heritability were very high within a specific cross in carrot (from 89 to 98%), but moderate in a different cross (38 to 45%). Other studies were carried out separately on heritability of carotene, cryptoxanthin, zeaxanthin and lutein concentration in maize (*Zea mays* L.; Wong, 1999), of β-carotene concentration in durum wheat (*Triticum durum* Desf.; Santra, et al. 2005) and in sorghum (*Sorghum bicolor* L., Reddy et al., 2005), and of carotene in lettuce (*Lactuca sativa* L., Gupta et al., 2008). Estimates of β-carotene heritability in durum wheat and sorghum varied from 67 to 99%. This variation depended on the crosses studied, indicating the presence of additive gene effects. In maize, the heritability estimate for carotene were 33%, for cryptoxanthin 47%, for zeaxanthin 59%, and 78% for lutein (Wong, 1999). Heritability of proanthocyanidin levels had

been previously studied in American cranberry (*Vaccinium macrocarpon* Ait.; Vorsa et al., 2003). Heritability studies on condensed tannins, including procyanidins and leucoanthocyanidins, were also conducted in sorghum (*Sorghum bicolor* L., Paroda et al., 1975; Woodruff et al., 1982) and common bean (*Phaseolus vulgaris* L., Ma and Bliss, 1978). In these species, heritability for overall proanthocyanidins or condensed tannins varied from intermediate to high (43% to 89%).

Quantitative genetic analyses in avocado populations

Heritability studies as well as morphological, biochemical, molecular and QTL studies are being conducted using a replicated experimental avocado population of over 800 trees established in 2000-2001. The experimental population is composed of 204 genotypes that all share cultivar Gwen (G) as their maternal parent. The paternal parents are: 54 Gwen x Fuerte (G x F), 58 Gwen x Zutano (G x Z), 44 Gwen x Bacon (G x B) and the remainder are Gwen progeny having a wide assortment of male parents. Each avocado cultivar is highly heterozygous, leading to considerable diversity among the progeny of these crosses. Each progeny genotype was propagated via bud grafting and replicated four times, with two replicates grown at the UC South Coast Research and Extension Center (SCREC) in Irvine, CA, and two replicates grown at Agricultural Operations (AgOps) at UC, Riverside, CA. Classical quantitative genetic analyses combined with the level of replication within this population allow for the estimation of the non-genetic component of a phenotypic variance as well as the total genetic component of variance.

Chen et al. (2007) investigated heritability in this experimental population for different traits of agronomic interest, based on segregating progeny from Gwen X Fuerte (G X F), Gwen X Zutano (G X Z) and Gwen X Bacon (G X B) crosses (described above). The traits studied included growth rate, fruit set and flower abundance. Growth rate is of particular interest considering the avocado's long juvenility period.

Traits studied in our current work on the G X F cross include fruit weight, seed weight, % pulp weight/fresh weight

(% pw/fw) and nutrient content. We also evaluated fruit set based on observations from published studies in cucumber (*Cucumis sativus* L.; Cramer and Wehner, 2000; Taha et al., 2003) that indicated that fruit set correlates to yield, and that success rate of hand-pollination depends on fruit set. Similarly, fruit size/weight was found to be correlated to yield in different fruit species, such as peanut (*Arachis hypogaea* L., Chiow and Wynne, 1983), watermelon (*Citrullus lanatus* (Thunb.) Matsum & Nakai, Gusmini and Wehner, 2007) and chilli (*Capsicum* spp., Sreelathakumary and Rajamony, 2002).

Recently consumers have become more concerned with the nutritional value of foods. Avocado has been singled out as having beneficial effects on human health. Avocado fruits have a unique biochemical profile and contain a variety of bioactive components, including carotenoids (Lassen et al., 1944; Lu et al., 2005), β-sitosterol (Duester, 2001), monounsaturated fatty acids, in addition to B vitamins, vitamins C and E (Slater et al., 1944), and proanthocyanidins or condensed tannins that come in a multitude of chemical structures and sizes consisting of base units called “monomers” (Vinson et al., 2001). These different constituents have an array of beneficial properties for human health acting as antioxidants, cholesterol-lowering and anticarcinogenic agents. Given the health advantage of avocado over many other crops, the evaluation of nutritional traits is highly desirable. Clearly, it is vital to enhance our understanding of the genetic underpinnings of these traits and to develop affordable biochemical and molecular assays for future breeding efforts.

Estimates of heritability

From 2002 to 2005 Chen et al. (2007) measured growth rate based on the following attributes: 1) tree height, measured from ground to treetop; 2) average canopy diameter, measured as the distance from the widest part of the tree canopy in two dimensions—parallel to the row and perpendicular to the row; 3) average tree trunk diameter, measured from about 10 cm above the ground in two perpendicular directions. In the same years, flower abundance was visually

estimated and was coded as: 0, none; 1, a few; 2, moderate; and 3, many. Also fruit set was coded as: 0, none; 1, a few; 2, moderate; and 3, many.

Heritability for tree height, canopy diameter and trunk diameter was 34.3, 29.7, and 28.5%, respectively, and corresponding values for flowering abundance and fruit set per tree accounted for 32.3 and 23.4%, respectively (Chen et al., 2007). These heritability values are typical of other studies cited above and may be sufficient to permit breeding advance. Genotype x environment interaction was weak for flowering and fruit set yield per tree. Growth rates did not correlate with flowering abundance, and only a moderate correlation was found between growth rate and fruit yield per tree. These results are reported in Table 1 and Table 2.

Table 1. Heritability (H^2) of growth rate, flower abundance, and fruit set of the progeny of 'Gwen' avocado and genotype-by-location interactions. *Significant at $P < 0.05$ (Chen et al., 2007)

	Heritability H^2 (%)	Genotype by location G x E
Tree height	35.5	0.003
Canopy diameter	30.3	0.05
Trunk diameter	26.6	0
Flower abundance	33.8	0.219*
Fruit set	23.0	0.171*

Heritabilities of all traits were moderate (around 30%), except for fruit set (H^2 = 23%). The environmental variance was almost certainly inflated by the different planting times of trees, environmental differences between the locations, and the low level of genotype replication in the two locations. The low heritability of fruit set may be also determined by the limited data available, since data were collected for one year. Genotype-by-location interactions in tree growth rate for height, canopy diameter, and trunk diameter are small and non-significant. However, significant but modest location by genotype interactions for flower abundance and fruit set were detected (Table 2).

Table 2. Correlation among growth rate, flower abundance, and fruit set of the progeny of 'Gwen' avocado. ***Significant at P < 0.001 (Chen et al., 2007)

	Tree height	Canopy diameter	Trunk diameter/ Stem girth	Flower abundance	Fruit set
Tree height	1	0.681***	0.662***	0.095	0.524***
Canopy Diameter		1	0.665***	0.089	0.488***
Trunk diameter/Stem girth			1	0.081	0.411***
Flower abundance				1	0.179***
Fruit set					1

From fall 2008 to winter 2010, mature fruits, after reaching more than 20% dry weight, were harvested and samples frozen, for subsequent evaluation of nutritional phenotypes, pulp weight, seed weight, and the average of whole fruit weight over two seasons (winter 2009 and winter 2010). Percent pw/fw was estimated from these data. Subsequently, ripe fruit mesocarp tissue from the same fruits was analyzed for nutritional composition for one season (winter 2009). The methods for the extraction, identification and quantification of proanthocyanidin monomers, carotenoids (including β-carotene, α-carotene, β-cryptoxanthin, lycopene, lutein and zeaxanthin) and β-sitosterol have been elaborated. A modification of the carotenoid assays described by Luterotti et al. (2006) and Lichtenthaler (1987) was used to estimate total carotenoid content. The protocol described by Hagerman (2002) and Porter et al. (1989) was used to estimate proanthocyanins content. A modification of the sterol assay described by Jeong and Lachance (2001) was used to measure sterols content. Sterols were analyzed by Thin Layer Chromatography (TLC), while proanthocyanidins and carotenoids were analyzed by spectrophotometry. For each data set we calculated heritability and related parameters including genotype x environment interaction effects. For growth

rate, flower abundance and fruit set, we also calculated trait correlations following the methods of Chen et al. (2007).

Heritability of % pw/fw was high (63%). Environment and year also have a highly significant effect ($p<0.001$), as well as genotype x environment and genotype x year interactions ($p<0.0001$, Table 3).

Heritability of β-sitosterol content is high (80.6%) and genotype has by far the largest effect on sitosterols ($p<0.001$) while environment has a more limited effect ($p < 0.05$) and genotype x environment interaction is not significant (Table 3).

Heritability of total carotenoid content (including β-carotene, α-carotene, β-cryptoxanthin, lycopene, lutein and zeaxanthin) content is also high (76.3%). Both genotype and environment have a highly significant effect on carotenoid content ($p < 0.0001$), while the genotype x environment interaction is not significant ($p = 0.53$, Table 3).

Heritability of proanthocyanidin monomers content is moderate (30.7%). Both genotype and environment have a significant effect on proanthocyanidin content ($p < 0.01$). Genotype effect ($p < 0.0001$) is higher than environmental effect ($p= 0.0087$). Genotype x environment interaction is highly significant as well ($p < 0.001$, Table 3).

Table 3. Heritability values of fruit traits measured in clonal replicates of Gwen x Fuerte avocado genotypes. Data in bold were not normally distributed and therefore transformed.

	Heritability (H^2)%	G	E	G x E	Y	G x Y
% pw/fw	63%	***	**	***	**	**
β-sitosterol	80%	***	*	NS	^	^
Carotenoids	76%	**	**	NS	^	^
Proanthocyanidins monomers	30%	***	**	***	^	^

Abbreviations: H^2 = heritability, G = genotype, E = environment, G x E = effect of interaction between G and E, Y = year, G x Y = effect of interaction between G and Y, NS = non-significant. Significance at $p < 0.001$, < 0.01 and < 0.05 are indicated as ***, ** and *. ^ = assayed in 2009 samples only.

There are no published studies on heritability of % pw/fw. In avocado our estimates of heritability from avocado trees fall in the mid range of previous results in mango, chili and watermelon (81% to 99%; Brown et al., 2009; Manju and Sreelathakumary, 2002; Sreelathakumary and Rajamony, 2004; Gusmini and Wehner, 2007). We found a significant effect of the environment on heritability of % pw/fw, in agreement with the study by Gusmini and Wehner (2007) on watermelon, that showed that a high number of factors influence fruit weight.

To our knowledge, only one heritability study on sitosterol was carried out previously, and this study was conducted in oilseed rape (Amar et al., 2008a). High heritability (90%) was detected, similarly to our findings (80.6%). In agreement with Amar et al. (2008a), we found a mild effect of the environment on β-sitosterol ($p<5\%$), but contrary to Amar's findings, in our study G x E interaction was not significant.

Our results showed that in avocado total carotenoid content heritability is intermediate between heritability estimates of total carotenoid content in two carrot crosses (Santos and Simons, 2006). Consistent with our findings, β-carotene heritability in durum wheat (Santra et al. 2005) and in sorghum (Reddy et al. 2005) varied from 67 to 99%. On the contrary, carotene heritability estimates in lettuce and maize were lower (50.81%) than total carotenoid heritability in avocado. In agreement with our results, previous studies showed that environment had comparatively little influence on carotenoids (Santra et al., 2005). The total carotenoids assay used in this study includes β-carotene, α-carotene, β-cryptoxanthin, lycopene, lutein and zeaxanthin. Based on standard USDA carotenoid measurements (http://www.ars.usda.gov/ba/bhnrc/ndl) lutein and zeaxanthin are expected to be the most abundant components detected by the total carotenoids assay presented in this study. This could explain why our estimates of heritability for total carotenoids are intermediate between zeaxanthin (59%) and lutein (78%) as described for maize in Wong (1999) study.

Overall proanthocyanidins or condensed tannins heri-

tability estimates from American cranberry (Vorsa et al., 2003), sorghum (Paroda et al., 1975; Woodruff et al., 1982) and common bean (Ma and Bliss, 1978) varied from intermediate to high (43% to 89%). Our estimates of heritability from avocado fruits fall below the estimates of previous studies (30%). Overall, these data showed that growth rate, fruit set and flowering abundance, and proanthocyanidins have a moderately low heritability (around 30%). The level of heritability of measured traits also depends on whether trait evaluation can be replicated well across different environments and periods of time. Traits with moderate heritability may be controlled by many QTL of small effects and the presence of complex interaction networks between these may limit the possibility to detect and clone QTLs. However, marker-assisted selection can be made efficient if the effects of G xE interactions and epistasis[G] can be predicted (Openshaw and Frascaroli, 1997). When G x E and epistasis are important, it may be necessary to re-evaluate QTL effects within each breeding program (Podlich et al., 2004). Breeding for low penetrance or complex traits can greatly benefit from the application of marker-assisted selection.

Carotenoid and sitosterol content in avocado showed high heritability estimates and more limited but significant location effects, in general agreement with previous quantitative genetic studies in other plant species. These results suggest that carotenoids and sitosterol content in avocado may be oligogenic traits, in agreement with indications from previous QTL studies in different crop species (Abbo et al., 2005; Amar et al., 2008b; Rousseaux et al., 2005). High heritability estimates indicate that carotenoid and sitosterol content in avocado could be efficiently improved by selection. Marker-assisted selection has the potential to increase selection efficiency and to pyramid multiple monogenic traits or several QTLs for a single target trait with complex inheritance.

Discovery of SNPs in genes involved in nutritional pathways

The high heritability observed for some nutrient traits

suggests that Single Nucleotide Polymorphisms within nutrient-related candidate genes might predict the transmission of these traits. Candidate nutrient-related genes were identified using functionally characterized Arabidopsis gene sequences deposited at TAIR (www.arabidopsis.org) with high similarity to avocado EST/cDNA sequences from fruit-, flower- and other organs-specific libraries, developed by Cornell University (www.floralgenome.org), HortResearch (New Zealand), and CINVESTAV (Mexico). PCR amplification and sequencing of candidate genes was first done using our fruit cDNA library, and subsequently using genomic DNA extracted from Gwen x Fuerte progenies from our experimental population. SNPs were identified by standard resequencing using the Sanger method within gene fragments of about 500 bp. To date we have identified 78 SNPs by resequencing 36 candidates gene fragments that are part of pathways leading to the biosynthesis of sitosterol, isoprenoids, carotenoids, flavonoids, anthocyanins, phenylpropanoids, fatty acids, B vitamins, vitamin C and vitamin E (Table 4). The sequences obtained represent promising candidate genes that can be tested for associations with nutritional phenotypes.

Table 4. Identification of SNPs in genes related to nutritional compounds in avocado fruits

Biochemical Pathway or Gene Function Catagory	Gene Name	Expressed in fruit	Detected SNPs
Carotenoids			
	1. Lycopene beta cyclase (LBC)	X	1
	2. Phytoene synthase (PSY)	X	2
	3. Carotene beta-ring hydroxylase (LUT5)	X	1
	4. Zeta-carotene desaturase (ZDS)	X	1
Vitamin B complex			
Vitamin B1 (thiamine)	5. 1-deoxy-D-xylulose-5-phosphate synthase (DXPS1)	X	5
Vitamin B2 (riboflavin)	6. GTP-Cyclohydolase II; 3,4-dihydroxy-2-butanone-4-phosphate synthase (GCH)	X	
Vitamin B2 (riboflavin)	7. COI1 suppressor 1 (COS1)	X	
Vitamin B5 (pantothenic acid)	8. Branched-chain aminotransferase 3 (BCAT3)	X	1
Vitamin B5 (pantothenic acid)	9. Branched-chain aminotransferase 5 (BCAT5)		
Vitamin B6 (pyridoxin)	10. Pyridoxin biosynthesis 1 (PDX1)	X	4
Vitamin B6 (pyridoxin)	11. Pyridoxin biosynthesis 2 (PDX2)	X	2
Vitamin B9 (folic acid)	12. Thymidylate synthase 1 (THY-1)	X	
Vitamin B9 (folic acid)	13. Thymidylate synthase 2 (THY-2)	X	

Vitamin B9 (folic acid)	14. Aminotransferase class IV family (atrans)	X	6
Vitamin B9 (folic acid)	15. 10-formyltetrahydrofolate synthetase	X	
Vitamin C			
GDP-D-mannose biosynthesis	16. Phosphoglucose isomerase	X	4
GDP-D-mannose biosynthesis	17. Phosphomannomutase	X	
GDP-D-mannose biosynthesis	18. GDP-mannose pyrophosphorylase (VITAMIN C DEFECTIVE 1)	X	5
GDP-D-mannose biosynthesis	19. Mannose-6-phosphate isomerase	X	
GDP-L-galactose biosynthesis	20. GDP-mannose-3',5'-epimerase	X	2
Ascorbate biosynthesis	21. L-galactose-1-phosphate phosphatase	X	
Ascorbate biosynthesis	22. GDP-L galactose phosphorylase (VITAMIN C DEFECTIVE 2)	X	3
Ascorbate biosynthesis	23. L-galactose dehydrogenase	X	
Isoprenoid & sitosterol			
Farnesyl diphosphate biosynthesis	24. Farnesyl diphosphate synthase	X	3
Sterol biosynthesis	25. Cycloeucalenol cycloisomerase	X	5
Squalene biosynthesis	26. Squalene synthase (SQS1)	X	1
Campestrol and sitosterol biosynthesis	27. 24-dehydrocholesterol reductase	X	3
Vitamin E			
Alpha and gamma tocopherol biosynthesis	28. Tocopherol cyclase (VITAMIN E DEFECTIVE 1=VTE1)	X	
	29. Homogentisate phytyltransferase (VTE2)	X	4
	30. MPBQ/MSBQ methyltransferase (VTE3)	X	1
	31. Gamma-tocopherol methyltransferase (VTE4)	X	4
	32. 4-hydroxyphenylpyruvate dioxygenase (PHYTOENE DESATURASE 1)	X	3
Fatty acid pathway			
	33. LACS9, long chain acyl-CoA synthetase 9	X	
	34. LACS2, long chain acyl-CoA synthetase 2	X	
	35. LACS7, long chain acyl-CoA synthetase 7		
	36. Acyltransferase, Cuticular 1 (CUT 1)	X	2
	37. ECR, enoyl-CoA reductase	X	
	38. Lipoxygenase (LOX)	X	
Flavonoid, anthocyanin & phenylpropanoid pathways			
	39. Anthocyanidin synthase (ANS)	X	
	40. Phenylalanine ammonia-lyase 2 (PAL2)	X	2
	41. Flavonol 3'-O-methyltransferase 1 (OTM1)	X	7
	42. Caffeoyl-CoA O-methyltransferase (caff3)	X	5
	43. Chalcone synthase (CHS)	X	
	44. Flavonol 3-hydroxylase (F3H)	X	
Pulp browning			
	45. Polyphenol oxidase (PPO)	X	
	46. Putative laccase/diphenol oxidase	X	
Totals			77

Future goals

a. Linking SNPs to traits

To evaluate whether the Single Nucleotide Polymorphisms (SNPs) identified within putatively nutrient-related genes and SSR markers (Ashworth et al., 2004) are correlated to fruit nutrient content, we intend to run QTL analyses (PROC QTL) using a genetic map we are developing based on the Gwen x Fuerte population. Although a significant correlation of a SNP is likely to be highest for nutrients from its source pathway, we will check for correlations with nutrients from the other pathways. We will also search for any correlations of SSR markers with all the phenotypic traits described above.

b. Searching for regulatory and other genes involved in the expression nutritional traits

A second phase of the project, now underway, will search for new molecular markers using microarray[G] analyses. In this phase we seek to identify candidate transcription factors[G] that control the expression of nutrient-related genes at the unripe and ripe fruit stage. Subsequently, gene expression variation at these loci will be examined across different Gwen x Fuerte progeny. A total of 536 avocado unigenes matching 240 unique Arabidopsis transcription factor loci, and 246 metabolism-related genes, including the nutrient-related genes targeted for SNP discovery as described above (Tab. 4), were included in the microarray. The microarrays also included 15 genes that are involved in the sucrose and galactose degradation pathways, and a polyphenol oxidase that is putatively involved in tissue browning. This microarray platform can detect gene expression in avocado fruits by hybridizing its probes to antisense RNA copied from messenger RNA extracted from fruit samples. Red and green fluorophores are used to characterize respective genotypes in each pair-wise genotype comparison. The array design follows the interwoven loop design by Kerr and Churchill (2001) and includes 10 genotypes chosen to represent extremes of

nutrient levels with four biological replicates for each genotype and a total of 40 array hybridizations, 80 targets and two developmental stages (immature, defined as 14.0-16.3% dry weight and very mature defined as 27-32% dry weight).

c. Marker-assisted selection

The ultimate goal of our work is to identify SNPs or SSRs that predict the transmission of valuable traits. Traits with high heritabilities are the most likely to show significant trait-marker associations. The statistical power to detect associations is very limited when the environmental component of variance is high, so traits with low to moderate heritabilities are unlikely to be attractive targets for MAS. Accordingly, we expect that any significant trait-marker associations detected will be associated with traits that show high heritability. The final phase of the project is to select approximately eight of the "best" parent genotypes based on nutrient content, % pw/fw, tree growth, fruit set and flowering abundance. About 200 seedlings from each of the selected trees will be assayed for transmission of markers associated with desirable traits. Seedlings that received the desirable markers will be saved for further testing. We expect to use this selection criterion to remove about 98% of all seedlings and to therefore apply intense selection at an early stage based on markers that predict the transmission of valuable fruit-related traits that are not expressed until the tree reaches maturity. This Marker-Assisted Selection scheme will allow the application of intense selection without the land, labor and time costs associated with conventional selection. The prospects for accelerating the rate of avocado improvement are very good and it is likely that improved avocado cultivars will emerge from this program over the next decade.

GLOSSARY

Backcross breeding:
Backcross breeding is a crossing of a hybrid with one of its parents or an individual genetically similar to its parent, in order to achieve offspring with a genetic identity which is closer to that of the parent.

Breeding cycle:
A breeding cycle is composed of a series of steps: creation of variation, selection, evaluation, release, multiplication, and distribution of the new variety.

DNA microarray:
A multiplex technology that consists of an arrayed series of thousands of microscopic spots of DNA sequences (oligonucleotides) than can be a short section of a gene or other DNA element that are used to hybridize a cDNA or cRNA sample (called target). Probe-target hybridization is usually detected and quantified by detection of labeled targets to determine the relative abundance of nucleic acid sequences in the target. A microarray experiment can accomplish many genetic tests in parallel, accelerating many types of investigation.

Epistasis:
The masking of the phenotypic effect of alleles at one gene by alleles of another gene is defined as epistasis. A gene is said to be epistatic when its presence suppresses the effect of a gene at another locus.

Gene pyramiding:
The act of breeding together genes contained in different loci.

Genotype x environment interaction effect:
The influence of specific combinations of genetic and environmental factors on a trait that goes beyond the additive action of these factors. This can refer to genes that control sensitivity to the environment or environmental factors that influence gene expression.

Heritability (H^2):
Heritability is the proportion of the total variation between individuals in a given population due to genetic variation. This number can range from 0 (no genetic contribution) to 1 (all differences on a trait reflect genetic variation).

Heterozygous:
Two possible states of a gene for a diploid organism. Each gene is made up of two representative alleles - one inherited from the maternal source and the other inherited from the paternal source. When a gene is homozygous, both alleles for that gene are the same genotype consisting of two different alleles of a gene for a particular trait.

Monogenic trait:
A character/trait determined by a single gene.

Narrow-sense heritability (h^2):
Narrow-sense heritability is the ratio of the additive component of genetic variance to total phenotypic variance. Additive genetic variance is the part on genetic variance controlled by alleles that contribute a fixed value to the metric value of a quantitative trait.

Oligogenic trait:
A phenotypic trait produced by two or more genes working together.

Polygenic trait:
A trait determined by many genes at different loci (called Quantitative Trait Loci, QTLs), with small additive effects. Polygenic traits typically show a continuous variation.

Transcription factors:
A transcription factor is a protein that binds to specific DNA sequences, thereby controlling the transfer (or transcription) of genetic information from DNA to mRNA. A defining feature of transcription factors is that they contain one or more DNA-binding domains, which attach to specific sequences of DNA adjacent to the genes that they regulate.

References

Abbo S, Molina C, Jungmann R, Grusak MA, Berkovitch Z, Reifen R, Kahl G, Winter P, Reifen R. 2005. Quantitative trait loci governing carotenoid concentration and weight in seeds of chickpea (*Cicer arietinum* L.). Theor Appl Genet 111:185–195.

Akande SR and Balogun MO. 2007. Evaluation and heritability studies of local Lima bean (*Phaseolus lunatus* L.) cultivars from south-west Nigeria. Revista UDO Agrícola 7(1):22-28.

Amar S, Becker HC and Möllers C. 2008a. Genetic variation and genotype x environment interactions of phytosterol content in three doubled haploid populations of winter rapeseed. Crop Sci 48:1000-1006.

Amar S, Ecke W, Becker HC, Möllers C. 2008b. QTL for phytosterol and sinapate ester content in *Brassica napus* L. collocate with the two erucic acid genes. Theor Appl Genet 116(8):1051–1061.

Ashworth VETM, Kobayashi MC, de la Cruz M, Clegg MT. 2004. Microsatellite markers in avocado (*Persea americana* Mill.). Developing dinucleotide and trinucleotide markers. Sci Hortic - Amsterdam 101:255–267.

Bergh BO and Lahav E. 1996. Avocados. pp. 113–166. In: J. Janick and J.N. Moore (Eds.). Fruit Breeding. Vol. I: Tree and Tropical Fruits. John Wiley and Sons, Inc., West Lafayette.

Brown JS, Schnell RJ, Avala-Silva T, Moore JM, Tondo CL, and Winterstein MC. 2009. Broad-sense heritability estimates for fruit color and morphological traits from open-pollinated half-sib mango families. HortScience 44(6):1552-1556.

Chen H, Ashworth VETM, Xu S, Clegg MT. 2007. Quantitative Genetic Analysis of Growth Rate in Avocado. Amer Soc Hort Sci 132(5):691–696.

Chen H, Morrel PL, De La Cruz M, Clegg MT. 2008. Nucleotide Diversity and Linkage Disequilibrium in Wild Avocado (*Persea americana* Mill.). J Hered 99(4):382–389.

Chiow HY and Wynne JC. 1983. Heritabilities and Genetic Correlations for Yield and Quality Traits of Advanced

Generations in a Cross of Peanut. Peanut Sci 10:13-17.
Cramer CS and Wehner TC. 2000. Path analysis of the correlation between fruit number and plant traits of cucumber populations. ASHS Southern Region. Annual Meeting No.60, Lexington, Ky., USA, Vol. 35, No 4, pp. 550-570 (27 ref.), pp. 708-711.
Duester KC. 2001. Avocado fruit is a rich source of beta-sitosterol. J Am Diet Assoc 101:404–405.
Ericsson T and Fries A. 2004. Genetic analysis of fiber size in a fullsib *Pinus sylvestris* L. progeny test. Scand J For Res 19:7–13.
Godoy IJG and Norden AJ. 1981. Shell and seed size relationship in peanuts. Peanut Sci 8:21-24.
Gupta AJ, Dolma T, Chattoo MA, Yasmin S. 2008. Estimation of Genetic Variability and Heritability in Lettuce (*Lactuca sativa* L.). Indian Journal of Plant Genetic Resources 21(2)
Gusmini G and Wehner TC. 2007. Heritability and Genetic Variance Estimates for Fruit Weight in Watermelon. HortScience 42(6):1332–1336.
Gwaze DP Harding KJ, Purnell RC, Bridgwater FE. 2002. Optimum selection age for wood density in loblolly pine. Can J For Res 32:1393–1399.
Hagerman AE. 2002. Acid butanol assay for proanthocyanidins (pp. 48-49) in: Tannin analysis. http://www.users.muohio.edu/hagermae/.
Jansson G, Li BL, Hannrup B. 2003. Time trends in genetic parameters for height and optimal age for parental selection in Scots pine. For Sci 49:696–705.
Jeong WS and Lachance PA. 2001. Phytosterols and fatty acids in fig (*Ficus carica*, var. *mission*) fruit and tree components. J Food Sci 66(2):278-281.
Kerr MK, Churchill GA. 2001. Experimental design for gene expression microarrays. Biostatistics 2:183–201.
Khan NU, Marwat KB, Hassan G, Farhatullah, Batool S, Makhdoom K, Ahmad W, Khan HU. 2010. Genetic variation and heritability for cotton seed, fiber and oil. Traits in *Gossypium hirsutum* L. Pak J Bot 42(1):615-625.
Kheradnama M and Niknejada M. 1974. Heritability estimates and correlations of agronomic characters in cowpea (*Vi-*

gna sinensis L.). J Agr Sci 82:207-208.
Lassen D, Bacon K, Sutherland J. 1944. Chromatographic investigation of the carotenoid pigments of the avocado. Food Res 9: 427–433.
Lavi U, Lahav E, Degani C, Gazit S, Hillel J. 1993. Genetic variance components and heritabilities of several avocado traits. J Amer Soc Hort Sci 118(3):400-404.
Lavi U, Lahav E, Genizi A, Gazit S, Hillel J. 1991. Quantitative genetic analysis of traits in avocado cultivars. Plant Breed 106:149-160.
Lichtenthaler HK. 1987. Chlorolphylls and Carotenoids: Pigments of Photosynthetic Biomembranes. Methods in Enzymology 148:350-382.
Luterotti S, Bicanic D, Požgaj R. 2006. New simple spectrophotometric assay of total carotenes in margarines. Anal Chim Acta 573–574:466–473.
Ma Y and Bliss FA. 1978. Tannin Content and Inheritance in Common Bean. Crop Sci 18:201-204
Manju PR and Sreelathakumary I. 2002. Genetic variability, heritability and genetic advance in hot chilli (*Capsicum Chinese*, Jacq.). Journal of Tropical Agriculture 40:4-6.
Openshaw S and Frascaroli E. 1997. QTL detection and marker assisted selection for complex traits in maize. Annu. Corn Sorghum Research Conference Proceedings 52:44–53.
Paroda RS, Saini ML and Arora SK. 1975. Inheritance of tannin content in Eu-Sorghums Z Pflanzenzuchlg 74:251-256.
Persson T and Andersson B. 2003. Genetic variance and covariance patterns of growth and survival in northern *Pinus sylvestris*. Scand J For Res 18:332-343.
Phillips RL and Vasil IK (ed.). 2001. DNA-based markers in plants. Kluwer Academic, Dordrecht, The Netherlands.
Podlich DW, Winkler CR, and Cooper M. 2004. Mapping as you go: An effective approach for marker-assisted selection of complex traits. Crop Sci 44:1560–1571.
Porter LJ, Methods in Plant Biochemistry. 1989. Vol. 1, Plant Polyphenols (ed. J.B. Harborne), pp. 389-419, Academic Press London.
Reddy BVS, Ramesh S, Longvah T. 2005. Prospects of breed-

ing for micronutrients and beta-carotene-dense sorghums. Journal of SAT Agricultural Research, 14 pages. DOI:10.3914/ICRISAT.0126

Rousseaux MC, Jones CM, Adams D, Chetelat R, Bennett A, Powell A. 2005. QTL analysis of fruit antioxidants in tomato using *Lycopersicon pennellii* introgression lines. Theor Appl Genet 111:1396–1408.

Santra M, Santra DK, Rao VS, Taware SP, Tamhankar SA. 2005. Inheritance of β-carotene concentration in durum wheat (*Triticum turgidum* L. ssp. *durum*). Euphytica 144: 215–221.

Santos CAF and Simon PW. 2006. Heritabilites and miminimum gene number estimates of carrot carotenoids. Euphytica 151:79-86.

Slater GG, Shankman S, Shepherd JS, Alfin-Slater RB. 1975. Seasonal variation in the composition of California avocados. J Agric Food Chem 23:468–474.

Sreelathakumary I, Rajamony L. 2002. Variability, heritability and correlation studies in chilli (*Capsicum annuum* L.) under shade. Indian J Hortic Sci 59(1):77-83.

Sreelathakumary I and Rajamony L. 2004. Variability, heritability and genetic advance in chilli (*Capsicum annuum* L.). Journal of Tropical Agriculture 42(1-2):35-37.

Syvänen AC. 2005. Toward genome-wide SNP genotyping. Nat Genet 37:S5–S10.

Taha M, Omara K, El Jack A. 2003. Correlation among growth, yield and quality characters in *Cucumis melo* L.. Cucurbit Genetics Cooperative Report 26:9-11.

Tester M and Langridge P. 2010. Breeding technologies to increase crop production in a changing world. Science 327: 818–822.

U.S. Department of Agriculture, Agricultural Research Service. 2009. USDA National Nutrient Database for Standard Reference, Release 22. Nutrient Data Laboratory Home Page, http://www.ars.usda.gov/ba/bhnrc/ndl.

Vinson JA, Su X, Zubik L, Bose P. 2001. Phenol antioxidant quantity and quality in foods: fruits. J Agric Food Chem 49:5315–5321.

Vorsa N, Howell AB, Foo LY and Lu Y. 2003. Food Factors in Health Promotion and Disease Prevention. ACS Sympo-

sium Series, Vol. 851, Chapter 26, pp.298–311.
Wong JC. 1999. Molecular marker mapping of chromosomal regions associated with levels of carotenoids and tocopherols in maize kernels. MS. Thesis. University of Illinois at Urbana-Champaign.
Woodruff BJ, Cantrell RP, Axtell JD, Butler LG. 1982. Tannin quantity in sorghum. Journal of Heredity 73:214-218.
Yu QB, Yang DQ, Zhang SY, Beaulieu J, Duchesne I. 2003. Genetic variation in decay resistance and its correlation to wood density and growth in white spruce. Can J For Res 33:2177–2183.
Zhang SY, Yu QB, Beaulieu J. 2004. Genetic variation in veneer quality and its correlation to growth in white spruce. Can J For Res 34:1311–1318.