2010 Production Research Report California Avocado Commission

Breeding

Linking Candidate Genes to Biochemical Phenotypes in Avocado

Michael T. Clegg
UC Irvine

Cooperators: Mary Durbin, Carlos Calderon, Livia Tommasini, and Vanessa Ashworth (UC Irvine)

Overview and Summary

This 3-year project, funded jointly by the UC Discovery Program and the CAC, is approaching the end of its final year. It was designed to identify genetic markers that track the nutritional composition of avocado fruit ("nutritional phenotypes") for implementation via marker-assisted selection. The discovery of candidate genes and development of SNP (Single Nucleotide Polymorphism) markers is complete, though additional markers are expected from transcription factors involved in the regulation of gene expression. Determination of nutritional composition, phenotypic heritability, and associations between nutritional phenotypes and SNP markers has progressed well. Heritability analyses suggest that proanthocyanidin (the precursor compound of anthocyanins), carotenoid and especially sitosterol contents have a genetic component, and selection for genes from their biosynthetic pathways is expected to produce rapid breeding advance. SNPs of genes from the anthocyanin, carotenoid and sitosterol pathways were found to be associated with elevated nutritional contents of the fruit; the association was very highly significant for a gene from the sitosterol pathway.

Assays of nutritional composition of the fruit ("nutritional phenotypes")

Nutritional assays on the fruit collected in 2009 are complete, and the contents of carotenoids and proanthocyanidins have been determined in fruit of the 2010 growing season (Table 1). Carotenoid contents averaged 7.93 and 11.8 µg/g fresh weight in 2009 and 2010, respectively (Figs. 1 and 2).

Table 1. Summary of the gene discovery process, including status of nutritional assays for fruit collected in 2009 and 2010, numbers of genes sequenced, number of SNPs detected, number of SNPs showing significant associations with nutritional phenotypes, and number of diagnostic SNPs (the minimum number needed to provide genotypic differentiation).

Biochemical Pathway from which genes were recruited	Status of Phenotypic assays	Number of candidate genes	Total number of SNPs	SNPs correlated with phenotype*	Number of diagnostic SNPs*
Vitamin B complex	Assay	11	23	tba	tba
Vitamin C	2009	8	14	tba	tba
Flavonoids, anthocyanins[1], phenylpropanoids	2009 & 2010	6	14	5	2

Isoprenoids/sitosterols	2009 & 2010	4	12	3	1
Vitamin E	Tba[2]	5	12	tba	tba
Carotenoids	2009 & 2010[3]	7	7	3	2
Fatty acids	-	5	2	tba	tba
Pulp browning	Tba	2	0	tba	tba
Totals		**48**	**84**	**11**	**5**

Abbreviations: Assay = streamlining assay methodology; Tba = data collection in progress. *Based on 2009 data; [1]New assay being tested; [2]Approximately 50% done for both years; [3]New assays completed for both years.

Determination of heritability following the statistical methods of quantitative genetics

Broad-sense heritability has been estimated for all available data using the SAS statistical software (Table 2). Estimates show that proanthocyanidin, carotenoid and especially sitosterol contents have a strong genetic component, and selection for these traits is thus expected to produce rapid breeding advance. Significant location effects were found for all measurements except dry weight. The genotype x environment interaction effect for sitosterol content was non-significant, suggesting that this nutrient is stable across environments.

Table 2. Broad-sense heritability values of traits measured in clonal replicates of Gwen x Fuerte avocado genotypes. Data in bold were not normally distributed and therefore transformed.

	H^2 (%)	G	E	G x E	Y	G x Y
Sitosterols[1]	**80.6**	*******	*****	**NS**	^	^
Pulp Wt (PW) / Fruit Wt (FW) %	63.0	***	**	***	**	**
Carotenoids[2]	47.2%[3]	***	***	**	***	NS
Fruit Weight (FW)	**49.2**	*******	*******	*****	******	**NS**
Dry Weight (DW)	35.4	***	NS	NS	***	NS
Pulp Weight (PW)	**35.3**	*******	*******	******	******	**NS**
Seed Weight (SW)	35.3	***	***	NS	***	NS
Proanthocyanidins	**30.7**	*******	******	*******	^	^
Vitamin C	16.2	**	**	**	^	^

Abbreviations: H^2 = heritability, G = genotype, E = environment (= location), 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 *, respectively. ^ = Assayed in 2009 samples only; [1]Data collection complete for 2010 but updated heritability values not yet available. [2]A new assay was used to re-determine values for both years of data; [3]Heritability for the two years separately are 61.4 and 69.3%.

Discovery of candidate genes and SNP marker development

The discovery of candidate genes from key biochemical pathways and identification of SNPs in their DNA sequences are complete, but we are seeking additional SNPs from regulatory genes using a microarray approach. Construction of a microarray that targets transcription factors (Tfs) was completed over the summer. TFs are genes that regulate the expression of other genes. Results from this microarray experiment, which is designed to reveal expression specifically of fruit ripening genes, are expected any time now, and TFs that correlate with the expression profiles of fruit ripening genes and nutritional phenotypes are a potential source of many additional SNP markers.

.

Looking for associations between nutritional phenotype and markers

We are running association analyses (general linear model implemented in the software TASSEL) to see whether SNPs are correlated with the content of nutrients in the fruit. Although correlation of a SNP is expected to be highest for nutrients from its source pathway, we do check for correlation with nutrients from the other pathways. Of particular promise are 3 SNPs in the 24-methylenecholesterol reductase gene of the sitosterol pathway that show a statistically highly significant correlation with sitosterol content. SNPs from the genes caffeoyl-CoA O-methyltransferase, phenyl alanine-ammonia lyase, phytoene synthase, and zeta-carotene desaturase show promise as well, but their associations with carotenoid and proanthocyanidin levels are of marginal statistical significance.

Linkage analysis

Linkage analysis is designed to infer the distribution of genetic markers across each of the 12 avocado chromosomes. Markers are assigned to linkage groups (which approximate the chromosomes) according to their pattern of segregation in a progeny array (here the Gwen x Fuerte progeny trees). Linked markers share the same segregation pattern and are assigned to the same linkage group. Our 130 markers (47 SSRs and 83 SNPs) locate to 16 linkage groups (Table 3); additional markers—from the Tfs and markers published by other research groups—will align the number of linkage groups with the number of chromosomes.

Table 3: Distribution of SSR (microsatellite) markers and SNPs on the linkage groups determined using the software OneMap.

Linkage Group	Number of SSR markers	Number of SNP markers	Number of distinct genes	Biochemical pathways*
1	9	1	1	FAP
2	3	6	2	Carotenoids, Vitamin E
3	1	4	1	Fatty acids
4	5	6	2	Carotenoids, Vitamin C
5	2	2	1	Fatty acids
6	4	3	1	Vitamin C
7	9	35	11	FAP, Pulp browning, Sitosterol, Vitamin B1, Vitamin B5, Vitamin B9, Vitamin C, Vitamin E
8	2	-	-	–
9	3	6	2	FAP, Carotenoids
10	1	1	1	Vitamin B6
11	1	2	1	Carotenoids
12	0	3	1	Sitosterol
13	0	3	1	Vitamin E
14	0	3	1	Vitamin E
15	0	4	1	Vitamin B6
16	0	3	1	Vitamin C
unlinked	7	1	1	FAP
	47	83	29	

*Abbreviations: FAP = flavonoid, anthocyanin & phenylpropanoid pathways

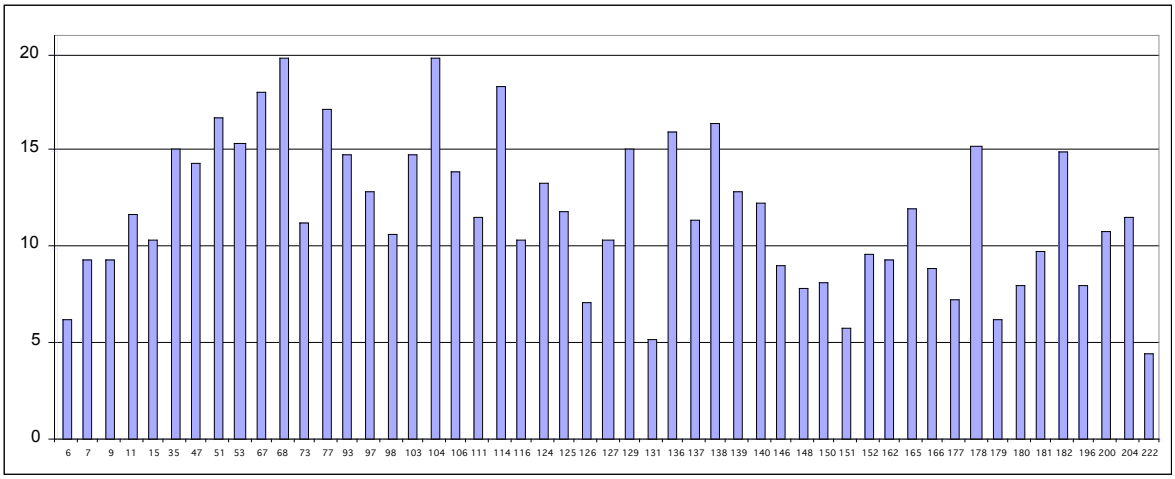

Fig. 1. Total carotenoid content (μg/g fresh weight) in the fruit collected in 2009. Genotype designations (Gwen x Fuerte progeny) appear on the x-axis.

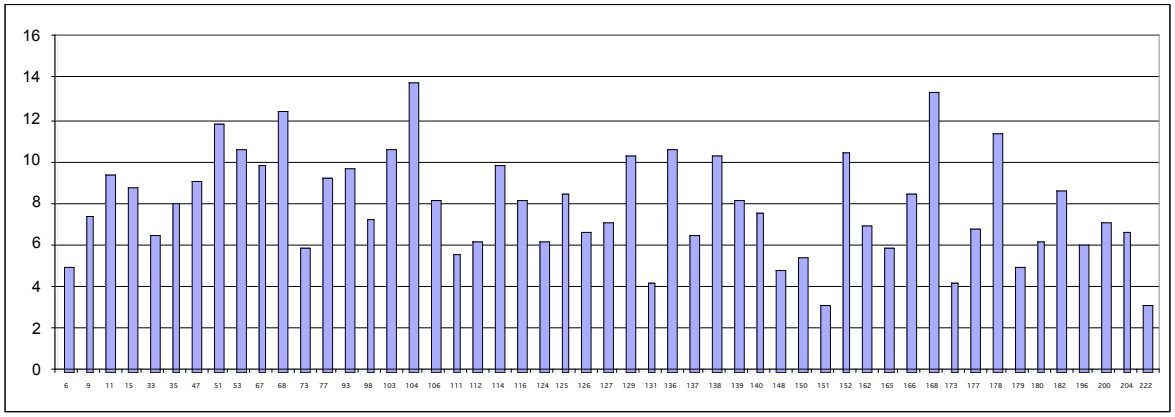

Fig. 2. Total carotenoid content (μg/g fresh weight) in the fruit collected in 2010. Genotype designations (Gwen x Fuerte progeny) appear on the x-axis.