Proceedings of the California Avocado Research Symposium, October 29, 2005. University of California, Riverside. Sponsored by the California Avocado Commission. Pages 38-43.

Assessing the Genetic Determination of Valuable Avocado Traits Using Microsatellite (SSR) Markers and Quantitative Trait Locus (QTL) Analysis

Ongoing Project: Year 4 of 8

Project Leader: Michael T. Clegg (949) 824-4490
e-mail: mclegg@uci.edu
Department of Ecology and Evolutionary Biology, UC Irvine, and
Department of Botany and Plant Sciences, UC Riverside

Supporting Staff; Haofeng Chen (UCR) and Vanessa Ashworth (UCI)
e-mail: Haofengchen@yahoo.com, vashwort@uci.edu

Benefits to the Industry

Modern genetic technologies have the power to greatly accelerate the pace of varietal improvement in avocado (*Persea americana* Mill.) and other tree crops that are slow to reach maturity. The first step is to identify molecular marker alleles that are associated with agriculturally interesting traits. Second, using a process known as marker-assisted selection (MAS), the presence of desirable marker alleles is assayed in seed or very early seedling stages. Presence of the desired alleles connotes the presence of desirable traits. The transmission of a commercially valuable trait can thus be predicted without the time, land resource, and labor costs associated with growing trees to the appropriate stage of development. Undesirable seedlings are immediately removed from the greenhouse bench or breeding block, while more promising genotypes are retained, thereby increasing the intensity of selection by orders of magnitude.

Objectives

Our research is laying the foundations for quantitative trait locus (QTL) analysis, a statistical procedure that reveals an association between productivity-related traits on the one hand and molecular marker alleles on the other. The method relies on a genetically characterized population grown in an experimental design with sufficient replication to eliminate all but the genetically determined components of the measurements. Only when a large amount of data from replicated field trials is available can an association between particular alleles and selected productivity-related (quantitatively inherited) traits be discerned.

Our research objectives are (1) to collect measurement data from our experimental trees, (2) assess the genetic determination of measurement traits using traditional quantitative genetic methods, (3) gather molecular marker (allele) data from each genotype of the experimental trees, (4) subject measurement and molecular data to QTL analysis, and (5) designate marker alleles suitable for the implementation of MAS.

Summary

Measurement data

During the 2004–2005 season we intensified our efforts to collect phenotypic measurement data from the experimental populations at South Coast Research and Extension Center (SCREC, Irvine) and Agricultural Operations (Ag Ops), UC Riverside. This focus was essential because the tree growth data are particularly important in the early years, and the measurement data must be collected at several points during each year to accurately characterize tree development. As the trees get larger their flowering and, most recently, fruiting characteristics are assuming greater importance in our data collection activities.

Tree growth data: Our cumulative data now span 2002–2005 for tree height, canopy diameter, and stem girth. In most cases we measured the trees once a year. Tree height and canopy diameter are measured between the months of April and October. For stem girth, we have four data sets, taken 6 months apart in 2002 and 2004, respectively. Owing to an uneven stem radius, each data point represents an average of two measurements taken in two opposite orientations (N–S and E–W). Table 1 summarizes our data collection details.

Table 1. Cumulative measurement data for 2002-2005, itemized by the number of individual readings that feed into the final measurement. Details on fruit data are preliminary (see text for additional information). Incipient fruit data applies to some 120 trees (at SCREC) that bore fruit in 2004 for the first time.

	# Years studied	Readings per trait	Readings (cumul.)	Tree #	Data points
Tree height	5	1	5	700	3,500
Canopy diameter	4	2	5	700	7,000
Trunk diameter	4	2	4	700	5,600
Flowering data	3	6	18	700	75,600
Fruit data	2	8+	tba	ca. 120	tba
				Total:	91,700

Fruit data: Since February of this year we have been monitoring trees showing precocious fruit bearing. At SCREC, 62 trees (out of a total of 393 = 15.8%) produced mature fruit, with fruit number ranging from 1 to 216 per tree. The most productive tree yielded 69.28 kg of fruit. At Ag Ops, none of the fruit set earlier in the season survived to maturity. In view of the later planting date of Ag Ops trees relative to SCREC trees, this difference is most likely a normal developmental effect, and time will tell whether site/climate-specific effects play an additional role. Fruit set has been prolific this spring at both locations: new fruit destined to reach maturity next year is present on 359 (91.3%) of trees at SCREC and on 211 (72.8%) of trees at Ag Ops. Significantly, trees at SCREC have set a greater number of fruit than last year.

For individual fruits, measurements have been taken of fruit weight, fruit length, diameter and circumference. In the more productive trees we randomly sampled 15 fruits and pooled the measurements. We also evaluated fruit quality in terms of fruit shape, seed shape, seed size, the seed: fruit size ratio, dry weight, flavor, and skin thickness. Salient fruit data are summarized in Table 2.

Table 2. Summary of fruit data for 2005.

	Fruit number per tree	Fruit production per tree (kg)	Fruit weight (g)	Fruit diameter (mm)	Fruit length (mm)	Dry weight (%)	Seed : fruit weight ratio
Average	53.3±58.4	17.17±19.3	347±128.7	78.0±10.1	105.8±39.6	32.4±4.8	0.213±0.049
Max	216	69.28	690.6	101.5	190	45.4	0.321
Min	1	0.26	80.3	48.8	48.1	22.9	0.05

Flowering data: Flowering data have been collected for 2003, 2004 and 2005, with weekly or bimonthly measurements that aim to capture the entire flowering season until fruit set is initiated. We record discrete categories for bloom stage, bloom intensity, bloom open percentage, bloom drop percentage, and fruit crop intensity. This year, data collection began at the beginning of February and ended in mid-May.

Statistical analyses

Two charts depicting analyses of heritability for various measurement data are shown in Figures 1 and 2, respectively. Figure 1 illustrates the need for collecting data over several seasons: an unusually wet winter may account for the decline in the heritability value for tree height in the 2004/5 season at Irvine relative to the two prior seasons, although other factors may be at play. Figure 2 shows a heritability of over 0.45 for early fruit set for SCREC trees for 2004.

Molecular marker data

The collection of molecular marker data from our experimental trees (progeny of variety Gwen) resumed in mid-January. We started by analyzing the DNA from the trees growing at Ag Ops. These consist of two of the four clonal replicates per genotype. As data accumulated it soon became clear that our experimental population represented not selfed—but almost exclusively outcrossed—progeny, presumably reflecting the origin of the trees: all had been grown from fruit of a single Gwen tree located in an experimental field at Ag Ops or SCREC. This tree was part of the Dr. B. Bergh/Gray Martin variety selection program. Research plots/breeding blocks differ from commercial plantings in that there is often a greater diversity of tree varieties in the immediate vicinity, whereas commercial groves contain a (Hass) monoculture with occasional

interplanted rows of pollinizer trees. The high level of outcrossing does illustrate that given sufficient pollinators in close proximity, most fruit arise from outcrossing.

At this time we have determined the allelic composition of our experimental trees at 7 marker loci (AUCR418, AVO102, AVD013, AVD089, AVT005b, AVT226, and AVT436). The primary pollen donors that sired the Gwen progeny appear to be Fuerte (23.4%), Bacon (22.3%), and Zutano (14.9%), with other varieties represented in small quantities. In some cases, the pollen source has been narrowed down to a small panel of 2–5 potential varieties, and further loci will be necessary to pinpoint the actual pollen donor (see Table 3). Some 12.6% of genotypes do not trace to any of our 37 archived pollen donors.

For QTL analysis, our outcrossed progeny will be grouped by paternal genotype. Outcrossed parentage does not affect the identification of trait-marker associations that underpin MAS. Indeed, the presence of abundant outcrossed progeny gives us the valuable opportunity to assess the contribution of the male and female parent separately and to learn more about the impact of different pollen sources in a Gwen maternal genetic background. Moreover, our pool of genetic markers has increased, as we do not need to confine our analyses to those loci for which Gwen is heterozygous.

Conclusion

This past year has seen tremendous progress in the first three of our five research objectives. Sufficient measurement data have now been accumulated to soon permit a first pass of QTL analysis. The molecular data have been accumulating since mid-January, but much work remains to be done in this area. A priority will be to extend the number of markers assayed against each Gwen progeny genotype and to keep refining the inferred pollen source in instances of ambiguity.

Table 3. Allelic composition of 30 progeny genotypes at 7 molecular marker loci, combined with measurement data (fruit weight per tree and fruit number per tree in 2004, the first year of fruit bearing at SCREC). The table is sorted by fruit weight. Only one of the two clonal replicates bore fruit, except genotypes 121 and 126, for which fruit weight and number have been averaged. Abbreviations in the final column: PD = inferred pollen donor; B = Bacon; Z = Zutano; F = Fuerte, H = Hass; P. sch = *Persea schiedeana*; * = the assigned variety is one (the most likely) of several potential pollen donors.

Genotype	Fruit wt (g)	Fruit #	AVO102	AVD013	AVD089	AVT436	AVT005b	AUCR418	AVT226	PD
121.1	650.14	7	h/g	d/d	a/b	d/b	c/d	h/g	b/b+3	B
185.3	586.2	5	h/b	a/d	a/c	d/b	c/d	h/e	e/b	Z*
11.3	535.35	8	g/e	d/f	a/a	d/b	c/a	h/e	b/a	F
78.4	503.4	8	h/b	d/d	a/b	c/a	c/c	h/d	c/b	B
126.2	498.82	5.5	I/b	d-2/d	b/b	b/a	c/a	h/h		F
23.2	472.46	21	g/b	a/d	a/c	c/a	c/a	h/h	c/b	Z
184.1	452.71	30	g/g	d/d	b/b	c/b	c/c	h/h	c/b	Z*
49.4	442.75	2	k/g	d/d	a/c	d/a	c/c	h/d	d/c	P. sch
112.1	435.85	3	I/g	d/f	a/a	d/a	c/a	h/h	c/a	F*
58.3	409.15	2	h/b	a/d	a/b	c/a	c/c	g/e	d/b	B
76.2	407.1	32	g/b	d/d	a/c	c/b	c/c	h/g	d/c	B
41.2	402.5	1	b/a	a/d	b/b	c/a	c/a	h/e-2	b+3/b+3	?
170.2	389.6	2	h/b	d/d	a/a	d/b	c/a	h/g	c/b	B
149.4	383.4	1	g/g	a/d	a/c	b/b	a/a	h/e	b/a	?
197.2	381.18	12	h/g	d-2/d	a/a	d/a	c/a	h/h	e/c	Z*
178.2	363.6	1	e/b	d/f	a/a	d/b	c/a	h/e	c/a	F
43.4	338.95	2	g/g	d/d	a/c	c/b	c/c	h/g	c/b	B
164.2	336	6	I/g	a/d	a/c	b/b	c/c	h/e-2	c/b	H
81.1	332.5	6	g/g	a/a	a/c	c/b	a/a	e/c	b/a	?
156.2	330.92	17	g/b	d-2/d	a/c	d/a		h/c	d/c	Z
92.3	322.1	1	h/b	a/d	b/b	b/a	a/d	h/c	e/b	?
29.3	310.15	2	h/b	a/d	b/b	c/b	c/a	h/h	c/c	Z
116.2	286.2	8	e/b	d-2/d	a/b	d/a	c/c	h/h	b/b	F
198.4	284.71	8	g/b	d/d	a/c	c/b	c/c	g/e	c/b	B
153.1	278.83	3	b/b	d/f	a/c+2	b/a	c/c	h/e	c/b	H
86.2	264.7	1	h/g	d-2/d	a/a	d/a	c/d		b/b	?
115.4	248.55	2	g/g	d-2/d	b/b	d/b	c/c	h/c	b/b+3	Z
24.1	240.2	3	h/b	a/d-2	a/b	c/b	c/c	e/c	e/b	Z
163.1	227.73	3	b/b	a/d	b/b	b/b	c/a	h/e-2	c/b	H*
5.2	208.81	6	g/b	a/d	b/b	d/a	c/a	h/e	d/b	Z*

Fig. 1 show heritability values for trees at Ag Ops, Riverside, and blue bars for trees at SCREC, Irvine, calculated for three consecutive seasons (2002–2003, 2003–2004, and 2004–2005 from left to right).

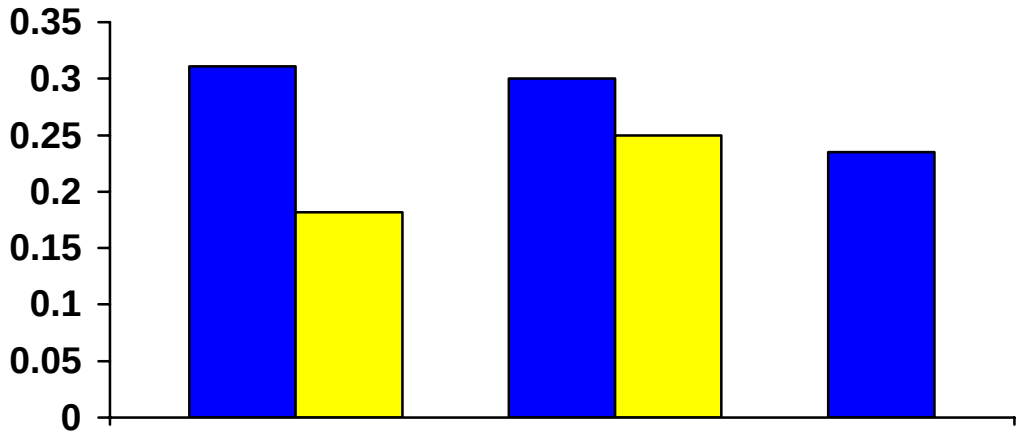

Fig. 1. Heritability of tree height over time (random model). Yellow bars show heritability values for trees at Ag Ops, Riverside, and blue bars for trees at SCREC, Irvine, calculated for three consecutive seasons (2002–2003, 2003–2004, and 2004–2005 from left to right).

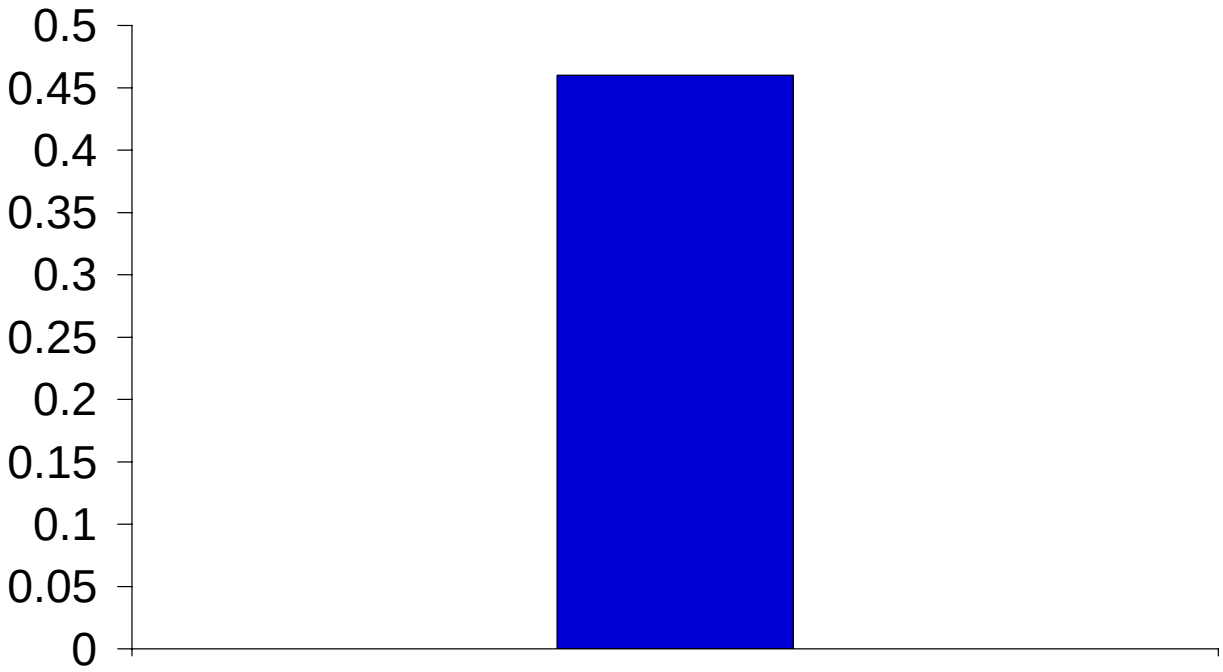

Fig. 2. Heritability of precocious fruit set at SCREC, Irvine, in 2004 (first year of data).