Quantitative Trait Locus Analysis in Avocado: The Challenge of a Slow-maturing Horticultural Tree Crop

Vanessa E.T.M. Ashworth5
Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

Haofeng Chen1 and Carlos L. Calderón-Vázquez2
Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Mary Lu Arpaia
Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

David N. Kuhn3
USDA-ARS Subtropical Horticulture Research Station, Miami, FL 33158

Mary L. Durbin3 and Livia Tommasini4
Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Elizabeth Deyett and Zhenyu Jia
Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

Michael T. Clegg3
Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Philippe E. Rolshausen5
Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

ADDITIONAL INDEX WORDS. Persea americana, β-sitosterol, candidate gene, linkage map, marker-assisted selection, vitamin E

ABSTRACT. The glossy, green-fleshed fruit of the avocado (Persea americana) has been the object of human selection for thousands of years. Recent interest in healthy nutrition has singled out the avocado as an excellent source of several phytonutrients. Yet as a sizeable, slow-maturing tree crop, it has been largely neglected by genetic studies, owing to a long breeding cycle and costly field trials. We use a small, replicated experimental population of 50 progeny, grown at two locations in two successive years, to explore the feasibility of developing a dense genetic linkage map and to implement quantitative trait locus (QTL) analysis for seven phenotypic traits. Additionally, we test the utility of candidate-gene single-nucleotide polymorphisms developed to genes from biosynthetic pathways of phytonutrients beneficial to human health. The resulting linkage map consisted of 1346 markers (1044.7 cM) distributed across 12 linkage groups. Numerous markers on Linkage Group 10 were associated with a QTL for flowering type. One marker on Linkage Group 1 tracked a QTL for β-sitosterol content of the fruit. A region on Linkage Group 3 tracked vitamin E (α-tocopherol) content of the fruit, and several markers were stable across both locations and study years. We argue that the pursuit of linkage mapping and QTL analysis is worthwhile, even when population size is small.
conferred by three main groups of compounds: β-sitosterol, carotenoids, and vitamin E.

Most of these phenotypic traits are inherited in a quantitative fashion; i.e., they are controlled by many genes of small effect and are typically under strong environmental influence. Yet only the genetic component of a phenotype will respond to breeding. Two studies in avocado (Calderón-Vázquez et al., 2013; Chen et al., 2007) used quantitative genetics to tease apart the genetic and the environmental components of the phenotypic value of a suite of quantitative traits. Chen et al. (2007) demonstrated for the progeny of cultivar Gwen that major growth-related traits, such as plant height and trunk- and canopy diameter, were under genetic control and showed sufficient heritability to respond to selection. Similarly, Calderón-Vázquez et al. (2013) showed for a ‘Gwen’ × ‘Fuerte’ experimental population—a subset of the population studied by Chen et al. (2007)—that β-sitosterol, carotenoids, and vitamin E of the fruit are likely to respond to breeding.

On theoretical grounds, therefore, breeding avocado for growth-related traits and enhanced levels of fruit nutrients is feasible. However, breeding in this long-lived tree crop is frustrated by an outcrossing breeding system, high heterozygosity, long generation times (up to 15 years [Bergh and Lahav, 1996]), and the need for costly field trials to accommodate tree size and a protracted maturation (Van Nocker and Gardiner, 2014). Moreover, controlled pollination is impracticable (Degani et al., 2003; Lammerts, 1942) owing to a profusion of tiny flowers and immature fruitlets—most of which are shed prematurely, and conventional breeding populations (e.g., doubled haploids, recombinant inbred lines) do not exist. At this time, avocado breeders have no option but to use phenotypic selection, which is associated with slow breeding advance. A move toward molecular breeding is a promising alternative to accelerate selection progress and to reduce costs associated with the maintenance of breeding populations.

When designing large-scale experiments leading to molecular breeding, the problem of high land and labor costs loom large, so genetic mapping populations tend to be small and poorly replicated, predisposing data to low statistical power. Yet many horticultural tree crops produce high-value fruit for which the genetic dissection of phenotypic traits is of considerable interest, raising the question whether mapping and quantitative trait locus studies may nonetheless be worthwhile, given adequate precautions. With the advent of next-generation technologies, the costs associated with developing abundant genetic markers have declined significantly, and a shortage of markers no longer represents a constraint. We explore the possibility of generating a linkage map and of estimating QTLs for seven phenotypic traits collected in a mapping population of 50 trees using over 5000 molecular markers. We ask whether a modestly sized mapping population can be used to estimate significant QTL loci and whether these loci are likely to be sufficiently robust.

Materials and Methods

Mapping Population. The experimental population of avocado trees consisted of the full-sib progeny of a ‘Gwen’ (G) × ‘Fuerte’ (F) cross. The G × F progeny is a subset of a larger population of open-pollinated trees raised from the fruit of a ‘Gwen’ maternal tree. Each progeny tree was screened using 10 simple sequence repeat (SSR) markers (Ashworth et al., 2004) to verify the origin of the pollen source. Of more than 200 progeny genotypes analyzed, 50 were the result of the cross G × F and were set aside for the mapping project. The remainder consisted of about 50 individuals each of G × ‘Bacon’, G × ‘Zutano’, and a miscellaneous group of largely unidentified pollen origin (Chen et al., 2007) that are not considered further here.

Four clonal replicates of each G × F progeny tree were grafted on ‘Duke 7’ rootstock and planted at two sites in southern California: two of the four replicate trees were grown in a randomized block design at a coastal location [University of California (UC) South Coast Research and Extension Center, Irvine, CA] and the other two replicate trees at an inland location (Agricultural Operations, UC Riverside campus, Riverside, CA), also in a randomized block layout. Each location, therefore, contained two replicates of 50 tree genotypes (100 trees). All trees were planted in the ground between Fall 2001 and Spring 2003.

Trees were spaced at 6.1 m between rows and at 4.6 m between trees within the same row. At the coastal site, fertilizer was applied at 0.45 kg/tree as a granular formulation of 15N–6.5P–12.5K in late March/early April. At the inland site, a 32N–0P–0K fertilizer solution was introduced into the irrigation water at 284.24 L ha⁻¹ in January. At both locations, the fertilizer regime was managed to industry standard. Irrigation water was dispensed from two microsprinklers per tree following guidelines established by California Irrigation Management Information System (CIMIS, 2003). The coastal location (Irvine) differed from the inland location (Riverside) by higher average rainfall, cooler average summer temperatures, and warmer average winter temperatures (Table 1). Soils at both locations were sandy loams. The Riverside site followed a gentle hillside contour that consisted of three different sandy loam subtypes (Table 1).

Phenotypic Traits. Seven datasets were collected from the experimental trees, including one qualitative (flowering type) and six quantitative (three measures of tree dimension, and three nutrients assayed in the avocado fruit flesh). Descriptive statistics for each quantitative trait are provided in Fig. 1.

Flowering type was recorded in Apr. 2013 at the coastal location in 100 trees. Avocado flowers exhibit protogynous dichogamy, a mechanism designed to prevent self-pollination by temporally separating stigma receptivity and pollen release (Sedgley, 1985). A tree was recorded as having B-type flowering if its flowers were in the male phase in the morning and as having A-type flowering if flowers were in the female phase in the morning. In commercial orchards, optimal pollination and fruit set in cultivars with A-type flowering (e.g., ‘Hass’ and ‘Gwen’) is achieved by interplanting with B-type pollinator cultivars (e.g., ‘Fuerte’ and ‘Bacon’) (Alcaraz and Hormaza, 2009). This trait was scored as a discrete character (presence or absence), with A-type flowering recorded as “1” and B-type flowering as “2.”

Measures of tree growth were collected at both locations each year from 2003 to 2005, but only the final year’s data were used in this study because the later-planted trees were still very immature during the first two years. Three measurements of tree dimension—trunk diameter, tree height, and canopy diameter—were recorded as a way of characterizing the three-dimensional aspect of early tree growth (Chen et al., 2007). Trunk diameter was determined at 10 cm aboveground in two perpendicular orientations, with values averaged. Plant height
was measured from ground level to the tip of the tree. Canopy diameter was determined at the widest part of the canopy in two orientations: parallel to the orchard row and perpendicular to the row, with the two values averaged.

Fruit nutrient composition [α-tocopherol (the most biologically active form of vitamin E in humans), β-sitosterol, and carotenoids] was assayed in fruit collected at both locations in 2009 and 2010. Fruit preparation and chemical assays for determination of the contents of α-tocopherol, β-sitosterol, and carotenoids in fruit tissue were adapted from Jeong and Lachance (2001), Mäeorg et al. (2007), and Ryan et al. (2007) and are detailed in Calderón-Vázquez et al. (2013). For any given tree, five fruit were picked at an optimum dry weight of 20% and then allowed to ripen in the laboratory. At ripeness, the flesh from the five fruit was pooled and homogenized, and aliquots were frozen and set aside for further analyses. Total carotenoids, which include α-carotene, β-carotene, β-cryptoxanthin, lutein, and zeaxanthin, were isolated using two extractions in hexane/petroleum ether (1:1). An aliquot of the resulting aqueous phase was analyzed by taking a spectrophotometric reading at 456 nm and comparing it to a standard curve for β-carotene (C4582; Sigma-Aldrich, St. Louis, MO) according to Luterotti et al. (2006). Beta-sitosterol and α-tocopherol contents were determined by application of the organic phase fraction to thin-layer chromatographic plates. Bands were visualized by dipping in phosphomolybdic acid (02553, Sigma-Aldrich) and quantified on an AlphaImager HP System (ProteinSimple, Santa Clara, CA) using standard curves generated from reference samples [β-sitosterol (S1270, Sigma-Aldrich), α-tocopherol (T3251, Sigma-Aldrich)]. Values for the parental cultivars Gwen and Fuerte were determined in trees growing at the coastal location using the same preparation and assay conditions as for the progeny (Calderón-Vázquez et al., 2013).

Statistical analyses of the phenotypic data were performed in R version 3.4.4 (R Core Team, 2019) using a nonparametric Kruskal–Wallis test to compare datasets, followed by a Wilcoxon test for pairwise comparisons and calculation of probability values.

**GENETIC MARKERS.** The genetic markers implemented in this study consisted of SSRs and single-nucleotide polymorphisms (SNPs) from several sources; the bulk of markers were SNPs developed by Kuhn et al. (2019). In our map, these SNPs were used to augment the total number of markers to ensure adequate map density. The second set of SNP markers was developed in a genetic discovery effort targeting candidate genes from several biosynthetic pathways involved in fruit nutrient composition. These candidate-gene SNPs (CG-SNPs) have not previously been published and their development is described in the following two paragraphs. In addition, we used published SSR markers developed by Sharon et al. (1997), Borrone et al. (2007), and Ashworth et al. (2004), as well as 28 SSR markers available from GenBank (V.E. Ashworth, C. Calderón-Vázquez, M.L. Durbin, L. Tommasini, and M.T. Clegg, unpublished data).

SNPs by Kuhn et al. [2019 (FL-SNPs)] originated by Illumina GAII sequencing (Illumina, San Diego, CA), and the individuals of our ‘Gwen’ × ‘Fuerte’ mapping population were included on the Illumina Infinium oligonucleotide array chip that assayed each tree genotype for 5050 FL-SNP markers. Details of marker development are provided in Kuhn et al. (2019).

Nutrient-related candidate genes were identified by aligning avocado expressed sequence tag (EST)/cDNA (complementary DNA) sequences from fruit-, flower-, and other organ-specific libraries developed by Cornell University [Ithaca, NY (Floral Genome Project, 2005), HortResearch (Mt Albert, New Zealand), and CINVESTAV (Irapuato, Mexico) to sequences of functionally characterized gene sequences deposited in TAIR (2005) or NCBI (2005). Avocado mRNA sequences showing high similarity to core enzymes in the flavonoid, carotenoid, fatty acid, and B-, C-, and E-vitamin biosynthesis pathways were retained. Their relevance in determining fruit nutritional composition was further verified by comparison with sequences from an avocado cDNA library developed from the fruit of cultivar Hass. Sequence alignment allowed design of amplification primers in conserved regions. Nested sequencing primers provided about 500 base pairs of high-quality DNA sequence.

SNP discovery was performed in sequences from a panel of 10 randomly chosen ‘Gwen’ × ‘Fuerte’ progeny genotypes. SNPs were identified by standard resequencing using the Sanger method. Sequence reads were assembled using Phred/Phrap/Consed (Ewing and Green, 1998; Gordon et al., 1998), and PolyPhred was used to detect the SNP sites (Nickerson et al., 1997). A total of 83 SNPs was developed from 28 candidate genes. Avocado genomic DNA of the 10 ‘Gwen’ × ‘Fuerte’ progeny was extracted from frozen young (flushing) leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Forward and reverse reads were generated during the sequencing phase. Sequences from the SNP phase were

---

Table 1. Climatic characteristics at Irvine and Riverside, CA, the two locations of the avocado mapping populations. Data are averages for 1981–2010 (U.S. Climate Data, 2018).

| Climate                        | Irvine   | Riverside |
|--------------------------------|----------|-----------|
| Annual high temperature (°C)   | 22.6     | 26.4      |
| Highest monthly average temperature—August (°C) | 28.3     | 35.0      |
| Annual low temperature (°C)    | 12.4     | 10.8      |
| Lowest monthly average temperature—December (°C) | 8.3      | 5.6       |
| Average temperature (°C)       | 17.5     | 18.6      |
| Average annual precipitation (mm) | 366.7    | 262.1    |
| Soil type                      | San Emigdio fine sandy loam | Arlington fine sandy loam; Hanford coarse sandy loam; Ramona sandy loam |
sequenced only in one direction (either 5' to 3' or 3' to 5'), either using polymerase chain reaction (PCR) amplification primers or nested primers (Supplemental Table 1). PCR amplification conditions were as follows: preheating at 94 °C for 2 min, then 35 cycles of 94 °C for 30 s, annealing at primer-specific temperatures (47 to 58 °C) for 30 s and extension at 72 °C for between 30 s and 1 min 45 s, ending with a final extension of 72 °C for 5 min. PCR products were purified using the QiAquick PCR purification kit (Qiagen) or ExoSAP-IT (USB-Affymetrix, Cleveland, OH). Sequencing products were run on a DNA sequencer (Applied Biosystems 3730xl DNA Analyzer; Thermo Fisher Scientific, Waltham, MA).

SSR markers included 53 published markers. They were sourced from Sharon et al. (1997; 1 marker), Borrone et al. (2007; 13 markers), and Ashworth et al. (2004; 39 markers). Twenty-eight new SSR markers are detailed in Supplemental Table 2; their development and assay conditions are identical to those given in Ashworth et al. (2004). SSR markers of Borrone et al. (2007) were developed from ESTs.

SSR markers originating at UC (CA-SSRs) were prefixed with AVO, AUCR, or AVD if developed from a genomic library enriched for dinucleotide repeats; a prefix of AVT denotes development from a trinucleotide-enriched genomic library (Ashworth et al., 2004). SSR markers developed by Borrone et al. (2007; FL-SSRs) are prefixed with SHRSPa (Subtropical Horticulture Research Station—Persea americana) followed by a three-digit number. AVMIX3 originated from Sharon et al. (1997). CG-SNPs are abbreviated in relation to the candidate gene name and numbered sequentially based on the SNP position within the gene sequence. The FL-SNPs (Kuhn et al., 2019) are prefixed by SHRSPaS00, followed by SNP numbers in the range 1000 to 6999. All CG-SNPs from the same candidate gene were retained unless a SNP showed strong segregation distortion or many missing data.

**LINKAGE MAP CONSTRUCTION.** Our linkage map [henceforth “California (CA)-map”] was generated using the regression mapping algorithm implemented in JoinMap version 4 (Van Ooijen, 2006) that allows analysis of a mixed set of marker types and segregation patterns. Population type was set to cross pollination (CP). We used regression mapping combined with the Kosambi function of transforming recombination frequencies into map units (centiMorgans). A log-odds (LOD) value of 5.0 was used for linkage group selection. MapChart version 2 (Voorrips, 2002) enabled markers to be graphically represented on their corresponding linkage group (LGs) based on the map distances determined via linkage analysis.

The chi-squared test implemented in Joinmap (Van Ooijen, 2006) was used to examine each marker for segregation distortion. Although distorted markers can be the cause of Type 1 Error (detecting false linkage), only markers with values of 8 or higher were pruned from the dataset, as modest amounts of segregation distortion are thought to contribute pertinent information (Hackett and Broadfoot, 2003; Wang et al., 2005).

**Fig. 1.** Variance statistics for six quantitative traits determined in an avocado mapping population growing at two locations in southern California [South Coast Research & Extension Center in Irvine, CA (SC) and Agricultural Operations of the University of California at Riverside (UCR)]. Dots represent samples, bars show means and SE. Numbers above brackets are probability values (no brackets are shown for $P > 0.05$).
To explore whether missing data may be affecting marker distribution and distances when working with small mapping populations, we developed a second map from which all markers with missing data had been removed. Additionally, we compared the CA-map to a high-density map integrated from four reciprocal mapping populations [514 progeny of ‘Tonnage’ × ‘Simmonds’, 249 of ‘Simmonds’ × ‘Tonnage’, 346 of ‘Hass’ × ‘Bacon’, and 230 of ‘Bacon’ × ‘Hass’; henceforth “FL-map” (Rendón-Anaya et al., 2019)] that included the same set of 5050 next-generation SNPs (Kuhn et al., 2019). The comparison was made using the VLOOKUP function in Excel (version 16.16.1; Microsoft, Redmond, WA) to check for marker distribution across and within linkage groups for markers common to both maps.

**QTL analysis.** QTL analysis was performed using both interval mapping [IM (Lander and Botstein, 1989)] and non-parametric mapping [Kruskal–Wallis (KW) test; Kruskal and Wallis (1952)] implemented in MapQTL version 5 (Van Ooijen, 2004). Under IM, QTL significance was assigned to a marker locus in relation to the LOD likelihood scores determined using 1000 permutations of the data at a significance level of $P = 0.05$. In the maximum likelihood mixture model of IM, where LOD scores are calculated using an iterative algorithm, an iteration number of 20 was used as a cut-off to declare a significant QTL, with values above 20 representing a poor fit of the data to the model (Van Ooijen, 2004). Markers exceeding the cutoff of 20 for iteration number were disregarded. The KW test evaluates each marker independently regardless of its location on the linkage map. It is recommended for data that are not normally distributed, such as qualitative data, counts, data with outliers, and truncated data probabilities (Kruglyak and Lander, 1995), and it assigns significance in relation to the test statistic $K^*$, with a value of $P \geq 0.005$ (denoted as **** in MapQTL) considered sufficiently stringent to declare a marker as being significantly associated with a QTL.

To verify significant QTLs, we performed an approximation of the multiple-QTL model (MQM) by manually selecting markers located close to a QTL as cofactors. The MQM model is more accurate and efficient at detecting QTLs than IM because the latter ignores the effects of other QTLs, but MQM suffers from being computationally intensive. A work-around was developed by Jansen (1993) and is implemented in MapQTL in the “rMQM” module. However, owing to the small population size and heterogeneously heterozygous population type (“CP” in MapQTL) of this dataset, we were not able to take advantage of the Automatic Cofactor Selection analysis available in MapQTL to perform backward elimination because it uses many degrees of freedom (df) and is computationally too demanding. Instead, we manually chose cofactors guided by the output from IM, sequentially selecting markers closest to a significant QTL and running rMQM. QTLs were retained if successive exclusion of cofactors did not alter the LOD values associated with the QTL.

Where multiple datasets were available, MapQTL analyses were performed for each location (coastal or inland) separately in the case of the growth-related traits (trunk diameter, plant height, and canopy diameter), as previous studies had shown significant location effects (Chen et al., 2007). For fruit nutrient content, analyses were also run on separate datasets (2 years and two locations) because Calderón-Vázquez et al. (2013) had demonstrated significant effects of harvest year on the contents of two of the three nutrients and a significant location effect on carotenoid contents, as well as interaction effects for genotype × environment (β-sitosterol and carotenoids) and genotype × year (β-sitosterol). Flowering was analyzed for a single year at the coastal location.

In all cases, we examined the output from both IM and the non-parametric KW test to declare significant QTLs, emphasizing those markers that were endorsed by both algorithms. Consideration of both the IM and KW output was deemed prudent (Kruglyak and Lander, 1995), given that the small population size ($n = 50$) may have affected the accuracy or power of the algorithms.

**Results.**

**Phenotypic traits.** Plots showing the distribution of tree measurements at both locations and of the fruit nutrient data at all four location/year combinations are presented in Fig. 1. Trees were consistently somewhat shorter at Riverside than at Irvine, averaging 1.97 ± 0.466 and 3.19 ± 0.639 m, respectively. Trees at Riverside also developed smaller canopies (2.1 ± 0.368 and 3.55 ± 0.691 m, respectively) and trunk diameters (75.13 ± 13.1 and 95.53 ± 18.0 mm, respectively).

Values of the three fruit nutrients responded differently depending on environment and year; α-tocopherol values were not significantly different for either year or location. Beta-sitosterol values were significantly different between years at the Riverside location, with higher values occurring in 2010. Differences between years at the Irvine location were not significant. Carotenoid contents were significantly different for all location/year comparisons, with values significantly higher at Riverside than at Irvine and significantly higher in 2010 than in 2009.

One genotype consistently produced fruit with the highest α-tocopherol concentrations at Irvine in both years and at Riverside in 2010 but failed to produce any fruit at Riverside in 2009, leading to a missing data point. The same genotype was also responsible for the highest β-sitosterol values at Irvine and Riverside in 2010 and the second-highest value in Irvine in 2009. In both years, almost half the progeny in Irvine exceeded α-tocopherol contents measured in the parental cultivars [19.5 and 19.0 μg·g⁻¹ FW] in ‘Gwen’ and ‘Fuerte’, respectively]. Two genotypes exceeded the value of their maternal parent more than 2-fold. Progeny values varied more than 6-fold (2009) and 8-fold (2010) at Irvine and more than 5-fold (2009) and 6-fold (2010) at UCR.

For β-sitosterol, values of the male parent (672 μg·g⁻¹ FW) consistently exceeded values in the progeny; but seven and five progeny genotypes, respectively, exceeded the value in ‘Gwen’ (469 μg·g⁻¹ FW) in 2009 and 2010. Progeny values varied more than 5-fold (2009) and 4-fold (2010) at Irvine and more than 7-fold (2009) and 4-fold (2010) at UCR.

Carotenoid contents were higher in ‘Fuerte’ (9.8 μg·g⁻¹ FW) than in ‘Gwen’ (8.37 μg·g⁻¹ FW). In 2009 and 2010, eight and 27 progeny genotypes, respectively, exceeded ‘Fuerte’ values. Values in the progeny varied 4-fold (2009) and 3-fold (2010) in Irvine and 3-fold (2009) and almost 4-fold (2010) at UCR.

Flowering type was determined at Irvine for 47 genotypes for which two replicate trees were available, 31 genotypes showing B-type flowering (as in ‘Fuerte’), and 16 showing A-type flowering (as in ‘Gwen’). All replicate pairs showed the same flowering type.
**Linkage Mapping.** We pre-screened 5050 FL-SNPs developed by Kuhn et al. (2019) to eliminate markers that were invariant or uninformative in the parental genotypes ‘Gwen’ and ‘Fuerte’. The remaining FL-SNP markers (2608) were then combined with 146 informative SNP and SSR markers; 83 SNPs developed to eight candidate genes of nutritional pathways and 63 SSR markers. In total, 2754 markers were imported into a JoinMap version 4.0 (Van Ooijen, 2006) data matrix for linkage mapping, of which 1346 markers (49%) placed on 12 linkage groups at a LOD value of 5.0, constituting the CA-map (Supplemental Fig. 1).

A total of 1399 markers were eliminated because of identical segregation or because of strong segregation distortion (38 markers with $\chi^2 = 8.00-31.04$, $P = 0.01-0.0000001$, df = 1–3). The placed markers consisted of 1235 FL-SNPs (91.8%), 58 CG-SNPs (4.3%), and 53 SSR markers [AVMIX3, 13 FL-SSRs, and 39 CA-SSRs (3.9%)]. Of the 1346 markers on the map, 616 (45.8%) were heterozygous in both parents, of which six segregated with four alleles (SSRs), 20 with three alleles (SSRs), and 590 with two alleles (SNPs and SSRs). Markers segregating in only one of the parents (730; 54.2%) numbered 309 in ‘Gwen’ and 421 in ‘Fuerte’.

Marker number per linkage group averaged 112, ranging from 56 loci (LG12) to 207 loci (LG2). Combined linkage group length was 1044.7 cM, ranging from 61.483 cM on LG2 to 121.125 cM on LG3, and averaging 87.06 ± 19.77 cM/linkage group. The mean number of loci/cM was 1.32. Gaps larger than 5 cM occurred on four linkage groups. The densest linkage group was LG2 (3.37 loci/cM). Sparse coverage characterized distal portions of LG7 (Supplemental Fig. 2). Supplemental Table 3 shows marker order on the 12 avocado linkage groups obtained in this study.

An exploratory map made up exclusively of markers containing no missing data closely resembled the CA-map. Also using a LOD value of 5.0 to assign markers to linkage groups, this map contained 1238 markers on 12 linkage groups with a combined length of 1036.3 cM. Linkage groups averaged 103 loci and 86.35 ± 27.44 cM. Of the 1238 placed markers, one SSR marker segregated with four alleles, four SSRs segregated with three alleles, 555 were of JoinMap segregation type hh×hh, 289 of type lm×ll, and 389 of type nn×np.

Comparison of the CA-map with the highly saturated FL-map (Rendón-Anaya et al., 2019) showed excellent agreement between the two maps, as markers common to both maps were assigned to the same linkage group and marker order was comparable (Supplemental Fig. 2). Although a few linkage groups showed inverted segments (Supplemental Fig. 2), we did not adopt the FL-map marker order. FL-map linkage groups contained ≥2.0 to 3.3 times as many marker loci as their CA-map counterparts. Overall, the number of loci on the FL-map was about 2.6 times greater than that on the CA-map, and total linkage group length (cM) of the FL-map was 1.73 times greater. The average marker density for the FL- and CA-maps was 1.97 and 1.32 markers/cM, respectively.

Of the 58 CG-SNPs assigned to a linkage group, the greatest number (13 SNPs; 22.4%) mapped to LG2. SNPs of the same candidate gene always mapped to the same linkage group. In most cases SNPs from the same candidate gene mapped in close proximity. Exceptions were the SNPs of CUT1 (12.569 cM apart), MEP (8.119 cM apart), PSY (6.731 cM apart), and VTE1_687 (6.015 cM from the nearest SNP, VTE1_573).

**QTL Analysis.** The number of markers showing a significant association (based on KW and IM) with each of the seven phenotypic traits is summarized in Table 2. IM failed to identify any markers associated significantly with canopy diameter, tree height, or trunk diameter at either location. KW identified five significant markers for trunk diameter and three for canopy diameter at Irvine and a single significant marker for canopy diameter and tree height at Riverside.

The content of total carotenoids in the fruit did not show significant association with any marker based on IM (Table 2). Based on KW, significant QTLs were located on LG1, 3, and 6. QTL analysis of fruit β-sitosterol content at Riverside in 2010 revealed one marker (SHRSPaS006673) at 61.087 cM on LG1 to be significantly associated using IM at a LOD of 3.72 (Fig. 2), explaining 35.6% of the variance (Table 2). This marker also achieved significance in the KW analysis in the same location and year, and at Irvine in 2009 (Table 2). Marker SHRSPaS001205 (LG1), less than 2 cM away from SHRSPaS006673, was also significantly associated with β-sitosterol content at Irvine in 2009 and Riverside in 2010, based on KW analysis. Figure 2 compares the IM LOD profiles of markers on LG1 for β-sitosterol in all four datasets (Irvine and Riverside in 2009 and 2010).

In IM analyses, markers on LG3 were significantly associated with α-tocopherol content at Irvine in both years—12 in 2009 and 15 in 2010—achieving LOD values of up to 4.52 and 4.61, respectively, and explaining up to 37.7% and 38.3% of the variance, respectively (Table 2). No marker attained significance based on IM at Riverside in 2009. Two markers, SHRSpa01282 and SHRSpa003314, were declared significant at both locations and in both years, based on IM and/or KW. Significant QTLs resided on the proximal end of LG3 at 7.968 to 18.601 cM (IM) and at 0 to 27.638 cM (KW; Fig. 2). Three HPT1 CG-SNPs were declared significant based on KW only (Supplemental Table 3). Flowering type showed significant association with many markers under IM, with LOD values far exceeding the permutation-based thresholds for significance. IM showed a significant association with 45 markers, all of which resided on LG10 (Table 2; Supplemental Table 3; Fig. 2). Twenty-four markers on LG10 exceeded the genome-wide LOD threshold of 7.1 and explained 50.4% to 100% of the variance in flowering type. Six of these markers achieved LOD scores of 99.99 in IM and explained 100% of the variance—were disregarded because they did not track phenotypic values and represented an artifact of the IM maximum likelihood algorithm applied to non-normal (discrete) data (Van Ooijen, 2009). A further 21 markers on LG10 exceeded the LG-specific LOD threshold, including the CG-SNP DXPS1_1593. All markers on LG10 declared significant at the genome-wide cutoff were located between 26.808 to 53.308 cM (Supplemental Table 3; Fig. 2). Eight of the 24 markers exceeding the genome-wide threshold under IM received no support in the KW test, including the six markers with a 99.99 LOD score. KW analysis identified 22 markers associated significantly with flowering type (Table 2), all but one also residing on LG10: a single QTL-associated marker, SHRSPaS003811, located to LG6 (Supplemental Table 3). The two highest-scoring markers in the KW test had $K^*$ values of 38.251 (SHRSPaS001390 and SHRSPaS004380) and were declared significant at $P = 0.0001$. Their validity as QTLs was endorsed by IM, which assigned LOD values of 18.66 and 18.34, respectively. Among the markers associated significantly
Table 2. Evaluation of quantitative trait loci (QTLs) identified by interval mapping (IM) or Kruskal–Wallis analysis (KW) implemented in MapQTL version 5 (Van Ooijen, 2004) for avocado mapping populations growing at two locations in southern California (Irvine and Riverside). Comparisons are made for all markers declared to be significant under the interval mapping (IM) or Kruskal–Wallis (KW) algorithms. Column headings details are as follows. IM = the number of significant loci declared by IM; in parentheses is the percentage of the variance explained by the locus with the highest log-of-odds (LOD) score. KW = the number of significant loci with a significance of **** or higher, based on KW. LGs-IM = the number of different linkage groups (LGs) from which significant markers were drawn, based on IM. LGs-KW = the number of different LGs from which significant markers were drawn, based on KW. QTL ≥ two environments = the number of QTLs present in at least two environments (two locations and 2 years for nutrients; two locations for tree measurements).

| Nutrient         | Location, yr | IM [no. (%)] | KW (no.) | LGs-IM (no.) | LGs-KW (no.) | QTL ≥ two environments |
|------------------|--------------|--------------|----------|--------------|--------------|------------------------|
| Alpha-tocopherol | Irvine, 2009 | 12 (37.7)    | 21       | 1            | 3            | 21 (5, 2)*              |
|                  | Irvine, 2010 | 15 (38.3)    | 24       | 1            | 4            |                        |
|                  | Riverside, 2009 | 0 (39.5) | 11 | n/a | 3            |                        |
|                  | Riverside, 2010 | 0 (37.4) | 14 | n/a | 1            |                        |
| Beta-sitosterol  | Irvine, 2009 | 0 (34.8)    | 17       | n/a          | 1            | 11                     |
|                  | Irvine, 2010 | 0 (33.9)    | 6        | n/a          | 1            |                        |
|                  | Riverside, 2009 | 0 (35.0) | 5 | n/a | 2            |                        |
|                  | Riverside, 2010 | 1 (35.6) | 12 | 1 | 3            |                        |
| Carotenoids      | Irvine, 2009 | 0 (28.5)    | 1        | n/a          | 1            | 1                      |
|                  | Irvine, 2010 | 0 (31.4)    | 3        | n/a          | 2            |                        |
|                  | Riverside, 2009 | 0 (35.8) | 8 | n/a | 1            |                        |
|                  | Riverside, 2010 | 0 (35.4) | 3 | n/a | 2            |                        |
| Trunk diameter   | Irvine, 2005 | 0 (26.6)    | 5        | n/a          | 3            | 0                      |
|                  | Riverside, 2005 | 0 (25.3) | 5 | n/a | 3            |                        |
| Canopy diameter  | Irvine, 2005 | 0 (26.3)    | 3        | n/a          | 2            | 0                      |
|                  | Riverside, 2005 | 0 (34.1) | 1 | n/a | 1            |                        |
| Height           | Irvine, 2005 | 0 (33.9)    | 0        | n/a          | n/a          | 0                      |
|                  | Riverside, 2005 | 0 (31.8) | 1 | 0 | 1            |                        |
| Flowering type   | Irvine, 2013 | 45 (24) (100.0) | 22 | 1 | 2 | n/a |

*In parentheses: number of QTLs shared by three and four environments, respectively.

24 QTLs for flowering type were declared significant using the genome-wide permutation threshold [18 after elimination of 6 QTLs with artifactually high LOD values (Van Ooijen, 2009)] and 45 using the linkage-group specific threshold (39 after adjusting for artificial LOD values).

Discussion

Despite the limited statistical power associated with small sample sizes, this study provided useful mapping information on two important phenotypic traits: flowering type and vitamin E (α-tocopherol) content of the fruit.

Flowering type is not a quantitative trait, and Lavi et al. (1993) suggested control by several loci with several alleles at each locus. A closer look at our data for flowering type uncovered a one-gene Mendelian model that likely governs this important trait in avocado. Using the 13 top-scoring loci on LG10 endorsed by both IM and KW, pairwise analysis showed that they were highly correlated with one another, suggesting a single causal locus with flanking loci linked through linkage disequilibrium (LD). Moreover, 29 (100%) individuals with genotype “ll” had B-type flowering, whereas—among individuals with genotype “lM”—16 (89%) individuals had A-type flowering and 2 (11%) individuals had B-type flowering. These results indicate that “M” is the dominant allele while “l” is the recessive allele. The two individuals with genotype “lM” showing the unexpected phenotype likely reflect the effect caused by a gene × environment interaction, which may reduce the penetrance of the dominant trait. This assumption is well supported by Sedgley and Annells’ findings (1981), which indicated that avocado flowering was affected by cold temperature, allowing the male and female phases of the flower to overlap. Elucidation of the genes determining flowering type would provide greater flexibility to growers in their choice of pollinator cultivars.

Alpha-tocopherol content exhibited moderate to high heritability in quantitative genetic analyses (Calderón-Vázquez et al., 2013; Chen et al., 2007) and might be expected to yield some success in breeding programs. The current mapping studies suggest that the variation underlying flowering type and α-tocopherol may be the result of mutations at a single genetic locus. A third trait (β-sitosterol content of the fruit), also with a substantial heritability (Calderón-Vázquez et al., 2013), provided promising, although not entirely consistent, evidence for a particular chromosomal location.

Not surprisingly, traits of low to moderate heritability do not give consistent results in the QTL analyses, as is the case for plant height, canopy diameter, and trunk diameter [broad-sense heritability estimates in the low- to medium range (0.266 to 0.366; Chen et al., 2007)]. Variation underlying these morphological traits is likely to be controlled by many loci throughout the genome and to be subject to substantial environmental variation. So, the failure to map variants associated with these traits is to be expected. Moreover, the high positive correlations between these three measurement traits (Chen et al., 2007) suggest that breeding for tree architecture may not be straightforward. The fact that QTL analysis for these three growth traits
revealed few significant QTLs under KW analysis (and none under IM) suggests that marker-assisted selection (MAS) for these growth-related traits is not worthwhile.

**NUTRITIONAL TRAITS.** Appreciable genetic determination of the fruit nutrient phenotypes was shown by Calderón-Vázquez et al. (2013), who determined broad-sense heritability for α-tocopherol, β-sitosterol, and carotenoids to be 0.76, 0.61, and 0.47, respectively. Considerably higher values than those of the tree measurements, these values are consistent with the fact that nutritional traits are the outcome of specific biochemical pathways. Additionally, correlations among the three nutritional traits were low, the highest arising between α-tocopherol and β-sitosterol at R = 32% (Calderón-Vázquez et al., 2013). Low correlation also may be due to the discrete biochemical pathways underlying the biosynthesis of these nutrients and will facilitate independent breeding. Significant genotype effects were found for all three nutritional traits (Calderón-Vázquez et al., 2013), but for the other variance components (year, location, and interaction effects), each nutrient responded differently. Combined with the current results, these findings argue that a focus on nutritional/biochemical traits can be effective, despite limited population sizes.

Among the nutrient data, few QTLs performed well across all four environments (two locations and 2 years). Significant QTLs for carotenoid and β-sitosterol contents were never shared by more than two environments (1 and 11 QTLs, respectively, were shared by 2 environments; Table 2). Of 21 QTLs for α-tocopherol that were common to at least two environments, five were present in three environments, and two were present in all four environments (Table 2). The discovery of QTL loci that tracked nutrient content across multiple environments is encouraging and presumably reflects genes with stable expression under different environmental conditions.

For β-sitosterol, the QTL achieving significance at Riverside in 2010 did not stand out in the other year/location combinations, calling into question whether this QTL will be amenable to MAS. It is worth noting, however, that this sole significant marker on LG1 was located adjacent (within 0.49 cM) to an EST-derived FL-SSR marker (SHRSPa102; Supplemental Table 3) that had a very low LOD value in most IM datasets, suggesting SHRSPa102 may not have been correctly placed on the CA-map (Van Ooijen, 2006). The position is visible as an abrupt deep incision on the LOD graph (Fig. 2). It is conceivable that the proximity of an incorrectly placed marker affected the LOD value within the interval surrounding the significant QTL.

**POPULATION SIZE CONSIDERATIONS.** As noted earlier in the section on QTL analysis, one aspect of this study—the small population size—clearly limited the power to generate a robust linkage map and to detect QTLs in avocado. Small population size exerts its primary effect by reducing the number of recombination events, leading to identical segregation of many markers, which results in their elimination as identicals in JoinMap (Van Ooijen, 2006) and a loss of marker information.

---

**Fig. 2.** Plots charting the log-of-odds (LOD) values of markers significantly associated with avocado fruit α-tocopherol contents on linkage group (LG) 3, β-sitosterol contents on LG1, and flowering type on LG10. For β-sitosterol and α-tocopherol, separate LOD plots are shown for each of 2 years and two locations studied [South Coast Research & Extension Center in Irvine, CA (SC) and Agricultural Operations of the University of California at Riverside (UCR)]. X-axes show map positions (cM).
A paucity of recombination events also results in relatively large chromosomal segments. This result, in turn, will tend to reduce the accuracy of QTL markers identified by the mapping algorithms, because the markers may be at some distance from the functional gene. Scarcie recombination events may also make mapping and QTL analysis more sensitive to the stochastic nature of allelic segregation, potentially leading to the underestimation of marker distances. In outbreeding full-sib families (CP population type in JoinMap; Van Ooijen, 2006), the mapping algorithm estimates the consensus map by averaging the positions of anchor markers segregating in both parents. However, because “hk” genotypes cannot be used (in heterozygotes sharing the same two alleles, it is impossible to tell from which parent respective alleles originated), the number of informative recombination events is thus further reduced from an already small segregation pool. Segregation type will also affect QTL estimation via the IM algorithm where flanking markers are used in the calculation of LOD values for markers with uninformative segregation. While any population size will contain a proportion of markers with uninformative segregation, small populations are likely to be more heavily impacted. Because the CA-map and QTL analyses were based on the same segregating population, errors in the calculation of QTL probabilities due to a mismatch in these two components can be ruled out (Van Ooijen, 2009).

Segregation distortion (SD), a phenomenon describing loci whose alleles do not segregate according to Mendelian expectations, affects recombination between marker loci (Wang et al., 2005) and often is accused of leading to the detection of false linkage. We chose to exclude strongly SD-affected markers before generating the linkage map, though they represented <3% of the total number of markers. This exclusion may have inadvertently removed potential QTLs, because distorted regions are as—or more—likely to contain QTLs as SD-free regions (Wang et al., 2005; Xu, 2008). In particular, SD markers are thought to be linked to loci for viability selection (Vogl and Xu, 2000), including those causing inbreeding depression, a phenomenon common to outbreeding species such as avocado. While we cannot be sure that QTLs may have been missed, the loss of power arising from ignoring distorted markers is negligible in dense maps (Xu, 2008).

**Candidate gene analysis.** It is disappointing that the SNPs we developed from candidate genes did not show more significant association with the nutrient phenotypes whose production the causative genes are assumed to control. One reason may be that the shortage of recombination events in our mapping population failed to detect signal. However, other factors may also be responsible. Tabor et al. (2002) argued that the candidate gene approach relies on a priori hypotheses about the role of candidate genes that may not be supported by a sufficient body of knowledge. Moreover, assumptions of gene function are generally based on studies in model organisms or major crops; yet the information may not be pertinent in avocado, an early-diverging angiosperm lineage. Further factors may be modulating effects exerted by genes outside the candidate gene pathways. Studies in Arabidopsis thaliana (Gilliland et al., 2006) and maize (Zea mays; Wang et al., 2018) identified QTLs controlling seed tocopherol content that were not part of known vitamin E pathways. In our study, CG-SNPs developed to the gene encoding the enzyme homogentisate phytlyl transferase (HPT1), the first committed gene in the tocopherol VTE2 biosynthetic pathway, were located in close proximity to markers significantly associated with α-tocopherol content and were identified as significant under KW at both locations in 2010 but at neither location in 2009. Insufficient map resolution or uninformative segregation in the flavanking markers may be responsible for the failure of IM to declare significance for the HPT1 CG-SNPs.

The only other CG-SNPs showing significant association with a phenotype (flowering type) was DXPAS1, a SNP developed to a candidate gene from the vitamin B complex, that controls synthesis of a thiamine-dependent enzyme involved in cell metabolism.

Vitamin E, which consists of α-tocopherol and several other tocopherol isomers, has been targeted by breeders pursuing crop biofortification in barley (Hordeum vulgare), maize, rapeseed (Brassica napus), rice (Oryza sativa), soybean (Glycine max), and tomato (Solanum lycopersicum) (reviewed in Fritsche et al., 2017). Peraza-Magallanes et al. (2017) found considerable variation for α-tocopherol content in avocado germplasm from Sinaloa, Mexico. Aside from the nutritional benefits arising from elevated vitamin E levels in crops, α-tocopherol has also been associated with enhanced tolerance of salinity and drought stress in rice and tobacco (Nicotiana tabacum) (Munne-Bosch, 2007; Ouyang et al., 2011).

**Experimental populations.** Avocado is a large tree that requires significant space, water, and labor resources. It takes 5 to 8 years to become productive (Lahav and Lavi, 2009), and its breeding system is very difficult to experimentally manipulate (Degani et al., 2003; Lammerts, 1942). Such cost and time considerations make it difficult and expensive to create and maintain large experimental populations and, in turn, favor working with small preexisting populations. In this regard, the UC populations used here have several strengths: 1) replication of progeny genotypes on a single clonal rootstock provides an estimate of within-genotype error variances; 2) replication in two locations provides a measure of location effects; and 3) multiple-year measurements provide a measure of temporal variance. These design features help identify important sources of environmental variance and point to important management considerations.

The current data were generated for a ‘Gwen’ × ‘Fuerte’ progeny array, and findings may not be fully transferrable to other cultivars and germplasm. However, ‘Gwen’—a grandchild of ‘Hass’—is central to the existing UC Riverside Breeding Program, making the QTL data relevant for MAS in the future. A crucial question to be confronted is whether QTL studies on a difficult tree crop justify the cost of land, time, and labor resources. More advanced technologies such as transformation and clustered, regularly interspaced short palindromic repeats (CRISPR)-CAS9 are appealing; but basic information about potential target genes is deficient, so for the time being MAS seems like the most practical alternative to relatively inefficient phenotypic selection. We believe that our results will encourage expanded QTL studies to guide the breeding of future cultivars in California and elsewhere, and that our findings will bring into focus the role of fruit nutritional traits with the long-term goal of breeding high-value/nutritionally enhanced cultivars achieving a market premium.

**Literature Cited**

Alcaraz, M.L. and J.I. Hormaza. 2009. Selection of potential pollinizers for ‘Hass’ avocado based on flowering time and male–female overlapping. Scientia Hort. 121:267–271.
Ashworth, V.E.T.M., M.C. Kobayashi, M. de la Cruz, and M.T. Clegg. 2004. Microsatellite markers in avocado (Persea americana Mill.). Developing dinucleotide and trinucleotide markers. Scientia Hort. 101:255–267.

Bergh, B.O. and E. Lahav. 1996. Avocados, p. 113–166. In: J. Janick and J.N. Moore (eds.). Fruit breeding, vol. I: Tree and tropical fruits. Wiley, New York, NY.

Borraine, J.W., R.J. Schnell, H.A. Violi, and R.C. Ploetz. 2007. Seventy microsatellite markers from Persea americana Miller (avocado) expressed sequence tags. Mol. Ecol. Notes 7:439–444.

Calderón-Vázquez, C., M.L. Durbin, V.E.T.M. Ashworth, L. Tommasini, K.K.T. Meyer, and M.T. Clegg. 2013. Quantitative genetic analysis of three important nutritive traits in the fruit of avocado. J. Amer. Soc. Hort. Sci. 138:283–289.

Chen, H., V.E.T.M. Ashworth, S. Xu, and M.T. Clegg. 2007. Quantitative genetic analysis of growth rate in avocado. J. Amer. Soc. Hort. Sci. 132:691–696.

California Irrigation Management Information System. 2003. CIMIS: California Irrigation Management Information System. 17 June 2019. <https://cimis.water.ca.gov/Resources.aspx>.

Destefani, S.M. 2007. Phytonutrients: A more natural approach toward cancer prevention. Semin. Cancer Biol. 17:345–346.

Degani, C., E. Lahav, R. El-Batsri, and S. Gazit. 2003. Caging single avocado trees with a beehive does not guarantee exclusive formation of selfed progeny. HortScience 38:1433–1434.

D’Ambrosio, S.M. 2007. Phytotrienols: A more natural approach toward cancer prevention. Semin. Cancer Biol. 17:345–346.

Ewing, B. and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8:186–194.

Fiori, A.S. 2007. Persea americana Miller (avocado): expressed sequence tags. Mol. Ecol. Notes 7:439–444.

Jeong, W.S. and P.A. Lachance. 2001. Phytosterols and fatty acids in Ficus carica, var. Mission) fruit and tree components. J. Food Science 66:278–281.

Kruglyak, L. and E.S. Lander. 1995. Nonparametric approach for mapping quantitative trait loci. Genetics 139:1421–1428.

Kruskal, W.H. and W.A. Wallis. 1952. Use of ranks in one-criterion variance analysis. J. Amer. Stat. Assn. 47:583–621.

Lamer, E.S. and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Latorre, H., K.N. Risch, and R.M. Myers. 2002. Candidate-gene approaches for studying complex genetic traits: Practical considerations. Nat. Rev. Genet. 3:391–397, doi: 10.1038/nrg796.

Lopez-Ledesma, R., A.C. Frati-Munari, B.C. Hernandez-Dominguez, S. Cervantes-Montalvo, M.H. Hernandez-Luna, C. Juarez, and S. Moran-Lira. 1996. Monounsaturated fatty acid (avocado) rich diet for mild hypercholesterolemia. Arch. Med. Res. 27:519–523.

Luterotti, S., D. Bicanic, and R. Pozgaj. 2006. New simple spectro-photometric assay of total carotenes in margarines. Anal. Chin. Acta 573–574:466–473.

Macleod, E., P. Lääniite, J. Joudu, and U. Määg. 2007. Some important aspects of sterol analysis of vegetable oils. Proc. Estonian Acad. Sci. Chem. 56:59–66.

Munné-Bosch, S., E.W. Weiler, L. Alegre, M. Müller, P. Dütch, and J. Falk. 2007. α-Tocopherol may influence cellular signaling by modulating jasmonic acid levels in plants. Planta 225:681–691.

National Center for Biotechnology Information. 2005. NCBI: National Center for Biotechnology Information. 17 June 2019. <http://www.ncbi.nlm.nih.gov>.

Nickerson, D.A., V.O. Tobe, and S.L. Taylor. 1997. PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. 25:2745–2751.

Ouyang, S., S. He, P. Liu, W. Zhang, J. Zhang, and S. Chen. 2011. The role of tocopherol cyclase in salt stress tolerance of rice (Oryza sativa). Sci. China Life Sci. 54:181–188.

Peraza-Magallanes, A., M.A. Perea-Camacho, E. Sandoval-Castro, S. Medina-Godoy, M. Valdez-Morales, M.T. Clegg, and C.L. Calderón-Vázquez. 2017. Exploring genetic variation, oil and alpha-tocopherol content in avocado (Persea americana) from northwestern Mexico. Genet. Resources Crop Evol. 64:443–449, doi: 10.1007/s10722-016-0478-9.

R Core Team. 2019. R: A language and environment for statistical computing. 5 Apr. 2019. <http://www.R-project.org>.

Rendón-Anaya, M., E. Ibarra-Laclette, A. Méndez Bravo, T. Lan, C. Zheng, L. Carretero-Paulet, C.A. Perez-Torres, A. Chacón-López, G. Hernandez-Guzmán, T.H. Chang, K.M. Farr, W.B. Barbazuk, S. Chamala, M. Mutwil, D. Shihvare, D. Alvarez-Ponce, N. Mutter, A. Hayward, S. Fletcher, J. Rozas, A. Sánchez Gracia, D. Kuhn, A.F. Barrientos-Priego, J. Salojárvi, P. Librado, D. Sankoff, A. Herrera-Estrella, V.A. Albert, and L. Herrera-Estrella. 2019. The avocado genome informs deep angiosperm phylogeny, highlights introgressive hybridization, and reveals pathogen-influenced gene space adaptation. Proc. Natl. Acad. Sci. USA. doi: 10.1073/pnas.1822192116.

Ryan, E.K., K. Galvin, T.P. O’Connor, A.R. Maguire, and N.M. O’Brien. 2007. Phytosterol, squalene, tocopherol content and fatty acid profile of selected seeds, grains, and legumes. Plant Foods Hum. Nutr. 62:85–91.

Sedgley, M. 1985. Some effects of daylength and flower manipulation on the floral cycle of two cultivars of avocado (Persea americana Mill., Lauraceae), a species showing protogynous dichogamy. J. Expt. Bot. 36:823–832.

Sedgley, M. and C.M. Amnells. 1981. Flowering and fruit-set response to temperature in the avocado cultivar ‘Hass’. Scientia Hort. 14:27–33.

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

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

U.S. Climate Data. 2018. U.S. Climate Data. 17 June 2019. <https://www.usclimatedata.comclimate/>.

U.S. Department of Agriculture. 2018. Noncitrus fruits and nuts 2017 summary, June 2018. 17 June 2019. <http://usda.mannlib.cornell.edu/usda/current/NoncFruitNu/NoncFruitNu-06-26-2018.pdf>.

Van Nocker, S. and S.E. Gardiner. 2014. Breeding better cultivars, faster: Applications of new technologies for the rapid deployment of superior horticultural tree crops. Hort. Res. 1:14022, doi: 10.1038/hortres.2014.22.
Van Ooijen, J.W. 2004. MapQTL® 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma, Wageningen, The Netherlands.

Van Ooijen, J.W. 2006. JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma, Wageningen, The Netherlands.

Van Ooijen, J.W. 2009. MapQTL® 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma, Wageningen, The Netherlands.

Vogl, C. and S. Xu. 2000. Multipoint mapping of viability and segregation distorting loci using molecular markers. Genetics 155:1439–1447.

Voorrips, R.E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93:77–78.

Wang, C., C. Zhu, H. Zhai, and J. Wan. 2005. Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (Oryza sativa L.). Genet. Res. 86:97–106.

Wang, H., S. Xu, Y. Fan, N. Liu, W. Zhan, H. Liu, Y. Xiao, K. Li, Q. Pan, W. Li, M. Deng, J. Liu, M. Jin, X. Yang, J. Li, Q. Li, and J. Yan. 2018. Beyond pathways: Genetic dissection of tocopherol content in maize kernels by combining linkage and association analyses. Plant Biotechnol. J. 16:1464–1475.

Xu, S. 2008. Quantitative trait locus mapping can benefit from segregation distortion. Genetics 180:2201–2208.
Supplemental Fig. 1. Dot plots showing map positions [cM] of all single nucleotide polymorphism (SNP) markers shared between the avocado ‘Gwen’ × ‘Fuerte’ California (CA)-map (x axis) (this study) and the integrated consensus linkage map of a ‘Simmonds’ × ‘Tonnage’ and ‘Hass’ × ‘Bacon’ reciprocal cross [Florida (FL)-map] (y axis) (Rendón-Anaya et al., 2019). All shared markers located to the same avocado linkage group, but marker arrangements differed in some cases.
Supplemental Fig. 2. Avocado linkage map generated using JoinMap version 4 (Van Ooijen, 2006) and displayed with MapChart (Voorrips, 2002).
Supplemental Table 1. Parameters used for single nucleotide polymorphism (SNP) discovery from candidate genes in avocado. Details are presented in the order (1) abbreviation used on the linkage map, (2) full name of enzyme encoded by the candidate gene, (3) functionally characterized gene accession found in the public databases of the National Center for Biotechnology Information (NCBI) (with the organismal source of the sequence, where given) showing the highest similarity, (4) similarity score, (5) probability of being the same gene (E-value), (6) number of SNPs detected in the gene, (7) amplifying primer (forward), (8) amplifying primer (reverse), (9) sequencing primer, annealing temperatures listed in same order as the three primers, if different.

Carotenoids

B1: Beta-carotene hydroxylase 1; At4G25700.1, 78.8%, 9e-18, 1, GAA CGA TGT TTT TGC GAT CA (B1-F147), AAC AGC CGG TAT GGC ACT C (B1-R443), CGT ATG GCA CTC CAT TGA A (B1-nest509F), 64, 65, 62

LUT5: Carotene beta-ring hydroxylase, cytochrome P450-type monoxygenase; AT1G31800.1, 71.6%, 1e-15, 1, ACG GTG GTA GTC CTC GTG AT (LUT-F58), TTT TTC TCT GTG TGG ATT GGA (LUT-R473), ACG GTG GTA GTC CTC GTG AT (LUT-F58), 64, 63, 64 PSY: Phytoene synthase (PSY), geranylgeranyl-diphosphate geranylgeranyltransferase, At5G17230.1, 158%, 1, GCC GTG TCA GGA TTA GGA AT (PSY-F22), TTT CGA CCA TGA ATC GCA (PSY-R-GAA), GGG GAT TTT ATT AGA AAA TGA (PSY-nest658R), 64, 60, 56

ZDS: Zeta-carotene desaturase (ZDS), [Citrus sinensis mRNA for zeta-carotene desaturase]; emb|AJ319762.1, 89.7%, 5e-14, 1, TCC TCC AGG ACC TGA GCA CT (ZDS-F372), GGT TGT TGT AGC AGC CAA A (ZDS-R-GGT), CAC ATG CAG CTC CAT TAC A (ZDS-R-CAC), 64, 61, 63

Darkening-related

Symbols: DXPS1, Vitamin B1 (thiamine), 1-deoxy-D-xylulose-5-phosphate synthase (DXPS1); gi|311337316|gb|HQ380894.1| [Nelumbo nucifera polyphenol oxidase mRNA], 470, 1.00E-128, 1, ACC AGC TGC TTG TTT TCA TC TC 5093, CCC TCC CAT GGT TCT TAC CT 5094, CCC TCC CAT GGT TCT TAC CT (5094), 54

Fatty acid pathway

CUT1: Acyltransferase, Cuticular 1 (CUT 1); AT1G68530.1, 289, 2e-81, 2, CAT GGT GAT AGC TGG TGA CG, (CUT1-F27), TCT GGG ACA GAT AGG GGA TG (CUT1-R554), CATGGTAGACTGGTACG (CUT1-F27), 64

Flavonoid, anthocyanin & phenylpropanoid pathways

Caf3: Caffeoyl-CoA O-methyltransferase (caff3); Os09g30360[2009.02714] [unspliced-genomic caffeoyl-CoA O-methyltransferase 1, putative, expressed], 91.5%, 8e-22, 5, TGC GCA GAA GGA CAA CTA CA (caff3-F50), CCA TGA TGC CTC CAT CTC TAG CA (caff3-R483), CCA AAT ATC AGA AAC AG (caff3-nest658R), 64, 65, 60

OMT1: Flavonol 3’-O-methyltransferase 1 (OMT1); gb|GU324973.1| [Eucalyptus camaldulensis caffeic O-methyltransferase1 (COMT1) gene], 66.2%, 8e-07, 7, GCA GAT TTC CTA AGG GAA TTT CGC (OMT1-F103), GGT CGA CCT ACA ATG TGC G (OMT1-R568), GAT CAC TTT CTT ATG CCG (OMT1-nest70F), 61, 62, 61

PAL2: Phenylalanine ammonia-lyase 2 (PAL2); At3G3260.1, 702, 0.0, 2, CAG ATG GAA TGG CAC ACT TCC AA (PAL2-F17), AGC AAA TGG GAA TAG GAG CA (PAL2-R1065), CAT GGT GAT AGC TGG TGA CG (CUT1-R554), CATGGTAGACTGGTACG (CUT1-F27), 64

Isoprenoid & sitosterol

CYP: cycloeucalenol cycloisomerase; gi|225456279|ref|XM_002283523.1| [predicted: Vitis vinifera cycloeucalenol cycloisomerase-like (LOC100262783), mRNA], 659, 0, 5, GCT TCA TAC ACC TTT CCG TCA 6163, CAT GAT GCC TCA GCA ATC C 6162, TAG GCA TTA CGG AGT TGC AG 2130, 53

FPS: farnesyl diphosphate synthase; gi|212960745|gb|FJ415102.1| Chimonanthus praecox farnesyl pyrophosphate synthase (FPS) mRNA, complete cds, 690, 0, 1, TGG GTT GGT GTG GAT GAA ACC TC 634, TTG CCC AAG AAA GAC TTT AGC 737, TTG GAT GGT GTA GAC AGC TC 634, 53

MCR: 24-dehydrocholesterol reductase; gi|359473656|ref|XM_002271810.2| [predicted: Vitis vinifera delta(24)-sterol reductase-like (LOC100258158), mRNA], 592, 1.00E-165, 3, GGA AAG GTA TGC TCC CAA GG 20, TGT GAA GTT CAT ATA ACG AAT AGT CA 7963, TTG GCC TCA ATT TAG CTG 3878, 53

SQS: squalene synthase (SQS1); gi|359475094|ref|XM_002266114.2| [predicted: Vitis vinifera squalene synthase-like (LOC100265798), mRNA], 682, 0, 4, TGA AAG GTA TGG CAC ACT TCC AA (PAL2-F17), AGC AAA TGG GAA TAG GAG CA (PAL2-R1065), CAT GGT GAT AGC TGG TGA CG (CUT1-R554), CATGGTAGACTGGTACG (CUT1-F27), 64

Vitamin B complex

atrans, Vitamin B9 (folic acid), Aminotransferase class IV family (atrans), Aminotransferase class IV family (atrans); AT5G57850.1 | Symbols: | aminotransferase class IV family protein, 66.2%, 5e-14, 2, 6, CAT GAT CG CAG CCA CAA TGA TA (atrans-F-12), ACC ATG GGA GGC TTC ATT GG (atrans-R-457), TGA CAC TGC ATC TAT (atrans-R-457-ic-TGA), 64, 66, 51

BCAT3, Vitamin B5 (pantothenic acid), Branched-chain aminotransferase 3 (BCAT3), Branched-chain aminotransferase 3 (BCAT3); gb|EU194916.1| Nicotiana benthamiana branched-chain aminotransferase (BCAT) mRNA, 181%, 1e-41, 1, CAA GGT AAA ACA TCC TAG ATC (BCAT3-F6), ACC CTT TAC TGG TGT GGC CG (BCAT3-R-ACC), GAA CCA GAA AAG CAG CAG (BCAT3-nest513F), 57, 63, 61

DXPS1, Vitamin B1 (thiamine), 1-deoxy-D-xylulose-5-phosphate synthase (DXPS1), 1-deoxy-D-xylulose-5-phosphate synthase (DXPS1); At3G21500.1 | Symbols: DXPS1 | DXPS1; 1-deoxy-D-xylulose-5-phosphate synthase, 239%, 6e-66, 5, CAG GGT AAA ACA TCC TAG ATC (DXPS1-F34), AAG CAG CAC CCA AGC AGC TT (DXPS1-R-AAAG), AAA TGC ATC ATA TTT TAG GAA (DXPS1-F34-R389), 57, 69, 55

PDX1, Vitamin B6, Pyridoxin biosynthesis 1 (PDX1); gi|356543999|ref|XM_003542937.1| PREDICTED: Glycine max pyridoxal biosynthesis protein PDX1-like (LOC100816306), mRNA, 589, 1.00E-164, 4, CAC ACC CAA GCT GCA TCA 787, AAA TCA AGG AGG CCG TCA C 789, CAC ACC CAA GCT GCA TCA 787, 59

Continued next page
Supplemental Table 1. Continued.

PDX2, Vitamin B6, Pyridoxin biosynthesis 2 (PDX2); gi[359478338]ref[XM_002285059.2] PREDICTED: Vitis vinifera pyridoxal biosynthesis protein PDX2-like (LOC100267348), mRNA, 100, 2.00E-17, 1, AAA CAG GGA AAC CTG TGT GG 779, GCC TGG TGG AAC AGC ATA AT 784, AAA CAG GGA AAC CTG TGT GG 779, 54

Vitamin C
MEP: GDP-mannose-3',5'-epimerase, gi[359487867]ref[XM_002279341.2] PREDICTED: Vitis vinifera GDP-mannose-3',5'-epimerase (LOC100233034), mRNA, 437, 3.00E-119, 2, TGC TGG CAT ATA CCC AGA GTT 8889, AAG GAT TGT GGC AGA CC 3058, AAG GAT TGT GGC AGA CC 3058, 54
PGI: phosphoglucose isomerase, gi[225458304]ref[XM_002282738.1] PREDICTED: Vitis vinifera glucose-6-phosphate isomerase (LOC100252335), mRNA, 515, 1.00E-142, 4, TGA TAC TGG GAA AAT ACA TGA AAA CA 3881, TAA AGC CCT CAA CTG GTT CC 870, TGA TAC TGG GAA AAT ACA TGA AAA CA 3881, 54

VTC1: GDP-mannose pyrophosphorylase (VITAMIN C DEFECTIVE 1), gi[224038261]gb[FJ643600.1] Actinidia latifolia GDP-D-mannose pyrophosphorylase (GMP) mRNA, complete cds, 614, 4.00E-172, 3, GAA ACC GAG CCT CTA GGA AC 738, AGA AGC CCG GTA AGA CCA T 740, AGA AGC CCG GTA AGA CCA T 740, 54

VTC2: GDP-L galactose phosphorylase (VITAMIN C DEFECTIVE 2), gi[319739580]gb[HQ224948.1] Citrus unshiu putative GDP-L-galactose-pyrophosphatase mRNA, complete cds, 246, 2.00E-61, 3, AAA ATC AAG CAT TCG CAG AG 340, CAG GCT CTT GGA GAG GTG AG 5859, AAA ATC AAG CAT TCG CAG AG 340, 54

Vitamin E
HPT1: Homogentisate phytyltransferase (VTE2), gi[219842165]dbj[AB376091.1] Hevea brasiliensis hpt mRNA for homogentisate phytlytransferase, complete cds, 347, 7.00E-92, 3, AGG CCA TGG ATA TTC GCA AC 9827, GAA ACC AAT CCC ATC ACC AC 9825, AGG CCA TGG ATA TTC GCA AC 9827, 54

PDS1: 4-hydroxyphenylpyruvate dioxygenase (PHYTOENE DESATURASE 1), gi[359485346]ref[XM_002283239.2] PREDICTED: Vitis vinifera 4-hydroxyphenylpyruvate dioxygenase-like (LOC100248785), mRNA, 558, 3.00E-155, 3, GCT GGA AAT GTG ACT GA 991, TCC CAT GTC TTC ATT GAC 7960, GCT GGA AAT GTG ACT GA 991, 53

VTE1: Tocopherol cyclase (VITAMIN E DEFECTIVE 1, VTE1), gi[255550999]ref[XM_002516502.1] Ricinus communis Tocopherol cyclase, chloroplast precursor, putative, mRNA, 91.5, 2.00E-14, 5, GGG CAG TGC AAG AAT ATA ACT G 6564, CTC CAA GAT GGC AGA ATG AG 996, GGG CAG TGC AAG AAT ATA ACT G 6564, 54

VTE3: MPBQ/MSBQ methyltransferase (VTE3), gi[219842171]dbj[AB376094.1] Hevea brasiliensis mmgbqmt mRNA for 2-methyl-6-geranylgeranylbenzoquinone methyltransferase, complete cds, 814, 0, 2, TGG CTT CAA TGC TCA AT 350, GCA TAA TCA TGT GGG AAT GG 5758, TGG CTT CAA TGC TCA AT 350, 54

VTE4: Gamma-tocopherol methyltransferase (VTE4), gi[219842175]dbj[AB376096.1] Hevea brasiliensis gamma-tmt mRNA for gammatocopherol methyltransferase, complete cds, 381, 2.00E-102, 5, GAA CAC CAA GCC GGA AGA AT 3026, GAG AGC ACA TGC TCA ATA AA 996, GAA CAC CAA GCC GGA AGA AT 3026, 54
Supplemental Table 2. Information on simple sequence repeat (SSR) markers of avocado, featuring marker name, source, fragment sizes in cultivars Gwen and Fuerte, distorted segregation (if applicable), forward primer, 5’ to 3’, reverse primer, 3’ to 5’, nucleotide repeat unit, annealing temperature [°C], and GenBank accession number.

| Marker Name | Year | Frag Sizes | Distorted Segregation | Forward Primer, 5’ to 3’ | Reverse Primer, 3’ to 5’ | Nucleotide Repeat Unit | Annealing Temp [°C] | GenBank Accession |
|-------------|------|------------|-----------------------|--------------------------|--------------------------|------------------------|---------------------|------------------|
| AUCR008b    | new  | 268/278, 268/268 | CTT CCG TAT CTC ATC AAA TA, AAA TCA GAC TCA AAT CAG TG, (CT)22, 56 | KC768707             |                           |                        |                     |                  |
| AUCR017     | new  | 363/370, 363/376 | AAA AAG GAG TTC CAC AGT AGT A, TTC AAG TCA GAA ACC CAC TAT T, (TC)4,(AC)6, 58 | KC768708             |                           |                        |                     |                  |
| AUCR050     | new  | 323/329, 329/329 | GCA GAC CTG GGT GAT ATT GA, TTA GCA GCC TAT TAC TAC GAT G, (TG)18, 60 | KC768709             |                           |                        |                     |                  |
| AUCR053     | new  | 245/257, 245/265 | AGG TTA AAT AAA GGA TGG AAA ACC AGA, GCA GGG CTA CCC TTA ACC CT, (CT)6,(TC)11, 61 | KC768710             |                           |                        |                     |                  |
| AUCR089     | new  | 229/292, 265/265 | TTC TGG GAT ATT GTG TTG CT, GGC TTT ATT CTC CCC CCT AT, (TG)5,(TG)8,(GA)10, 61 | KC768711             |                           |                        |                     |                  |
| AUCR181     | new  | 173/183, 173/206 | AGA TAA TGA AGG TCG ATG ATG AGA TTT, (GA)9, 55 | KC768715             |                           |                        |                     |                  |
| AUCR202     | new  | 170/184, 184/184 | GGG TTA AGC AAG AGA AAG A, ATG GCA CAA GGA AGT TC, (AG)18, 65 | KC768716             |                           |                        |                     |                  |
| AVD010      | new  | 323/329, 329/329 | GCA GAC CTG GGT GAT ATT GA, TTA GCA GCC TAT TAC TAC GAT G, (TG)18, 60 | KC768709             |                           |                        |                     |                  |
| AVD032      | new  | 323/329, 329/329 | GCA GAC CTG GGT GAT ATT GA, TTA GCA GCC TAT TAC TAC GAT G, (TG)18, 60 | KC768709             |                           |                        |                     |                  |
| AVD036      | new  | 119/119, 125/125 | CTT CTC TCT TTG TTC ACC CA, TTA GAC TTC CTC ACC AT, (CA)3,(GA)15,(CT)8,(AA)12, 62 | KC768718             |                           |                        |                     |                  |
| AVD044      | new  | 311/313, 302/313 | CTG TCG TAT GGT GTG GAT GAC, CCA GAC GCA ATG TGA GGC TCT C, (CT)6,(TC)11, 61 | KC768719             |                           |                        |                     |                  |
| AUCR050     | new  | 186/193, 183/186 | CCA AAG TAA CTC ACC AAC CT, CTC TCA GAC TCG TGA CTC ATC, (GA)26, 59 | KC768721             |                           |                        |                     |                  |
| AVD065      | new  | 133/135, 133/135 | CCA TAA ACC CTC TCT ACC ATC, CGT GGG GAT GAT CCA AAA TG, (TG)7, 67 | KC768722             |                           |                        |                     |                  |
| AVD082      | new  | 113/128, 113/120 | GAC CTA CTT GGA TGA GTC CT, TTG TTA TAT TGA TCT CTT, (AT)5,(GT)14, 57 | KC768723             |                           |                        |                     |                  |
| AVD095      | new  | 256/266, 256/269 | CCT GCG TCA TCT CTC GTC CTC, TAA AAG GGG TTT GTG TTA CCA TC, (GT)13, 60 | KC768724             |                           |                        |                     |                  |
| AVD103      | new  | 181/197, 197/197 | TTC CGT TAT CTT TAA TCC CC, GTT TCG CAA AAG GCT TTT AT, (CT)20, 58 | KC768725             |                           |                        |                     |                  |
| AVD104      | new  | 190/221, 190/221 | **, TGA ACG AAA TGG AAA CAT AT, ATT TTA AAT GGT GGT, (CG)4,(TG)15,(AG)22, 58 | KC768726             |                           |                        |                     |                  |
| AVD106      | new  | 183/191, 183/186 | GGA CAC ATC AGT CTT AAA TG, TGC TAC AGG GAC AAC TTA AA, (TG)15,(AG)8, 61 | KC768727             |                           |                        |                     |                  |
| AVD116      | new  | 209/217, 193/217 | ACA AAT GTG TGG TGA CAT CAG A, ATT TCC AAG TGT CAC AAA T, (TG)12,(GA)13, 59 | KC768728             |                           |                        |                     |                  |
| AVD117      | new  | 231/231, 239/241 | CGA AAG ATA GCA GGT GAG TG, GCA GTA AAG GTA TGG AAG AAT C, (GA)22, 60 | KC768729             |                           |                        |                     |                  |
| AVD120      | new  | 192/206, 196/206 | TTC ACT ATT TTT CTT GTG GAC, AAC CAG ATG TTT CTA CAG AGA, (AG)14, 57 | KC768730             |                           |                        |                     |                  |
| AVO109      | new  | 152/154, 143/154 | AAC TGG CTT TCA TCT CTC ATT C, GGT GGG GAA CTT GGT TAG T, (TC)22, 59 | KC768731             |                           |                        |                     |                  |
| AVT005b     | new  | 184/188, 184/188 | TTA GCA GCA GCA GCA GGG AGG TG, (GA)26, 59 | KC768732             |                           |                        |                     |                  |
| AVT020gat   | new  | 158/162, 158/168 | CAT ATG AGG TCT TTTTT ATT TCA CTG AAG, (AG)6,(AT)5, 63, KC768734 |                           |                           |                        |                     |                  |
| AVT106      | new  | 126/136, 126/132 | CCA TCA CTC GCG GTG GTG AT, TAT GTA TGA TGC TTA GAA CC, (ATC)8, 65 | KC768735             |                           |                        |                     |                  |
| AVT158      | new  | 313/313, 313/313 | ACG AAG TTA GGA GAA AGG AAC, GCA AAG GAG CCG GTG TTA AG, (GA)7, 62 | KC768736             |                           |                        |                     |                  |

Continued next page
| GenBank Accession | Date | Forward/Reverse | Primer Sets | Predicted Segments | Gene ID | ORF Size | Predicted Segments | Notes |
|-------------------|------|-----------------|-------------|-------------------|---------|----------|-------------------|-------|
| Supplemental Table 2. Continued. |

**AVT191**, 2004, 215/218, 215/218, TCC ACA ACT TCT ACA GGG TCG T, GGA AGA TAA CGC ACC TTG AGT TC, (ATG)7(TGG)4, 69, KC795708

**AVT226**, 2004, 298/304, 294/298, GGC TGA TTT TTA TAG TCG ATG T, TCC GAT TGA CAG TGG ATT GTT, (TCA)6..(CTT)4, 60, KC795709

**AVT386**, 2004, 229/229, 219/229, ACA ACC CAA ACA ATG CT, AAT AGA GGT CAT CGA CC, (TGA)8, 60, KC795710

**AVT436**, 2004, 149/152, 139/149, ***, ACT AAA ATG AGG GGA GAC TAG, GAG TGT ATG GAG GAG TTT GG, (ATG)9, 56, KC795711

**AVT448**, 2004, 193/193, 183/193, ACG GTG TTT GGA AGA AGA TG, GCA CTT CAA CTA ATG CTT AC, (GAT)8, 60, KC795712

**AVT517**, 2004, 229/229, 219/229, AAT CCT TCC ACT CAG AAA CT, TAC ACA AAC GAC AAC AAT GG, (GAT)6, 59, KC795713

**AVMIX03**, 2009, 145/174, 145/174, GAT ATT CCT GTT GTC ACT GC, AAT GTT CCC CAT GAA AGT CTC C, (TG)16, (AG)20, 56

**SHRSPa043**, 2009, 160/180, 164/180, TCA CTG CTC TCT TCT TGC CC, ATC TAT TGC CCT CTT GTA CTC ACT, (ATG)2GCA(TCA)14(TG)2N6(CAAA)2, 56

**SHRSPa044**, 2009, 174/181, 175/175, GCC AAC GAG GGT CAT ATC AA, CGC AAA CCA ACC GCA GA, (CTT)3(TTTAT)4, 56

**SHRSPa055**, 2009, 108/123, 117/137, TCT CTT CAT CAA CTC GAC TGC, AAC GGT ATC CAA ACG CTA AT, CC(TTCT)2(TTA)2CAA(CT)16TT(T)2, 56

**SHRSPa073**, 2009, 123/125, 125/125, CTG CTT TTC CCA CTG CTC, CCA GAA CAA ACT GAA CCA CAA, (AG)7AA(AG)2, 56

**SHRSPa081**, 2009, 218/218, 218/220, GGG CTT CAA TTC AAT CCA ATC C, TCT TCA GCA CGC CAC GAG TCT, (C)2(GA)7, 56

**SHRSPa099**, 2009, 79/79, 79/94, TCA TCC CAA TTC CCA CTC TC, AGC GGA GCA TTT TAG CG, (AG)9AA(AG)2, 56

**SHRSPa102**, 2009, 95/113, 113/119, GCC ACA AAT CTT CAA AAT ACC A, TCT TCT TGA GTG GCA GCA GC, A(GAA)6AG, 56

**SHRSPa107**, 2009, 151/165, 151/177, CGC AGT CTT CAA TAC CA, CCC CTC TTC ACT TCC AA, (AT)4N4(AC)3Ta(AC)2(CT)2(TG)2(AGA)2AA(TG)2TAT(CT)8, 56

**SHRSPa197**, 2009, 164/178, 164/164, CTC TCT CTA GCA GTG CTC GC, GGA ATT CCG CAC AGT AGC AT, (CT)10CAC(CTT)3CTG(TC)2(CTT)2, 56

**SHRSPa203**, 2009, 111/117, 109/111, ATG GTT ACA AGA ATT GGC CG, ATG AGT GCA AAA GGA CCC TG, (TA)2(CATA)3(TA)4, 56

**SHRSPa212**, 2009, 304/310, 304/304, ATT CTT CCT TCT GTC CCA AA, TGT GCC ATT AAA GAC GAC GA, (TC)5N30(CAG)2N10(GA)2(AGAGAA)3AGA(AGC)2, 56

**SHRSPa243**, 2009, 260/264, 260/264, ACA GAT GAC GGT TTT CTC GC, CTC TCA GCA TCG ACC CTT TT, (ATGATT)2CAA(AG)8, 56

**SHRSPa245**, 2009, 149/151, 149/150, CCA TGA CGG AGG TTT TTT GT, GCC AAT GGC GAT TCA GTA AT, (GT)7(T)4A(AT)3(T)5(AG)3, 56

**SHRSPa249**, 2009, 272/276, 270/274, CCA GAA GCT GGC AAT CTA GC, CCA AAC GGG TTG TAA TGG TA, (TA)3TT(TA)9, 56

**SHRSPa262**, 2009, 192/195, 192/192, GGG GAA TCC ACG GCA T, TGG AGG GGA TTC TCC TCC TT, (CTT)3(CTC)4CTGCT(TCC)3, 56

**SHRSPa274**, 2009, 132/139, 139/139, GTG AGT CTG TAA CGC GCA GA, GCT ACA AGA TGC AGC AAC AA, (TC)21TT(T)2, 56

**SHRSPa285**, 2009, 255/264, 255/256, ACC GGT CTG TCG GAA ATC AG, GCC AAC AGT ACA TTC CCC AT, (AT)2(AGG)7(AAG)6, 56

6 J. Amer. Soc. Hort. Sci. 144(5):1–18. 2019.
Supplemental Table 3. Position of genetic markers on the twelve avocado linkage groups. Quantitative trait loci (QTLs) are highlighted in bold if inferred by Interval Mapping and underlined if inferred by Kruskal-Wallis analysis. Phenotypic traits are abbreviated to A (alpha-tocopherol), B (beta-sitosterol), C (carotenoids), CP (canopy diameter), H (tree height), T (trunk diameter), F (flowering type).

| group 1 |
|-----------------|-----------------|
| AVD028 | 0 |
| SHRSPaS003949 | 1.496 |
| SHRSPaS004383 | 1.782 |
| SHRSPaS001411 | 2.956 |
| SHRSPaS006205 | 3.596 |
| SHRSPaS002267 | 3.596 |
| SHRSPaS001479 | 4.75 C |
| SHRSPaS005923 | 4.817 C |
| SHRSPaS003997 | 4.879 |
| SHRSPaS003077 | 7.016 |
| SHRSPa212 | 7.442 |
| SHRSPaS001835 | 7.905 |
| SHRSPaS003122 | 7.905 |
| SHRSPaS003937 | 9.59 |
| SHRSPaS001255 | 11.981 |
| SHRSPaS003341 | 14.167 |
| SHRSPaS002400 | 14.425 |
| SHRSPaS001497 | 14.46 |
| SHRSPaS001760 | 15.963 |
| SHRSPaS001015 | 16.878 |
| SHRSPaS002216 | 16.981 |
| SHRSPaS002070 | 17.657 |
| SHRSPaS004066 | 18.294 |
| SHRSPaS003503 | 19.677 |
| SHRSPaS004945 | 21.771 |
| SHRSPaS001118 | 22.436 |
| SHRSPaS002191 | 24.148 |
| SHRSPaS003028 | 24.637 |
| SHRSPaS004904 | 26.288 |
| SHRSPaS002246 | 26.591 |
| SHRSPaS005298 | 28.294 |
| SHRSPaS002150 | 28.862 |
| SHRSPaS002075 | 32.832 |
| SHRSPaS003332 | 34.45 |
| SHRSPaS003987 | 34.69 |
| SHRSPaS001253 | 35.563 |
| SHRSPaS002125 | 35.592 |
| SHRSPaS003741 | 35.592 |
| SHRSPaS002478 | 36.73 |
| SHRSPaS003445 | 37.667 |
| SHRSPaS001353 | 38.098 |
| SHRSPaS004287 | 38.391 |
| SHRSPaS001130 | 38.614 |
| SHRSPaS004019 | 39.033 |
| SHRSPaS006916 | 39.476 |
| SHRSPaS003156 | 39.839 |
| SHRSPaS001286 | 40.345 |
| SHRSPaS006979 | 40.706 |
| SHRSPaS002056 | 41.216 |
| SHRSPaS005224 | 41.435 |
| SHRSPaS002667 | 41.923 |

Continued next page
| Gene ID         | Log2 Fold Change |
|----------------|-----------------|
| SHRSPaS003632  | 11.658          |
| SHRSPaS006670  | 12.004          |
| SHRSPaS002703  | 12.4            |
| SHRSPaS001771  | 12.549          |
| SHRSPaS001999  | 12.869          |
| SHRSPaS001453  | 12.909          |
| SHRSPaS004553  | 13.206          |
| SHRSPaS002686  | 13.738          |
| SHRSPaS003305  | 13.743          |
| SHRSPaS001552  | 13.994          |
| SHRSPaS004994  | 14.025          |
| SHRSPaS003599  | 14.239          |
| SHRSPaS003086  | 14.474          |
| SHRSPaS006435  | 14.68           |
| SHRSPaS003810  | 15.109          |
| SHRSPaS003496  | 15.362          |
| SHRSPaS002286  | 16.009          |
| SHRSPaS003528  | 16.178          |
| SHRSPaS001530  | 16.293          |
| SHRSPaS002014  | 16.946          |
| SHRSPaS003513  | 16.959          |
| SHRSPaS003206  | 17.12           |
| SHRSPaS003090  | 17.361          |
| SHRSPaS002731  | 17.361          |
| SHRSPaS001831  | 17.825          |
| SHRSPaS001883  | 17.895          |
| SHRSPaS002026  | 18.202          |
| SHRSPaS002140  | 18.202          |
| SHRSPaS001464  | 18.265          |
| SHRSPaS001404  | 18.436          |
| SHRSPaS004250  | 19.154          |
| SHRSPaS004471  | 19.467          |
| SHRSPaS002018  | 19.777          |
| SHRSPaS001343  | 20.195          |
| SHRSPaS003209  | 20.195          |
| SHRSPaS002961  | 20.242          |
| SHRSPaS005301  | 20.242          |
| SHRSPaS004677  | 20.43           |
| SHRSPaS001669  | 20.444          |
| SHRSPaS004053  | 20.788          |
| SHRSPaS005014  | 20.788          |
| SHRSPaS004298  | 20.8            |
| SHRSPaS004502  | 20.845          |
| SHRSPaS004944  | 21.024          |
| SHRSPaS004999  | 21.264          |
| SHRSPaS004772  | 21.264          |
| SHRSPaS003366  | 21.641          |
| CYP890         | 21.732          |
| SHRSPaS006206  | 21.773          |
| SHRSPaS003501  | 22.025          |
| SHRSPaS004303  | 22.36           |
| SHRSPaS002021  | 22.534          |
| SHRSPaS001199  | 22.712          |
| CYP967         | 22.764          |
| SHRSPaS001408  | 22.862          |
| SHRSPaS004868  | 22.99           |
| SHRSPaS002009  | 23.043          |
| SHRSPaS002890  | 23.065          |
| SHRSPaS001306  | 23.151          |
| SHRSPaS000235  | 84.028          |
| SHRSPaS001196  | 84.155          |
| SHRSPaS001842  | 84.331          |
| SHRSPaS001214  | 84.751          |
| SHRSPaS004934  | 85.315          |
| SHRSPaS004517  | 85.652          |
| SHRSPaS003315  | 86.869          |
| SHRSPaS002266  | 88.035          |
| SHRSPaS003269  | 89.143          |
| SHRSPaS001164  | 90.035          |
| SHRSPaS001330  | 91.25           |
| SHRSPaS001955  | 93.466          |
| AVD104         | 93.897          |
| SHRSPaS006607  | 94.461          |
| SHRSPaS002221  | 94.967          |
| SHRSPaS003054  | 95.261          |
| SHRSPaS002061  | 97.05           |
| SHRSPaS004896  | 98.516          |
| SHRSPaS002904  | 98.832          |
| SHRSPaS002076  | 99.926          |
| SHRSPaS003187  | 99.953          |
| SHRSPaS001229  | 100.777         |
| SHRSPaS001526  | 101.82          |
| SHRSPaS002118  | 101.82          |
| SHRSPaS001587  | 102.876         |
| SHRSPaS001873  | 102.887         |
| SHRSPaS002800  | 103.034         |
| SHRSPaS003802  | 103.911         |
| SHRSPaS003920  | 104.432         |
| PDX2_549       | 106.045         |
| SHRSPaS003269  | 107.05          |
| SHRSPaS001164  | 108.035         |
| SHRSPaS001330  | 109.25          |
| SHRSPaS001955  | 110.466         |
| AVD104         | 111.897         |
| SHRSPaS006607  | 112.034         |
| SHRSPaS002221  | 113.911         |
| SHRSPaS003054  | 114.432         |
| SHRSPaS002061  | 115.045         |
| SHRSPaS004896  | 116.516         |
| SHRSPaS002904  | 117.832         |
| SHRSPaS002076  | 118.926         |
| SHRSPaS003187  | 119.953         |
| SHRSPaS001229  | 120.777         |
| SHRSPaS001526  | 121.82          |
| SHRSPaS002118  | 122.876         |
| SHRSPaS001587  | 123.876         |
| SHRSPaS001873  | 124.887         |
| SHRSPaS002800  | 125.034         |
| SHRSPaS003802  | 126.911         |
| SHRSPaS003920  | 127.432         |
| PDX2_549       | 128.045         |

Continued next page

**Continued next page**
| Gene ID         | Value  |
|----------------|--------|
| SHRSPaS005940  | 23.23  |
| SHRSPaS002450  | 23.33  |
| SHRSPaS002909  | 23.71  |
| CYP1085        | 23.85  |
| SHRSPaS006613  | 24.45  |
| SHRSPaS006065  | 24.58  |
| AVT191         | 24.72  |
| SHRSPaS002171  | 25.03  |
| SHRSPaS002428  | 25.03  |
| SHRSPaS001615  | 25.04  |
| SHRSPaS00262   | 25.48  |
| SHRSPaS001822  | 25.57  |
| SHRSPaS004881  | 25.85  |
| SHRSPaS006379  | 25.99  |
| SHRSPaS004176  | 26.26  |
| SHRSPaS002196  | 26.67  |
| SHRSPaS006718  | 26.69  |
| SHRSPaS002374  | 26.72  |
| SHRSPaS008201  | 27.08  |
| SHRSPaS002787  | 27.08  |
| SHRSPaS002203  | 27.50  |
| SHRSPaS001976  | 27.60  |
| SHRSPaS003472  | 28.41  |
| SHRSPaS002183  | 28.70  |
| SHRSPaS001998  | 28.90  |
| SHRSPaS001078  | 29.18  |
| SHRSPaS001037  | 29.27  |
| SHRSPaS003743  | 29.49  |
| SHRSPaS002313  | 30.26  |
| SHRSPaS002762  | 30.34  |
| SHRSPaS004715  | 30.52  |
| SHRSPaS002561  | 30.52  |
| SHRSPaS005503  | 30.92  |
| SHRSPaS001768  | 30.99  |
| SHRSPaS001593  | 31.30  |
| SHRSPaS006845  | 31.54  |
| SHRSPaS001184  | 31.85  |
| SHRSPaS004000  | 31.89  |
| SHRSPaS002134  | 31.89  |
| SHRSPaS001029  | 31.91  |
| SHRSPaS002659  | 31.94  |
| SHRSPaS005876  | 32.28  |
| SHRSPaS003751  | 32.42  |
| SHRSPaS003433  | 32.82  |
| SHRSPaS001661  | 32.90  |
| SHRSPaS003294  | 33.23  |
| SHRSPaS001769  | 33.47  |
| SHRSPaS003627  | 33.58  |
| SHRSPaS003776  | 33.60  |
| AVD006         | 33.66  |
| SHRSPaS004324  | 34.07  |
| SHRSPaS006019  | 34.31  |
| SHRSPaS006959  | 34.35  |
| SHRSPaS002611  | 34.62  |

Continued next page
### Supplemental Table 3. Continued.

| Accession      | Value  |
|----------------|--------|
| SHRSPaS005346  | 55.582 |
| AUCR418        | 56.929 |
| SHRSPaS001523  | 57.267 |
| SHRSPaS001418  | 58.203 |
| SHRSPaS002539  | 58.75  |
| SHRSPaS006032  | 59.165 |
| SHRSPaS004951  | 60.469 |
| SHRSPaS003425  | 60.742 |
| SHRSPaS003554  | 61.27  |
| SHRSPaS004422  | 61.483 |

**Group 3**

| Accession      | Value  |
|----------------|--------|
| SHRSPaS003453  | 0.82   |
| SHRSPaS001761  | 2.938  |
| HRPT1_551      | 4.297  |
| HPT1_514       | 5.196  |
| SHRSPaS002447  | 7.242  |
| SHRSPaS002426  | 7.968 A|
| SHRSPaS001620  | 7.968 A|
| SHRSPaS003259  | 8.562 A|
| SHRSPaS003589  | 8.847 A|
| SHRSPaS001705  | 10.078 A|
| SHRSPaS00245   | 12.359 A|
| SHRSPaS006564  | 12.736 A|
| SHRSPaS004209  | 12.998 A|
| SHRSPaS003314  | 13.349 A|
| SHRSPaS002204  | 14.43  A|
| SHRSPaS002658  | 14.63  A|
| SHRSPaS004388  | 15.363 A|
| SHRSPaS005529  | 15.648 A|
| SHRSPaS003787  | 16.304 A|
| SHRSPaS004634  | 16.304 A|
| SHRSPaS001323  | 16.785 A|
| SHRSPaS001365  | 17.332 A|
| SHRSPaS004338  | 18.601 A|
| SHRSPaS005013  | 19.334 |
| SHRSPaS001566  | 20.682 A|
| SHRSPaS006054  | 24.144 B|
| SHRSPaS003561  | 25.5 A  |
| SHRSPaS004954  | 26.428 |
| SHRSPaS003645  | 27.638 A|
| SHRSPaS002817  | 28.082 |
| SHRSPaS003120  | 29.223 |
| SHRSPaS006371  | 29.223 |
| SHRSPaS001781  | 31.467 |
| SHRSPaS004933  | 33.176 |
| SHRSPaS003760  | 35.949 |
| SHRSPaS003017  | 37.112 |
| SHRSPaS005275  | 37.112 |
| SHRSPaS004350  | 38.714 |
| SHRSPaS005919  | 40.402 |
| SHRSPaS003739  | 42.778 |
| SHRSPaS005977  | 42.778 |
| SHRSPaS004018  | 43.234 |
| SHRSPaS006088  | 43.779 |
| SHRSPaS002393  | 43.793 |

Continued next page
| Gene   | Value  |
|--------|--------|
| SHRSpaS002825 | 92.984 |
| SHRSpaS001043 | 93.063 |
| SHRSpaS002042 | 94.008 |
| SHRSpaS001688 | 94.008 |
| SHRSpaS003869 | 94.203 C |
| SHRSpaS005910 | 94.666 |
| SHRSpaS003959 | 94.996 |
| SHRSpaS003708 | 94.996 |
| AVT020gat   | 95.543 |
| SHRSpaS004268 | 95.871 |
| SHRSpaS002912 | 96.08 |
| SHRSpaS001297 | 96.08 |
| SHRSpaS001349 | 96.114 |
| SHRSpaS003557 | 96.609 |
| SHRSpaS005922 | 97.148 |
| SHRSpaS003149 | 97.155 |
| SHRSpaS005725 | 97.867 |
| AVD107     | 98.297 B |
| SHRSpaS003161 | 98.297 B |
| SHRSpaS003012 | 98.538 C |
| SHRSpaS002578 | 99.379 |
| SHRSpaS004446 | 99.379 |
| SHRSpaS001881 | 99.906 |
| SHRSpaS003582 | 100.455 |
| SHRSpaS001316 | 100.455 |
| SHRSpaS001570 | 100.782 |
| SHRSpaS002153 | 102.052 |
| SHRSpaS004129 | 102.14 |
| SHRSpaS002786 | 102.494 |
| SHRSpaS001036 | 102.995 |
| SHRSpaS004561 | 103.711 |
| SHRSpaS005938 | 103.959 |
| SHRSpaS004802 | 104.375 |
| SHRSpaS002129 | 104.575 |
| SHRSpaS004329 | 104.764 |
| SHRSpaS004025 | 105.067 |
| SHRSpaS001908 | 105.659 |
| SHRSpaS001734 | 105.849 |
| SHRSpaS003405 | 106.226 |
| SHRSpaS001569 | 106.226 |
| SHRSpaS004323 | 106.98 |
| SHRSpaS006755 | 107.221 |
| SHRSpaS002047 | 107.833 |
| SHRSpaS003165 | 108.257 |
| AVD026      | 108.567 |
| SHRSpaS003582 | 108.602 |
| SHRSpaS004906 | 109.305 |
| SHRSpaS004540 | 109.725 |
| SHRSpaS005002 | 109.725 |
| SHRSpaS003623 | 110.087 |
| SHRSpaS002610 | 110.49 |
| cafl3_SNP745 | 110.505 |
| SHRSpaS003705 | 111.025 |
| SHRSpaS003191 | 111.782 |
| SHRSpaS001121 | 111.782 |
| SHRSpaS001298 | 111.813 |
| SHRSpaS004145 | 111.991 |
| SHRSpaS001695 | 112.477 |

**Group 4**

| Gene   | Value  |
|--------|--------|
| SHRSpaS003316 | 112.658 |
| SHRSpaS004215 | 112.867 |
| SHRSpaS004550 | 112.905 |
| SHRSpaS003378 | 114.667 |
| cafl3_SNP1012 | 115.113 |
| cafl3_SNP850  | 116.291 |
| cafl3_SNP1099 | 118.125 |
| ZDS_SNP_228   | 119.741 |
| cafl3_SNP814  | 121.125 |
| SHRSpaS004274 | 0        |
| SHRSpaS003489 | 3.71     |
| SHRSpaS002201 | 4.519    |
| SHRSpaS003694 | 5.476    |
| SHRSpaS001086 | 5.505    |
| SHRSpaS002947 | 5.505    |
| SHRSpaS001428 | 6.431    |
| SHRSpaS001224 | 6.431    |
| SHRSpaS002527 | 6.487    |
| SHRSpaS003560 | 7.928    |
| SHRSpaS002713 | 9.719    |
| SHRSpaS002073 | 10.118   |
| SHRSpaS003412 | 10.492   |
| SHRSpaS00249  | 12.654   |
| SHRSpaS002293 | 13.388   |
| SHRSpaS005507 | 15.622   |
| SHRSpaS004673 | 15.622   |
| SHRSpaS005574 | 17.487   |
| SHRSpaS003761 | 17.487   |
| SHRSpaS001966 | 18.916   |
| SHRSpaS001856 | 20.868   |
| SHRSpaS003225 | 23.349   |
| SHRSpaS002296 | 24.762   |
| SHRSpaS004699 | 27.68    |
| SHRSpaS001416 | 28.68    |
| SHRSpaS001734 | 29.556   |
| SHRSpaS005878 | 30.148   |
| SHRSpaS004065 | 30.5     |
| SHRSpaS004400 | 32.508   |
| SHRSpaS005892 | 33.172   |
| SHRSpaS003174 | 33.727   |
| SHRSpaS004731 | 34.796   |
| SHRSpaS002860 | 34.796   |
| SHRSpaS002120 | 35.133   |
| SHRSpaS003670 | 35.165   |
| SHRSpaS003904 | 35.947   |
| SHRSpaS003418 | 36.926   |
| SHRSpaS001309 | 37.139   |
| SHRSpaS004865 | 38.006   |
| SHRSpaS002697 | 38.573   |
| SHRSpaS003963 | 38.573   |
| SHRSpaS001020 | 38.656   |
| SHRSpaS002062 | 40.595   |
| SHRSpaS002151 | 42.569   |
| SHRSpaS0081  | 47.675 A |
| SHRSpaS003355 | 49.264   |
| SHRSpaS003210 | 49.264   |
### Supplemental Table 3. Continued.

| Gene ID     | Value |
|-------------|-------|
| SHRSaP004354| 23.46 |
| SHRSaP002409| 24.309|
| SHRSaP002124| 24.309|
| SHRSaP002253| 28.261|
| SHRSaP003370| 30.437|
| SHRSaP00285 | 30.495|
| SHRSaP002479| 32.822|
| SHRSaP001305| 33.132|
| SHRSaP0003250| 33.132|
| SHRSaP002167| 33.944|
| AUCR053     | 35.318|
| SHRSaP006160| 35.832|
| SHRSaP002639| 36.965|
| SHRSaP002219| 36.965|
| SHRSaP001192| 37.198|
| SHRSaP004357| 37.198|
| SHRSaP003290| 37.962|
| SHRSaP002399| 37.962|
| SHRSaP002862| 38.557|
| SHRSaP002756| 39.973|
| SHRSaP001387| 40.935|
| SHRSaP003744| 41.2|
| SHRSaP003177| 41.2|
| SHRSaP0004646| 42.222|
| SHRSaP001843| 42.222|
| SHRSaP001847| 43.704|
| SHRSaP003699| 44.482|
| SHRSaP001168| 44.892|
| SHRSaP004331| 45.536|
| SHRSaP004636| 46.567|
| SHRSaP003340| 46.567|
| SHRSaP001068| 46.994|
| SHRSaP003950| 47.079|
| SQS913      | 47.547|
| SHRSaP001953| 47.574|
| SQS843      | 47.689|
| SQS769      | 47.689|
| SHRSaP003345| 47.886|
| SHRSaP003134| 47.886|
| SHRSaP002676| 48.351|
| SHRSaP002297| 48.351|
| SHRSaP001405| 49.074|
| SHRSaP005955| 49.712|
| SHRSaP00107 | 50.306|
| SHRSaP003308| 51.546|
| SHRSaP001104| 51.922|
| SHRSaP002300| 52.798|
| SHRSaP001993| 53.855|
| SHRSaP001046| 54.152|
| SHRSaP003944| 54.152|
| SHRSaP003738| 55.035|
| SHRSaP001246| 55.99|
| SHRSaP004482| 57.458|
| SHRSaP003457| 57.458|
| 58          |     |
| SHRSaP002060| 58.814|
| SHRSaP002532| 59.919|
| SHRSaP002792| 60.217|
| SHRSaP002430| 62.505|

---

**Continued next page**
| Gene            | AUCR    |
|-----------------|---------|
| SHRSPaS003998   | 63.551  |
| AUCR008b        | 64.489  |
| SHRSPaS003800   | 64.817  |
| SHRSPaS002350   | 65.429  |
| SHRSPaS003637   | 66.141  |
| SHRSPaS003184   | 66.141  |
| SHRSPaS002416   | 66.886  |
| SHRSPaS003803   | 67.856  |
| AUCR181         | 68.798  |
| SHRSPaS003585   | 68.814  |
| SHRSPaS004203   | 69.545  |
| SHRSPaS003155   | 71.195  |
| SHRSPaS006171   | 73.836  |
| SHRSPaS00171    | 74.913  |
| SHRSPaS006026   | 77.238  |
| SHRSPaS002520   | 77.238  |
| SHRSPaS002837   | 79.158  |
| SHRSPaS001935   | 79.307  |
| SHRSPaS005098   | 79.552  |
| SHRSPaS001372   | 79.961  |
| SHRSPaS001711   | 79.995  |
| SHRSPaS001743   | 80.872  |
| SHRSPaS003297   | 81.644  |
| SHRSPaS0061195  | 81.644  |
| SHRSPaS001950   | 81.685  |
| SHRSPaS002905   | 82.637  |
| SHRSPaS003239   | 82.796  |
| AVDO82          | 83.12   |
| SHRSPaS003595   | 83.44   |
| SHRSPaS001779   | 83.561  |
| SHRSPaS004622   | 84.326  |
| SHRSPaS001654   | 84.624  |
| SHRSPaS003881   | 85.174  |
| SHRSPaS003415   | 85.174  |
| SHRSPaS001468   | 85.277  |
| SHRSPaS001671   | 85.662  |
| SHRSPaS005580   | 86.313  |
| SHRSPaS002714   | 86.313  |
| SHRSPaS001783   | 86.831  |
| SHRSPaS001267   | 87.253  |
| SHRSPaS002631   | 87.314  |
| SHRSPaS002235   | 87.937  |
| SHRSPaS005804   | 87.937  |
| SHRSPaS004575   | 87.942  |
| SHRSPaS001350   | 88.51   |
| SHRSPaS001306   | 89.121  |
| SHRSPaS001099   | 89.578  |
| SHRSPaS001478   | 90.057  |
| SHRSPaS006151   | 90.292  |
| SHRSPaS002783   | 90.798  |
| SHRSPaS002894   | 90.839  |
| SHRSPaS001683   | 91.68   |
| SHRSPaS001287   | 91.805  |
| SHRSPaS003098   | 92.094  |
| SHRSPaS002282   | 94.01   |
| SHRSPaS003136   | 94.101  |
| SHRSPaS005970   | 94.101  |

Continued next page
| Gene          | Value  |
|--------------|--------|
| SHRSPaS001538| 43.101 C |
| SHRSPaS003598| 44     |
| SHRSPaS004251| 44     |
| SHRSPaS001664| 44.722 |
| SHRSPaS001503| 45.129 |
| SHRSPaS001454| 46.13 C |
| SHRSPaS004712| 46.891 |
| SHRSPaS003733| 47.914 |
| SHRSPaS001664| 47.914 |
| SHRSPaS001162| 48.38 C |
| SHRSPaS001501| 49.827 |
| SHRSPaS003639| 49.827 |
| SHRSPaS003569| 50.771 |
| SHRSPaS001329| 50.771 |
| SHRSPaS004252| 51.504 |
| SHRSPaS003085| 52.282 |
| SHRSPaS001536| 53.217 C |
| SHRSPaS002728| 53.586 |
| SHRSPaS002154| 54.131 |
| SHRSPaS004713| 54.717 C |
| SHRSPaS006785| 55.63  |
| SHRSPaS006696| 56.019 |
| SHRSPaS002669| 57.187 |
| SHRSPaS002346| 57.548 C |
| SHRSPaS004674| 57.642 |
| SHRSPaS002735| 57.657 |
| SHRSPaS005679| 58.06  |
| SHRSPaS002031| 58.214 |
| SHRSPaS001544| 58.819 |
| SHRSPaS003653| 59.557 |
| SHRSPaS004439| 59.557 |
| AVT517       | 60.034 |
| SHRSPaS003264| 60.159 |
| SHRSPaS002852| 60.758 C |
| SHRSPaS004639| 61.419 |
| LUT5_SNP_1351| 63.284 |
| SHRSPaS005466| 64.024 |
| SHRSPaS001710| 66.054 |
| SHRSPaS002169| 66.054 |
| SHRSPaS001516| 66.303 |
| SHRSPaS001022| 67.002 |
| SHRSPaS004093| 67.613 |
| SHRSPaS002543| 68.008 |
| SHRSPaS003514| 68.611 |
| SHRSPaS003812| 71.592 |
| SHRSPaS003990| 71.716 |
| SHRSPaS002744| 71.959 |
| VTE4_1035    | 73.171 |
| VTE4_1257    | 73.181 |
| SHRSPaS006514| 74.199 |
| SHRSPaS001676| 76.112 |
| VTE4_1068    | 76.567 |
| group 7      |        |
| SHRSPaS002765| 0      |
| SHRSPaS003542| 0      |
| SHRSPaS002055| 0      |
| SHRSPaS002341| 1.43   |

Continued next page
| Gene ID     | Value  |
|------------|--------|
| SHRSPaS004243 | 66.458  |
| SHRSPaS001165 | 66.64   |
| SHRSPaS001155 | 67.112  |
| SHRSPaS003943 | 68.947  |
| SHRSPaS004825 | 70.072  |
| SHRSPaS004855 | 70.072  |
| SHRSPaS002421 | 71.113  |
| SHRSPaS002768 | 71.113  |
| SHRSPaS001397 | 71.324  |
| SHRSPaS003943 | 73.063  |
| SHRSPaS004942 | 73.206  |
| SHRSPaS001685 | 74.199  |
| SHRSPaS004326 | 75.207  |
| SHRSPaS001273 | 75.207  |
| SHRSPaS005004 | 75.407  |
| SHRSPaS003537 | 76.283  |
| SHRSPaS003087 | 76.323  |
| SHRSPaS004064 | 76.814  |
| SHRSPaS001334 | 77.288  |
| SHRSPaS005939 | 77.71   |
| SHRSPaS002405 | 82.25   |
| SHRSPaS002440 | 84.516  |
| SHRSPaS001561 | 87.923  |
| SHRSPaS003426 | 92.232  |
| SHRSPaS003665 | 93.013  |
| SHRSPaS002041 | 96.47   |
| SHRSPaS006248 | 100.563 |
| SHRSPaS001417 | 102.78  |

| Gene ID     | Value  |
|------------|--------|
| SHRSPaS002313 | 18.26  |
| SHRSPaS003358 | 18.412 |
| SHRSPaS004398 | 19.77  |
| AVT038      | 20.061 |
| SHRSPaS004095 | 20.379 |
| SHRSPaS003107 | 21.122 |
| SHRSPaS004942 | 23.729 |
| SHRSPaS004328 | 24.552 |
| SHRSPaS006517 | 24.818 |
| SHRSPaS001081 | 26.287 |
| SHRSPaS006531 | 27.54  |
| SHRSPaS001685 | 74.199  |
| SHRSPaS004326 | 75.207  |
| SHRSPaS001273 | 75.207  |
| SHRSPaS005004 | 75.407  |
| SHRSPaS003537 | 76.283  |
| SHRSPaS003087 | 76.323  |
| SHRSPaS004064 | 76.814  |
| SHRSPaS001334 | 77.288  |
| SHRSPaS005939 | 77.71   |
| SHRSPaS002405 | 82.25   |
| SHRSPaS002440 | 84.516  |
| SHRSPaS001561 | 87.923  |
| SHRSPaS003426 | 92.232  |
| SHRSPaS003665 | 93.013  |
| SHRSPaS002041 | 96.47   |
| SHRSPaS006248 | 100.563 |
| SHRSPaS001417 | 102.78  |

| Gene ID     | Value  |
|------------|--------|
| SHRSPaS001273 | 29.771  |
| SHRSPaS002782 | 29.727  |
| SHRSPaS003537 | 30.736  |
| SHRSPaS004064 | 32.272  |
| SHRSPaS002211 | 36.536  |
| SHRSPaS001334 | 39.587  |
| SHRSPaS005939 | 42.546  |
| SHRSPaS002082 | 42.557  |
| SHRSPaS001674 | 43.082  |
| SHRSPaS001178 | 43.098  |
| SHRSPaS002529 | 43.098  |
| SHRSPaS001549 | 43.958  |
| SHRSPaS002812 | 43.958  |
| SHRSPaS001936 | 43.958  |
| SHRSPaS002812 | 43.958  |
| SHRSPaS003140 | 44.416  |
| SHRSPaS001561 | 44.416  |
| SHRSPaS003426 | 44.416  |
| SHRSPaS003665 | 44.416  |
| SHRSPaS002041 | 46.3    |
| SHRSPaS006248 | 46.998  |
| SHRSPaS001417 | 47.706  |
| SHRSPaS001549 | 49.552  |
| SHRSPaS001561 | 49.552  |
| SHRSPaS003420 | 49.552  |
| SHRSPaS0002405 | 50.838  |
| SHRSPaS004543 | 50.838  |
| SHRSPaS0002405 | 52.642  |
| SHRSPaS0002405 | 53.02   |
| SHRSPaS0002405 | 53.441  |
| SHRSPaS0002405 | 53.859  |
| SHRSPaS0002405 | 54.435  |
| SHRSPaS0002405 | 54.094  |
| SHRSPaS0002405 | 56.834  |
| SHRSPaS0002405 | 56.834  |
| SHRSPaS0002405 | 57.098  |
| SHRSPaS0002405 | 57.935  |
| SHRSPaS0002405 | 58.62   |
| SHRSPaS0002405 | 59.001  |
| SHRSPaS0002405 | 59.554  |
| SHRSPaS0002405 | 60.417  |
| SHRSPaS0002405 | 60.885  |
| SHRSPaS0002405 | 61.085  |
| SHRSPaS0002405 | 62.094  |
| SHRSPaS0002405 | 62.606  |
| SHRSPaS0002405 | 63.012  |
| SHRSPaS0002405 | 63.921  |
| SHRSPaS0002405 | 65.152  |
| SHRSPaS0002405 | 66.105  |
| SHRSPaS0002405 | 68.303  |

Continued next page
| SHRS001638 | 90.057 |
|------------|--------|
| SHRS001646 | 91.135 |
| SHRS001672 | 92.098 |
| SHRS004539 | 92.928 |
| SHRS001021 | 92.986 |
| SHRS001139 | 92.986 |
| SHRS006403 | 96.073 |

**group 9**

| SHRS001638 | 0 |
|------------|---|
| SHRS0243   | 2.563 |
| SHRS002814 | 4.658 |
| SHRS005406 | 5.705 |
| SHRS003487 | 7.177 |
| SHRS001914 | 9.174 |
| SHRS003251 | 9.174 |
| SHRS001580 | 9.501 |
| SHRS004956 | 11.214 |
| SHRS001421 | 11.636 |
| SHRS003573 | 12.287 |
| SHRS004831 | 12.287 |
| SHRS003344 | 13.353 |
| SHRS005963 | 13.487 |
| SHRS001237 | 13.826 |
| SHRS004520 | 15.419 |
| SHRS002439 | 15.419 |
| SHRS002709 | 16.501 |
| SHRS005735 | 18.52 |
| SHRS002538 | 21.062 |
| FPS1135    | 21.199 |
| SHRS003427 | 22.533 |
| SHRS006483 | 22.753 |
| SHRS001395 | 23.351 |
| SHRS001013 | 23.351 |
| SHRS002574 | 24.633 |
| SHRS001284 | 24.888 |
| SHRS001364 | 28.945 |
| SHRS002012 | 28.996 |
| SHRS005992 | 30.586 |
| SHRS002544 | 33.984 |
| PDH1_775   | 34.29  |
| PDH1_1001  | 34.316 |
| PDH1_941   | 34.412 |
| SHRS003093 | 36.045 |

**group 10**

| SHRS001638 | 0 |
|------------|---|
| SHRS002439 | 0.664 |
| SHRS002777 | 0.773 |
| SHRS005151 | 2.409 |
| SHRS001785 | 3.195 |
| SHRS004226 | 4.397 |
| SHRS001911 | 4.891 |
| SHRS001463 | 7.541 |
| SHRS004821 | 8.803 F |
| SHRS001692 | 9.796 F |
| SHRS001876 | 11.626 F |
| SHRS005289 | 15.671 F |
| SHRS001228 | 17.08 F |
| SHRS004991 | 18.179 F |
| SHRS001648 | 19.193 F |
| SHRS006707 | 19.909 F |
| SHRS002720 | 23.273 F |

Continued next page
| SHRSPaS002034 | 24.112 F | SHRSPaS002034 | 15.75 |
| ------------- | -------- | -------------- | ------ |
| DXPS1_SNP1593  | 25.756 F | SHRSPaS001709  | 15.75 |
| SHRSPaS002337  | 25.91 F  | SHRSPaS001235  | 16.485 |
| SHRSPaS002152  | 26.808 F | SHRSPaS001977  | 18.371 |
| DXPS1_SNP1328  | 26.876   | SHR006118      | 19.875 |
| SHRSPaS002994  | 30.261 F | SHRSPaS001494  | 20.982 |
| SHRSPaS001393  | 31.674 F | SHRSPaS004272  | 21.992 CP |
| SHRSPaS000395  | 32.854 F | SHRSPaS002263  | 23.552 |
| SHRSPaS003892  | 33.914 F | SHRSPaS004983  | 24.169 |
| SHRSPaS001500  | 35.058 F | SHRSPaS00261   | 24.275 |
| SHRSPaS002920  | 35.434 F | AVD116         | 25.043 |
| SHRSPaS001512  | 36.688 F | SHRSPaS004232  | 25.151 CP |
| SHRSPaS004112  | 36.71 F  | VTCI_1121      | 26.8   |
| SHRSPaS002197  | 37.499 F | VTCI_1084      | 27.267 |
| SHRSPaS001256  | 37.532 F | VTCI_1187      | 28.072 |
| SHRSPaS001931  | 38.485 F | SHRSPaS001665  | 29.627 |
| SHRSPaS003940  | 40.615 F | SHRSPaS001649  | 29.65 |
| SHRSPaS003414  | 42.421 F | SHRSPaS002491  | 29.789 |
| SHRSPaS002815  | 44.139 F | SHRSPaS003895  | 30.222 |
| SHRSPaS004380  | 44.68 F  | SHRSPaS002621  | 30.634 |
| AVD010         | 45.094 F | SHRSPaS001260  | 30.855 |
| SHRSPaS006391  | 45.452 F | SHRSPaS006777  | 31.904 |
| SHRSPaS002938  | 45.511 F | SHRSPaS003138  | 32.112 |
| SHRSPaS002466  | 45.617 F | SHRSPaS006702  | 32.209 |
| SHRSPaS002742  | 46.511 F | SHRSPaS001151  | 32.761 |
| SHRSPaS001577  | 47.115 F | SHRSPaS004039  | 32.761 |
| SHRSPaS004654  | 47.196 F | SHRSPaS002602  | 33.055 A |
| SHRSPaS001390  | 47.349 F | M1022          | 34.256 |
| SHRSPaS001445  | 47.949 F | SHRSPaS002545  | 34.413 A |
| SHRSPaS004170  | 48.06 F  | SHRSPaS003786  | 34.81 |
| SHRSPaS004995  | 50.098 F | SHRSPaS001352  | 34.984 A |
| SHRSPaS002997  | 50.098 F | SHRSPaS003977  | 37.506 |
| SHRSPaS002903  | 51.317 F | SHRSPaS004049  | 37.609 |
| SHRSPaS004214  | 53.308 F | SHRSPaS001989  | 38.343 |
| SHRSPaS004955  | 53.308 F | SHRSPaS002813  | 38.474 |
| SHRSPaS001432  | 55.544 F | SHRSPaS003082  | 39.352 |
| SHRSPaS002875  | 59.658 F | SHRSPaS003374  | 40.396 A |
| SHRSPaS006283  | 63.9     | SHRSPaS001213  | 41.746 |
| SHRSPaS002351  | 65.464   | SHRSPaS005008  | 42.21 |
| SHRSPaS003095  | 67.268   | SHRSPaS002803  | 42.863 |
| SHRSPaS004747  | 68.473   | SHRSPaS002011  | 43.178 A |
| SHR001122      | 15.593   | SHRSPaS002403  | 45.033 |
| SHRSPaS003442  | 0        | SHRSPaS001789  | 46.037 |
| AVD022         | 3.311    | SHRSPaS001429  | 46.57 |
| SHRSPaS003135  | 4.039    | SHRSPaS00122   | 47.204 |
| SHRSPaS002683  | 6.661    | SHRSPaS001120  | 47.655 |
| AVT448         | 6.884    | SHRSPaS002895  | 49.289 |
| SHRSPaS003783  | 8.863    | SHRSPaS003304  | 49.913 |
| SHRSPaS002750  | 9.273    | SHRSPaS001234  | 54.212 |
| SHRSPaS001233  | 10.35    | SHRSPaS002438  | 54.212 |
| SHRSPaS004529  | 10.421   | SHRSPaS001317  | 55.615 |
| SHRSPaS002839  | 11.257   | SHRSPaS002265  | 55.849 |
| SHRSPaS004285  | 12.698   | SHRSPaS002328  | 56.457 |
| SHRSPaS004920  | 13.306   | SHRSPaS002588  | 57.347 |
| SHRSPaS005726  | 14.805   | SHRSPaS002038  | 57.903 |
|                |          |                | 57.913 |
### Supplemental Table 3. Continued.

| Gene Symbol | Value 1 | Value 2 |
|-------------|---------|---------|
| SHRSPaS002609 | 58.345 |         |
| SHRSPaS001388 | 58.538 |         |
| SHRSPaS003428 | 58.719 |         |
| AVT001 | 59.309 |         |
| SHRSPaS004508 | 59.726 |         |
| SHRSPaS004295 | 60.143 |         |
| SHRSPaS004625 | 60.53  |         |
| SHRSPaS003479 | 61.856 |         |
| SHRSPaS003327 | 62.278 |         |
| SHRSPaS003317 | 63.21  |         |
| SHRSPaS001802 | 63.21  |         |
| SHRSPaS001650 | 64.337 |         |
| SHRSPaS001745 | 64.852 |         |
| SHRSPaS001270 | 65.225 | A       |
| PSY_SNPs29or945 | 67.225 | B       |
| SHRSPaS003100 | 67.335 |         |
| SHRSPaS001623 | 69.225 |         |
| SHRSPaS002303 | 70.506 |         |
| SHRSPaS003888 | 71.047 |         |
| SHRSPaS001863 | 71.71  |         |
| SHRSPaS001328 | 72.389 |         |
| SHRSPaS003180 | 72.64  |         |
| SHRSPaS001815 | 73.147 |         |
| PSY_SNPs37or686 | 73.956 |         |
| SHRSPaS001643 | 75.612 |         |
| SHRSPaS000656 | 78.049 |         |
| SHRSPaS000327 | 80.349 |         |

**group 12**

| Gene Symbol | Value 1 | Value 2 |
|-------------|---------|---------|
| SHRSPaS003393 | 0      |         |
| SHRSPaS003248 | 2.081  |         |
| AVDO117       | 2.89   |         |
| SHRSPaS002662 | 4.505  |         |
| SHRSPaS003402 | 4.505  |         |
| SHRSPaS003265 | 6.588  |         |
| SHRSPaS001356 | 8.656 B |        |
| SHRSPaS001322 | 10.728 B|        |
| SHRSPaS000517 | 12.781 B|        |
| SHRSPaS003368 | 14.855 B|        |
| SHRSPaS002902 | 14.855 B|        |
| AVT386        | 16.374 B|        |
| SHRSPaS003965 | 17.445  |         |
| SHRSPaS003179 | 19.447  |         |
| SHRSPaS002003 | 19.447  |         |
| SHRSPaS000313 | 21.643 H|         |
| SHRSPaS002592 | 24.593  |         |
| SHRSPaS006852 | 25.916  |         |
| SHRSPaS005416 | 25.966  |         |
| SHRSPaS002243 | 26.645  |         |
| SHRSPaS003434 | 27.133  |         |
| SHRSPaS001655 | 27.961  |         |
| SHRSPaS004584 | 28.702  |         |
| SHRSPaS002624 | 29.09   |         |
| SHRSPaS005587 | 29.971  |         |
| SHRSPaS003320 | 30.136  |         |

**Continued next page**