# Development of Molecular Markers, Microsatellite

## Continuing Project; Year 2 of 2

*Vanessa E. T. M. Ashworth, Marlene de la Cruz, and Michael T. Clegg*

*Project Leader: Michael T. Clegg (909) 787-3101*
*e-mail: michael.clegg@ucr.edu*
*Department of Botany and Plant Sciences, University of California, Riverside*

*Cooperating Personnel: Dr. Mary Lu Arpaia, Dr. John Menge, Brandon McKee, Fred Guillimet, Paul Robinson, and David Stottlemyer*

### Benefit to the Industry

The development of microsatellite markers serves two goals: (1) to trace the parentage of valuable crosses and varieties, and (2) to facilitate marker-assisted selection for desirable traits at the seedling stage. Once both tasks are accomplished, breeders will be able to intercross genotypes predictably and repeatably and within a far shorter period of time than is currently possible.

### Objectives

Our primary objective is to apply our microsatellite markers to two breeding programs, (1) the Phytophthora-resistant rootstock breeding program conducted by Dr. John Menge, and (2) the breeding program headed by Mary Lu Arpaia on B-type Hass-like varieties. In both cases, microsatellite markers will be used to infer the pollen parent of significant selections and to ascertain whether selfing or outcrossing has taken place.

A second objective focuses on the elaboration of a genetic map upon which microsatellite markers and economically useful traits are positioned in relation to their inheritance: if a desirable trait, such as a fruit character, is located close to a microsatellite marker, trait and marker will be inherited together. In other words, by selecting for the marker one is simultaneously selecting for the desired fruit trait. Presence of the marker may be determined at the seedling stage, eliminating the need to wait until the tree bears mature fruit.

### Summary and Discussion

Microsatellites are stretches of DNA characterized by highly repeated di-, tri-, or tetranucleotide motifs (e.g., AGAGAG..., GATGATGAT..., TCGATCGATCGA...). Different genotypes differ in the number of repeats present and hence in the size of the repeated region. Because each individual carries two chromosomes, which may contain two identical (e.g., 12) repeats or two different repeat numbers (e.g., 10 repeats on one chromosome and 12 repeats on the other chromosome), it is possible to distinguish heterozygous individuals from homozygous individuals (codominant inheritance).

We are screening a DNA library containing a large numbers of DNA fragments (loci) for the presence of microsatellite repeats. Initially, each DNA fragment is sequenced to determine whether a microsatellite is present. A suitable microsatellite should be positioned approximately in the center of the fragment, such that the flanks either side are long enough to design primers (short segments of DNA). The primers serve to guide the synthesis of the intervening repeat and its flanking sequences by a process called PCR (polymerase chain reaction). This intervening region of DNA will be slightly larger or smaller in individuals differing in the number of repeats. The size differences between different PCR products are visualized as bands on a gel (Figures 1 and 2), whereby band position is determined by repeat size.

Over the past year, we have been screening a DNA library enriched for trinucleotide repeats. We sequenced 586 DNA fragments of which about 124 (21%) contained microsatellite repeats of adequate repeat number. However, approximately half of these loci appeared to be identical, such that primers were designed to only 65 loci (11%). Preliminary results suggest that repeat number is lower for trinucleotide motifs (the highest repeat number for any trinucleotide motif in this study was 11) than for dinucleotide repeat motifs (Kijas et al. 1995), limiting the scope of detecting differences between closely related individuals. We have now resumed the screening of the dinucleotide-enhanced library, which has a higher yield of usable loci (primers were designed to 25% of DNA fragments sequenced). The first 48 DNA fragments have been sequenced, showing the presence of many new dinucleotide repeats.

Over the past three months, we have been evaluating all microsatellite loci for which we made primers. In particular, we are working toward a thorough knowledge of the utility and repeatability of each locus. Our three-step evaluation procedure first involves generating banding profiles for selfed offspring from Hass and Zutano trees. If the parental genotype is heterozygous at a locus (e.g., ab), its selfed offspring should show three possible genotypes (aa, ab, bb), approximately at a ratio of 1 : 2 : 1 (Figure 1). If only one or two of the possible genotypes are represented, we would not use this microsatellite locus as a marker. Second, some loci show background banding in addition to the 'real' region of bands. If this noise is barely distinguishable from that of the signal bands, and adjustment of experimental conditions yields no improvements, then the locus would also be dropped from further analysis. Typically, modifications in the temperature regime, magnesium concentration, and DNA concentration influence the strength and definition of bands, but each locus has different requirements. Over the past year, we ran temperature gradients for many loci; since June, we have been fine tuning the magnesium and DNA concentrations. Third, we examine whether a locus contains a sufficient number of distinct mobility forms (bands on a gel). We have done this using a small subset of distantly related genotypes representing the three groups of avocado (West Indian, Mexican, Guatemalan) and a wild relative of *P. americana*: if only a few repeat variants are present, the locus will not be useful for comparing a range of divergent genotypes.

To date we have identified 29 microsatellite markers that show promise (Table 1), compared with less than 20 some 4 months ago. This difference reflects in part the screening of new loci, but also a better understanding of experimental conditions required by each locus. The promising new markers are being applied to the *Phytophthora*-resistant rootstock program to determine the paternity of 15 rootstock selections of predominantly known maternal origin (Figure 2).

In connection with our objective to elaborate a genetic map for avocado, we have initiated the process of creating an experimental population for QTL mapping and marker-assisted selection based on the Gwen seedling population. During the spring, the Duke 7 rootstocks destined for our experimental population were grafted onto a Lula nurse seed with the help of Dr. Menge's assistant, Brandon McKee. The rootstocks are now ready and we are about to collect budwood from our Gwen trees.

In addition, we completed two articles on the use of markers in avocado improvement for the 4th International Avocado Congress (Clegg et al. 1999), and the use of RAPD markers in assessing the frequency of outcrossing in avocado groves in California (Kobayashi et al. 2000).

**References**

CLEGG, M. T., M. KOBAYASHI, AND J.-Z. LIN. 1999. The use of molecular markers in the management and improvement of avocado. Proceedings of the International Avocado Congress. Uruapan, Mexico. October 1999.

KIJAS, J. M. H., J. C. S. FOWLER, AND M. R. THOMAS. 1995. An evaluation of sequence tagged microsatellite site markers for genetic analysis within Citrus and related species. Genome 38: 349-355.

KOBAYASHI, M., J.-Z. LIN, J. DAVIS, L. FRANCES AND M. T. CLEGG. 2000. Quantitative analysis of avocado outcrossing and yield using RAPD markers. Scientia Horticulturae (in press).

**Table 1.** Microsatellite marker loci showing promise in the determination of crossing relationships in avocado genotypes. Prefixes AVO and AUCR represent dinucleotide markers, AVT denotes trinucleotide markers.

| Locus | Repeat Motif | Predicted Insert Size | Segregation in Hass/Zutano |
|---|---|---|---|
| AVO102 | $GA_{12}$ | 153 | yes/tba |
| AVO107 | $TC_{22}$ | 184 | yes/yes |
| AVO109 | $TC_{22}$ | 153 | yes/tba |
| AVO128 | $(TC)_9CC(TC)_{10}$ | 240 | no/tba |
| AVO129 | $(TC)_{20}A(CA)_9$ | 178 | no/tba |
| AVO317 | $(CAC)_2(CAT)_2CA(CAT)_7CA(CAT)_2$ | 138 | yes/tba |
| AUCR008 | $CT_{22}$ | 117 | yes/no |
| AUCR017 | $(TC)_9(AC)_9AAA(CA)_4AA(CA)_4TT(CA)_3$ | 359 | yes/yes |
| AUCR027 | $TC_{21}$ | 267 | yes/yes |
| AUCR046 | $GA_{12}$ ... $GAA_5$ | 351 | yes/no |
| AUCR050 | $TC_{18}$ | 323 | yes/tba |
| AUCR053 | $CT_6$ ... $TC_{11}$ | 247 | yes/tba |
| AUCR101 | $TG_{16}$ ... $AG_{16}$ | 131 | no/tba |
| AUCR107 | $CT_{15}$ | 219 | yes/yes |
| AUCR114 | $AG_{21}$ | 211 | no/yes |
| AUCR160 | $GA_{20}$ | 310 | no/yes |
| AUCR162 | $TC_{15}$ | 172 | yes/tba |
| AUCR168 | $GA_{17}$ | 167 | yes/yes |
| AUCR169 | $AG_{19}$ | 360 | yes/yes |
| AUCR181 | $GA_{16}$ | 244 | yes/yes |
| AUCR183 | $(AGT)_5AGA(AG)_{17}$ | 304 | yes/no |
| AUCR230 | $(AAG)_4AAAGTG(AG)_{12}$ | 269 | yes/no |
| AUCR403 | $CT_{24}$ | 331 | yes/tba |
| AUCR405 | $CT_{18}CA_{12}$ | 161 | no/yes |
| AUCR418 | $(GT)_{12}(GA)_{13}GG(GA)_3$ | 379 | no/yes |
| AUCR466 | $TG_{25}$ | 367 | yes/tba |
| AVT095 | $TCA_9$ | 277 | tba/yes |
| AVT129 | $(CTT)_5(CTT)_8(CTT)_7$ | 227 | tba/yes |
| AVT143 | $(GAA)_8(GAT)_6$ | 211 | tba/yes |

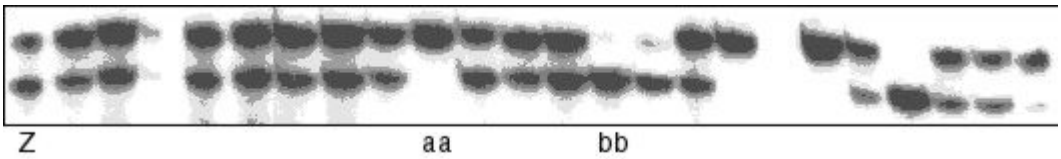


**Figure 1.** Autoradiogram of locus AUCR418 showing the selfed $F_2$ progeny of Zutano ("Z", far left). Zutano has two bands (ab), 15 $F_2$s also inherited the heterozygous pattern, 3 $F_2$s are homozygous for the upper band (aa), and 3 for the lower band (bb).

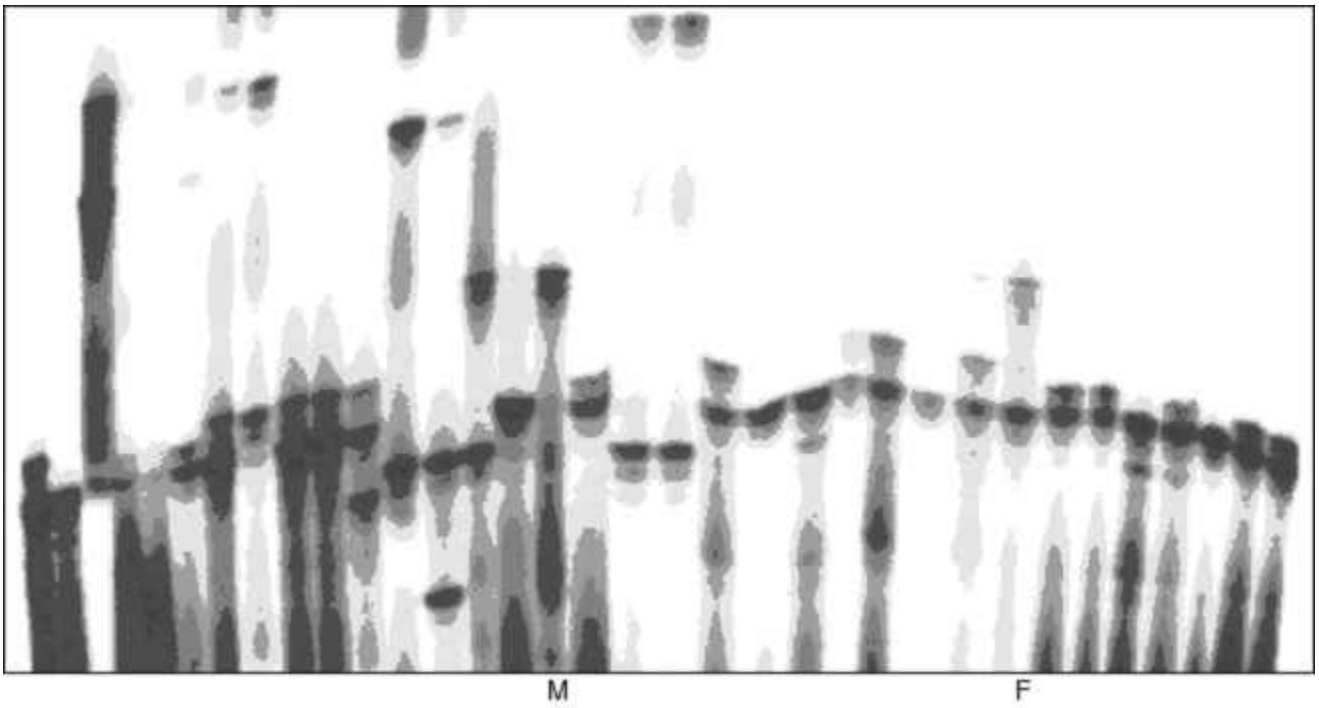


**Figure 2.** Autoradiogram of locus AVO102 depicting banding profiles of 15 *Phytophthora*-resistant (PP) selections (last 15 lanes) and their parental genotypes (maternal parent and possible pollen donors) on the left (first 19 lanes). Paternity is assessed based on band position (reflecting repeat number). For example, lane F represents a PP selection with two bands, the upper band originating from its maternal parent (M). The lower band could come from each of several pollen donors. By testing additional loci, the suite of possible pollen donors is narrowed down.