1999. Revista Chapingo Serie Horticultura 5: 227-231.

# THE USE OF MOLECULAR MARKERS IN THE MANAGEMENT AND IMPROVEMENT OF AVOCADO (*Persea americana* Mill.)

**M. T. Clegg[1]; M. Kobayashi[2] ; J.-Z. Lin[1]**
[1]Department of Botany & Plant Sciences, University of California Riverside, California 92521; Fax (909) 787-4437; michael.clegg@ucr.edu
[2]Department of Plant and Microbial Biology, University of California at Berkeley, Berkeley, California 94720

## SUMMARY

Molecular markers have become an important tool for the improvement and management of agricultural crops. Genetic markers can be developed using a number of different methods. We briefly discuss and compare several of these methods. We then consider several recent applications of molecular markers for avocado improvement. These applications range from simple problems of genotype identification to measuring the frequency of cross pollination in avocado groves in California to inferring the history of genealogical relationships among avocado cultivars.

**KEY WORDS:** Outcrossing estimation, RAPD, RFLP, microsatellites, genetic discrimination, geneological relationships.

## INTRODUCTION

Management of the reproductive system is central to the success of many crops. This is especially true for the cultivated avocado (*Persea americana* Mill.) where there is an absolute requirement for fertilization to initiate the developmental process that ultimately leads to a mature fruit. Reproductive success is a necessity for fruit yield in avocado, but does the pollen source matter? Is cross fertilization versus self-fertilization important in determining fruit yield in avocado? To address these questions it is necessary to determine the exact pollen and ovule genotypes that were joined at the moment of fertilization. In general, we can not observe the fertilization event directly so we must resort to indirect techniques to infer the history of fertilization after the fact. Genetic markers allow us to trace the pathways of genetic transmission from parent to progeny because the inheritance of the marker is known. Hence genetic markers allow us to retrospectively dissect the fertilization process.

Genetic markers have long been employed to assist in the management and improvement of agricultural crops. The first source of markers was morphological variants that had simple Mendelian patterns of inheritance. Early in the 20th century the transmission of morphological markers was used to monitor seed purity and to establish parentage in particular crosses. Unfortunately morphological markers have several significant drawbacks. First, very few morphological markers are available to assist the breeder of most crops. Second, many morphological markers confer some phenotypic disadvantage and their transmission into cultivars may be undesirable. Finally, most morphological markers are recessive which greatly limits their utility.

In the 1960s the invention of the isozyme technique provided a rich new source of genetic markers (Tanksley and Orton, 1983). Isozymes had several major advantages for the crop geneticist: First, isozymes are codominant and therefore provide much more information about patterns of genetic transmission; second, different allozyme variants (allelic forms of isozymes) do not appear to confer any disadvantageous property on the crop plant (their transmission can be regarded as functionally neutral); and third, it was possible to develop from ten to twenty isozyme markers for most crop species without great difficulty. Despite these advantages the isozyme method is limited by relatively small numbers of potential marker loci and by modest levels of polymorphism.

Beginning in the early 1980s the rapid elaboration of molecular biology began to yield new sources of genetic markers. The first new source was RFLPs (Restriction Fragment Length Polymorphisms) followed by RAPDs (Random Amplified Polymorphic DNA) followed by microsatellites and AFLPs (Amplified Fragment Length Polymorphism). Each of these new methods represented an incremental improvement over previous methods. All of these DNA based methods can produce a virtually unlimited number of marker loci and it is this unlimited number of marker loci that makes these methods so powerful. The chief distinctions between these methods are whether the method is PCR (Polymerase Chain Reaction) based (as is the case with RAPDs, microsatellites, AFLPs) or based on Southern transfers (the case with RFLPs); whether the resulting markers are codominant (RFLPs, microsatellites, AFLPs) or recessive (RAPDs) and the cost of developing the marker reagents (e.g. primers for microsatellites, cloned fragments for RFLPs). Table 1 gives a comparison of the various methods in terms of these factors.

**Table 1.** Comparison between different molecular approaches with regard to cost and genetic information content.

| Technique | **Microsatellite** PCR based assay | **RAPD** PCR based assay | **RFLP** Agarose gel/filter based |
|---|---|---|---|
| Amount of DNA used | Nanogram | Nanogram | Microgram |
| Inheritance of marker | Codominant | Dominant/Recessive | Codominant |
| Polymorphism | Highly | Moderately | Moderately |
| Development Time/Cost | High | Low/moderate | High |
| Cost per assay | Relatively low | Relatively low | Relatively high |

In this article we consider three case studies of the use of molecular markers in avocado improvement. The first case study involves the use of RFLP markers to distinguish among breeding lines. The second case study involves the use of RAPD markers to measure outcrossing in avocado orchards in southern California. The third case study involves the use of RFLP markers to unravel genealogical relationships among avocado cultivars.

## GENETIC DISCRIMINATION AMONG BREEDING LINES

*Persea schediana* Nees is a close relative of the cultivated avocado and it is relatively tolerant of avocado root rot disease (caused by the fungus *Phytophthora cinnamomi*). In 1975 several seed were collected from a Guatemalan market (Coffey, 1987). The seed were subsequently labeled as G755 A, B and C. G755A has been identified *as P. schiedeana* X *P. americana* hybrid through RFLP analysis (Furnier *et al*., 1990) and G755 B and C are also thought to be of hybrid origin. The G755 materials proved promising as root rot tolerant root stock and came into commercial use by the mid 1980s. Early workers did not distinguish between the three G755 lines and they tended to be used interchangeably. It was subsequently discovered that tolerance to root rot varied among the G755 sources and we were asked to use molecular markers to attempt to distinguish between these lines. Application of a battery of RFLP probes to DNAs prepared from each of the three sources showed that each was genetically distinct. Because each line could be associated with a unique genotype it is now possible to identify each of the three root stock sources based on a routine assay. This should assist future avocado breeders in selecting only the most root rot tolerant sources for propagation

## MEASUREMENT OF OUTCROSSING RATES IN HASS AVOCADO ORCHARDS

Avocado has a complex breeding system that is thought to be an adaptation to insure cross-pollination. This breeding system is referred to as synchronously dichogamous and it involves two flowering types (types A and B). The two floral types are distinguished by the times of pollen dehiscence and stigma receptivity. These times are complementary so that B type pollen is available to fertilize A type flowers and conversely A type pollen is available to fertilize B type flowers ( Stout, 1923; Davenport, 1986).

The Hass cultivar (which is type A) is typically grown in large monocultures in California. Because the ‘Hass’ genotype is clonally propagated through bud grafting these monocultures are composed of a single scion genotype and the complementary B mating type is unavailable as a source for pollination. In the past 15 years a slow decline in ‘Hass’ productivity has been observed in California. An important question is whether the reduced opportunity for pollination has caused reduced fruit yields in California? To address this question it is necessary to measure rates of cross-pollination (defined as pollination by a genotype different from the maternal parent) using marker loci and to correlate these rates with fruit yield. It is also desirable to ask whether different B type pollen sources vary in their efficacy as pollinators of ‘Hass’ maternal plants and whether cross-pollination frequencies depend on climatic factors. Accordingly, Kobayashi *et al*. (2000) initiated a series of experiments to measure these factors.

## MATERIAL AND METHODS

Fruit were collected from ‘Hass’ orchards over four years in the coastal and inland climatic regions of California. When comparing the average minimum temperature, maximum and mean temperatures during avocado flower bloom, March to mid-May, (Arpaia, 1997) of the two climatic regions, the coastal area is slightly cooler than the inland region. Precipitation in the coastal regions is approximately 2 to 5 times the precipitation in the inland area. In total, three inland groves in Riverside County were

chosen for each of the pollen sources (‘Bacon’, ‘Fuerte’, and ‘Zutano’). The coastal region was represented by three groves in Ventura county each of which contained the ‘Bacon’, ‘Fuerte’ and ‘Zutano’ pollen sources and one grove in Santa Barbara county with two sites, one for ‘Bacon’ and one for ‘Fuerte’.

The sites selected for study all had a tree (or trees) of ‘Bacon’, ‘Fuerte’ or ‘Zutano’ (type B pollen source) along the edge row of the orchard that could to serve as a pollen donor. At each site, 20 fruit were collected from two trees at a distance of one, five and fifteen rows from the potential pollen source. The number of total fruit on the six sample trees was also counted.

**RAPD analysis**

DNA was isolated from each embryo of the collected fruit using a modified DNA extraction procedure of Rawson *et al.* (1982) as described in Kobayashi *et al.* (2000). The typical preparation yielded between 2000 and 8000 nanograms of DNA suitable for PCR amplification. Approximately 15 nanograms of DNA was amplified as described in Kobayashi *et al.* (2000). More than 300 decamer primers (Operon Technologies**,** Alameda**,** CA) were screened against avocado DNAs to select primer pairs that yielded reliable fragment patterns and that discriminated among Hass and the three pollen sources. Five primers (OPC-2, OPC-7, OPE-12, OPE-13, OPE-18) produced a band in Bacon that was absent in Fuerte, Zutano and Hass. One primer, OPD-11, produced a band in Zutano that was absent in Bacon, Fuerte and Hass. Five primers (OPC-18, OPE-13, OPE-14, OPF-11, OPG-7) produced a band in Fuerte that was absent in Bacon, Zutano and Hass. The PCR products were electrophoresed in 1.7% agarose gel in TBE at 50 volts for 4 hours. The gel was stained in ethidium bromide and photographed using UV fluorescence to reveal RAPD bands of amplified DNA.

The Mendelian inheritance of each RAPD marker was verified by progeny testing seed progeny of the cultivars of interest (Kobayashi *et al*., 2000). This also established whether the cultivar was homozygous or heterozygous for each RAPD marker. Based on this criterion RAPD marker OPD-16 was present and homozygous for ‘Bacon’ and ‘Zutano’ while no band was present in ‘Hass’ or ‘Fuerte’, indicating a recessive homozygous genotype. The ‘Fuerte’ RAPD marker, OPE-14, was heterozygous, indicating that the observed outcrossing rate for ‘Fuerte’ must be doubled to account for all the outcrossing events. This RAPD band was absent in ‘Hass’, ‘Bacon’ and ‘Zutano’, indicating a recessive homozygous genotype.

**Statistical analysis**

The data set analyzed in this experiment derived from the RAPD assay of 2,393 individual fertilization events. Each progeny was scored as either an outcross or as the result of self-fertilization. Initial analysis indicated that outcrossing rate does did not vary by years. The outcrossing rate averaged over four years, all pollen sources, and all locations is was 0.371. A log-linear model for analysis of variance for categorical variables (Bishop *et al*., 1974) was fitted to the data combined over years, first with main effects and two-way interactions, and then by adding more terms to the model. The fit of the data to the model was tested by comparing the reduced model to the full model by a likelihood ratio test using PROC CATMOD in SAS (SAS Institute Inc., 1989).

The relationship between outcrossing rate and yield (number of fruit per tree) was examined for the three locations (Riverside, Ventura and Santa Barbara counties)

separately, where data for each year were treated as independent data points to retain sufficient sample size for each population. A Pearson correlation coefficient was calculated using PROC CORR in SAS (SAS Institute Inc., 1989).

Maximum likelihood analysis of variance showed that there were marginally significant effects of location and pollen sources on outcrossing rate, while there was a highly significant effect of row number on outcrossing rate. On average, outcrossing rate in the coastal regions (Ventura and Santa Barbara) is was higher than in the inland region (Riverside). Populations growing with the 'Fuerte' had the highest outcrossing rate while those growing with the 'Bacon' had the lowest outcrossing rate. A highly significant location x pollen source interactive effect on outcrossing rate was also observed. The interaction is explained by a change of ranks of outcrossing rate with pollen sources 'Bacon' and 'Zutano' between locations. The inland populations growing with 'Bacon' had a higher outcrossing rate than those growing with 'Zutano', whereas the coastal populations growing with 'Zutano' had a higher outcrossing rate than those growing with 'Bacon'. The average outcrossing rate decreased as the distance from pollen source increases increased from row 1 to row 15.

There was a marginally significant positive correlation between outcrossing rate and yield at Ventura, while there is was no significant correlation at Temecula and Santa Barbara. The proportion of variation in yield attributable to outcrossing rate was rather small ($R^2$=0.01 to 0.25, or 1 to 25%). We conclude that the changes in outcrossing rates have a small effect on yield.

An important feature of the experimental design was the ability to classify outcrosses by pollen source. This allowed an assessment of the efficacy of three different cultivars as pollen sources. The results clearly establish that 'Fuerte' is the most effective pollen source for 'Hass' maternal trees independent of climatic region. Accordingly, we recommend 'Fuerte' as a pollen source when it is deemed advantageous to have a mixed planting.

## GENEALOGICAL RELATIONSHIPS AMONG AVOCADO CULTIVARS

Markers can also be employed to trace more complicated patterns of genetic transmission that span many generations. It is of obvious interest to determine the history of genetic relationships that connect different avocado cultivars, because knowledge of these patterns is important in guiding breeding programs. To explore this possibility, Davis *et al*. (1998) cloned a number of anonymous DNA fragments from the avocado genome into a plasmid vector. These RFLP clones were then applied to Southern transfers of DNA from a panel of avocado cultivars to determine whether useful polymorphisms were associated with each RFLP pattern. The results indicated that virtually every anonymous clone was associated with polymorphism when applied to materials that represented the three races of avocado. These results suggest that cultivated avocado possess a very diverse gene pool.

To study the genealogical history of avocado cultivation RFLP patterns associated with fifteen anonymous clones were analyzed for 38 avocado cultivars. Every cultivar was found to be genetically distinct from every other cultivar based on one or more RFLP pattern. A measure of genetic similarity was then calculated based on the fraction of genes in common between pairs of cultivars. The resulting data were subjected to a cluster analysis to determine the patterns of genetic similarity across cultivars. These

analyses (Figure 1) clearly reveal three clusters that correspond to the three botanical races of avocado. Two additional clusters are also revealed in the analysis and these appear to represent cultivars that have an intervarietal hybrid origin. For instance, the cultivar Hass is believed to be a Guatemalan x Mexican x Guatemalan backcross. The genetic data appear to confirm this hypothesis because Hass falls into a cluster that is between the Guatemalan and Mexican cultivars but closer to the Guatemalan cluster (Figure 1).

The ability to average across a number of different nuclear loci provides a powerful means to assess the average genetic contribution to each avocado cultivar. It is this averaging across loci and across lineages that allows an accurate reconstruction of the genealogical history of avocado breeding relationships. These data are of particular utility because they can serve as a guide to future breeding program.

## CONCLUSIONS

The examples presented in this article illustrate the utility of genetic markers in avocado management and improvement. It is obviously important to have a means to unambiguously distinguish between genotypes. Similarly a detailed knowledge of the genealogical relationships among cultivars and other breeding materials is of substantial utility in guiding future breeding strategies. Genetic markers provide the tools to determine these relationships.

The successful management of many crops requires an ability to manage the system of pollination. In the case of avocado fruit development is absolutely dependent on fertilization. The work discussed in this article represents the most comprehensive effort to date to relate success in cross-pollination to pollen sources, climate and fruit yield. The data reveal that ‘Fuerte’ is the most effective pollen source for ‘Hass’ maternal plants in California. The correlation between yield and cross-pollination is weak and only a small proportion of the variance in yield can be accounted for by changes in rates of cross-pollination. This casts doubt on the value of interplanting type B pollen sources into California orchards.

Crop improvement has three fundamental objectives. The first objective is to manipulate sexual crosses so as to combine desirable genes from different lineages into a single lineage. The second objective is to select desirable genotypes from the enormous combinations of genotypes produced through sexual reproduction. And, the third objective is to maintain and propagate those genotypic combinations deemed most useful by the breeder (Clegg, 1985). Genetic markers provide valuable tools to assist in creating desired genetic combinations, to assist in the identification and selection desirable genotypes. Because of their great utility, genetic markers will become standard tools for avocado improvement in the future.

## ACKNOWLEDGEMENTS

This research was supported in part by the California Avocado Commission.

## LITERATURE CITED

ARPAIA, M.L. 1997. Development of a phenology model for California avocados. *In*: J. Shepherd (ed.), California Avocado Research Symposium. Saticoy, California.

BISHOP, U.M.M.; FIENBERG, S.E.; HOLLAND, P.W.. 1974. Discrete Multivariate Analysis: Theory and Practice. The MIT Press, Cambridge, MA.

CLEGG, M. T. 1986. Genetics of crop improvement. Amer. Zool. 26: 821-833.
COFFEY, M.D. 1987. *Phytophthora* root rot of avocado: An integrated approach to control in California. Plant Disease 71: 1046-1052.
DAVENPORT, T. L. 1986. Avocado flowering. Horticultural Reviews 8: 257-289.
DAVIS, J.; HENDERSON, D.; KOBAYASHI, M.; CLEGG. M.T. 1998. Genealogical relationships among cultivated avocado as revealed through RFLP analyses. J. Heredity 89: 319-323.
FURNIER, G. R.; CUMMINGS, M.P.; CLEGG, M.T. 1990 Evolution of the avocado as revealed by DNA restriction fragment variation. J. Heredity 81: 183-188.
KOBAYASHI, M.; LIN, J.-Z.; DAVIS, J.; FRANCIS, L.; CLEGG, M.T. 2000. Quantitative analysis of avocado outcrossing and yield in California using RAPD markers. submitted 6/99.
RAWSON, J. Y. F.; THOMAS, K.; CLEGG, M.T. 1982. Purification of total cellular DNA from a single plant. Biochemical Genetics 20: 209-219.
SAS INSTITUTE INC. 1989. SAS/STAT User's Guide, Version 6, 4th ed. Vol. 1. Cary, NC.
STOUT, A.B., 1923. A study in cross pollination of avocados in Southern California. California Avocado Society Yearbook 7: 29-45.
TANKSLEY, S.D.; ORTON, T.J. 1983. Isozymes in plant genetics and breeding. Volume 1. Amsterdam; New York: Elsevier.

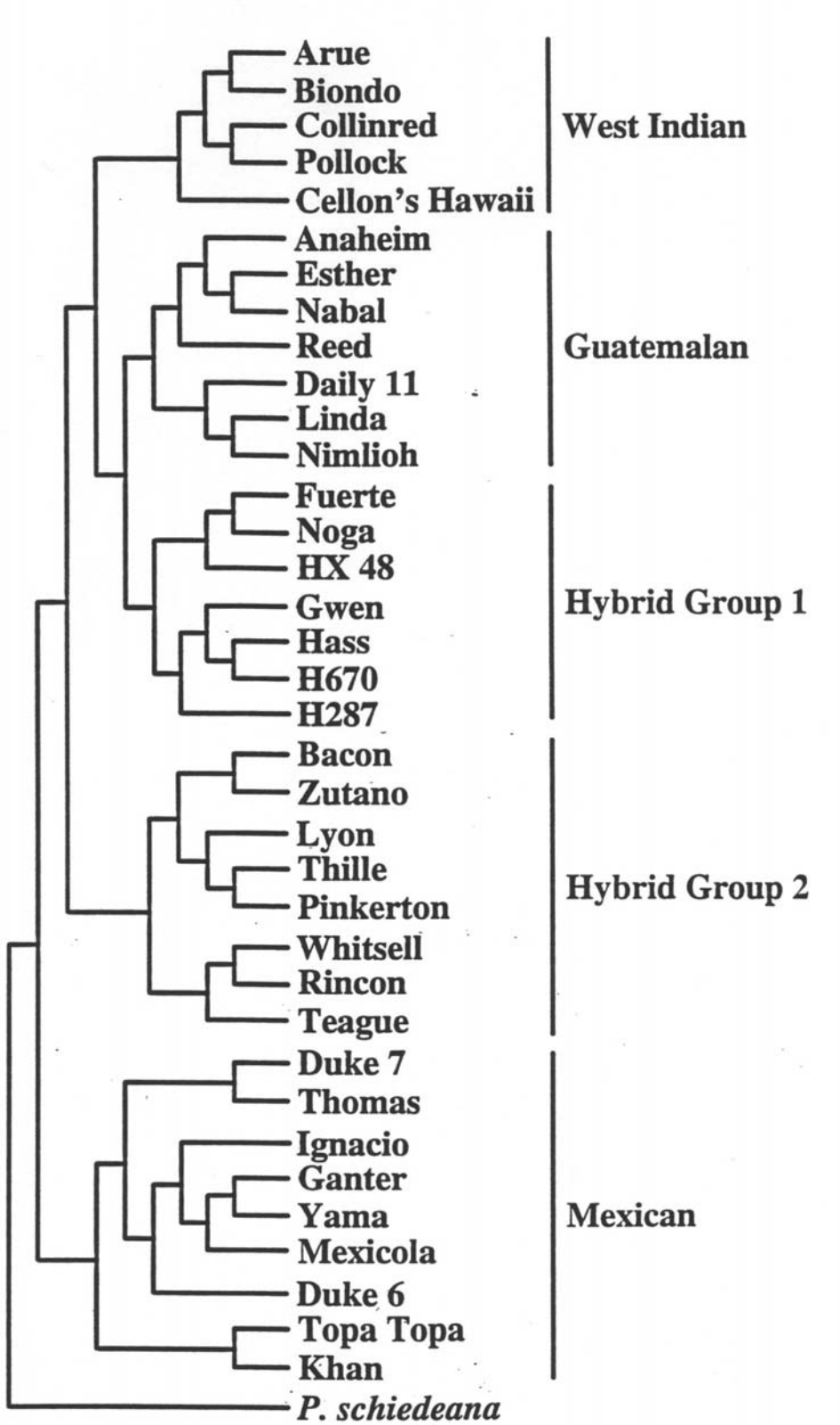


Figure 1. Cluster analysis of average number of genes shared among pairs of cultivars.