Development of Genomic Tools for Almond Rootstock Improvement

Dianne Velasco[1,3], Craig Ledbetter[2] & Malli Aradhya[3]

[1]UC Davis, [2]USDA-ARS, Parlier, CA and [3]USDA-ARS, Davis

In collaboration with
John Preece[3], Daniel Kluepfel[3], Gregory Browne[3], Carolyn DeBuse[3], Michael McKenry[4]
[4]UC Riverside (Kearney Ag. Center)

INTRODUCTION

PROBLEMS AND SIGNIFICANCE

The key soil-borne pathogens of almond, *Phytophthora* (PHY) rots/Agrobacterium induced crown gall (CG)/lesion and root knot nematodes (NEM) are the major limiting factors in orchard/nursery productivity and orchard longevity. The widely used fumigant in almond nurseries and orchards, methyl bromide is rapidly being phased out. Finding a sustainable and environmentally sound alternative is a top priority for the Almond Board of California. Rootstocks with resistance to soil-borne pathogens will significantly reduce our dependence on soil fumigation for optimal productivity. Development and deployment of genetic markers will facilitate rapid selection of resistant rootstock genotypes at juvenile stages without pathogen challenge and grow out. These advantages cannot be overstated when considering selection of a commercially viable woody perennial tree rootstock.

Availability of diverse germplasm, high throughput marker systems to produce high density genotypes, efficient foolproof disease testing schemes are the key to the success of this project. Single Nucleotide Polymorphisms (SNPs) are the marker of choice for high density genotyping. This research during the past three years has focused on: (1) screening and identifying reliable sources of durable combined resistance to these soil borne diseases; (2) hybridize potential donor species with peach and almond genotypes to produce novel rootstock genotypes; (3) SNP genotyping and disease testing; and (4) identify markers associated with soil borne diseases to develop and validate effective marker assisted selection strategies.

OBJECTIVES

- Develop molecular markers linked to disease-pest resistance for use in almond rootstock breeding programs.
- Discover single nucleotide polymorphisms (SNPs) using genotyping-by-sequencing strategy.
- Genotype commercial and experimental rootstocks and newly produced genetically diverse interspecific hybrids.

MATERIALS AND METHODS

Disease testing of rootstocks

During 2014, a new set of 20 hybrids out of a total of 34 197-series hybrids were subjected to CG evaluation (Figure 1). All 34 hybrids will be subjected to both CG and PHY during the Spring, 2015.

Genomics (SNP discovery and Genotyping)

Genotyping-by-sequencing (Elshire et al., 2011) of 190 diverse, currently used rootstocks that are under field testing and experimental hybrids developed in this project has yielded ~221 million reads of which 18 million unique reads were aligned to the published peach genome sequence and, after filtering, 164,742 SNPs were assembled. Further SNPs with low LD and low representation across hybrids were filtered out to arrive at 7444 (7k) SNPs for the association analysis.

Association Analyses

Disease screening data: Unpublished categorical screening data for CG was provided by Kluepfel lab at USDA-ARS. Unpublished quantitative screening data for PHY was provided by Browne lab at USDA-ARS. Quantitative root knot and root lesion nematode data was gleaned from online sources originating from the McKenry lab at UC Riverside. A final set of 7444 filtered SNPs was combined with disease data for CG/PHY/RLN for association analysis

TASSEL v5.0 (Bradbury et al., 2007): Mixed Linear Model (MLM) which uses population structure and coancestry among genotypes as covariates in computing marker-disease phenotypes associations was employed. Population structure was computed by generating eigen vectors using the principal components analysis (PCA) with PLINK software and coefficient of coancestry (kinship matrix) was generated using a module available in TASSEL software. These two parameters along with SNP genotype data and numerical (quantitative or categorical) phenotype data was used in association analysis.

PLINK v1.90 (Purcell et al., 2007; Chang et al., 2014): The association analysis using a general linear model (GLM), which accounts for only population structure, was used to compute associations. SNP data filtered to keep only biallelic loci. Loci with the minor allele at less than 5% frequency or missing data at greater than 10% were discarded. The SNPs were then filtered based on pair-wise linkage disequilibrium and subjected to a multidimension scaling (MDS) analysis (Figure 2). The coordinates extracted from the MDS analysis accounting for population structure was used in the association analysis. The quantile-quantile (QQ) plot of probability distribution of observed associations was generated to assess the goodness of fit with expected probability distributions. The chromosome-wise probability of associations between marker and phenotypes was displayed in a Manhattan plot (Figure 3).

REFERENCES

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633-2635.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2014. Second-generation PLINK: rising to the challenge of larger and richer datasets. arXiv:1410.4803

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, *6*(5), e19379.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, 81.

ACKNOWLEDGEMENTS

California Almonds — Almond Board of California
CDFA — California Department of Food and Agriculture
UC Davis Department of Plant Sciences

Table 1. The MLM association analysis of crown gall screening data using TASSEL.

Marker	Chr	Site	df	F	p	Error df	Marker R^2
S1_212063151	8	15627172	2	5.94159	0.00831	36	0.11168
S1_72042880	2	25165154	2	6.29985	0.00843	31	0.10136
S1_54210819	2	7333093	1	10.30516	0.00513	29	0.09392
S1_136457227	*5*	*10217200*	*2*	*6.0658*	*0.00969*	*31*	*0.09334*
S1_136457228	*5*	*10217201*	*2*	*6.0658*	*0.00969*	*31*	*0.09334*
S1_136457231	*5*	*10217204*	*2*	*6.0658*	*0.00969*	*31*	*0.09334*
S1_7294016	1	7294016	2	5.72594	0.00959	36	0.0907

Table 2. Adjusted GLM association analysis of crown gall data using PLINK.

CHR	SNP	Bonferroni p-vlaue	Holm_Bonferroni p-value	QQ p-value
1	S1_38007759	0.0447	0.0447	1.43E-05

RESEARCH HIGHLIGHTS

- **Developed SNP markers using GBS and genome resequencing approaches.**
- **Identified marker(s) associated with crown gall (CG), an important soil-borne disease.**
- **Generated a wide range of novel interspecific *Prunus* hybrids utilizing species that are potential donors of disease resistance.**
- **Initial disease testing showed a range of variation in response to soil-borne diseases indicating the potential for development of resistant rootstocks.**
- **Identified interspecific hybrids with resistance to CG and *Phytophthora*.**

Figure 1. New embryo rescued interspecific hybrids clonally propagated for disease testing

RESULTS & DISCUSSION

Identification of CG and PHY resistant interspecific hybrid rootstocks

Two selected rootstock genotypes, L-1-2 (open-pollinated *P. cerasifera*) and P-4-25, a hybrid 'Nemared' x *P. fenzliana*, (Figure 4) that showed resistance in the first round of disease screening were budded with 'Nonpareil' scion and planted at Wolfskill Experimental Orchards (WEO) in Winters, Calif. for further observations on graft compatibility. Additionally, an area within the almond germplasm collection at WEO known to have *Armillaria* infestation has been planted with experimental hybrids for observation. The first round of disease of testing of novel rootstocks generated in the project has yielded encouraging results with wide variation in response to CG and PHY screening (Figures 4, 5 & 6). A subset of 20 diverse interspecific hybrids (197 series) were tested for CG resistance during 2014 from a total of 34 interspecific hybrids. The hybrid 197-113 (*P. persica* x *P. tangutica*; Figure 6) was virtually immune to CG while the others, 197-11 (*P. persica* x *P. argentea*), 197-133 (*P. persica* x *P. tangutica*), 197-190 (*P. persica* x *P. dulcis*), 197-209 (*P. persica* x *P. kuramica*), and 197-214 (*P. persica* x *P. bucharica*) showed high levels of tolerance to CG. A second round of disease testing is planned for Spring 2015 for both CG and PHY. The hybrids to be tested include all ~ 190 clonal plants from a set of 34 diverse interspecific hybrids involving *P. persica*, *P. argentea*, *P. tangutica*, *P. dulcis*, *P. bucharica*, *P. kuramica*, *P. davidiana*, and *P. kansuensis*.

Association Analyses

The association analyses performed following the mixed linear model (MLM) implemented in TASSEL and general linear model (GLM) in PLINK software packages indicated association of several SNPs with crown gall (CG) with R^2 values ranging from 0.09 to 0.11 (Tables 1 & 2), which is considered significant for complex traits such as disease resistance with low heritability. While these analyses are still preliminary, at this time, we have identified significant association of markers with CG, but the analyses failed to come up with any markers for *Phytophthora*, root knot or lesion nematode infestations. We suspect that lack of consistent and adequate disease testing data has resulted in failure to detect markers associated with disease resistance. We are further filtering the data set using a number of filtering criteria to eliminate leaky data among the SNP loci identified in the GBS analysis for further analyses.

Genome resequencing and SNP calling

In the genome sequence data a minimally filtered SNP calling pipeline discovered from 2,026,306 to 24,206,662 SNPs, depending on species combination used for calling. The lowest number was from the group with only *P. persica* samples whereas the highest number was from the group of all samples across species. With over 10 million SNPs *P. dulcis* had approximately five times the number of SNPs as *P. persica*. This difference is expected when mating systems of the two species, outcrossing versus selfing, are considered. When all species are considered together for SNP calling and more stringent filtering results in 28,908 high quality SNPs.

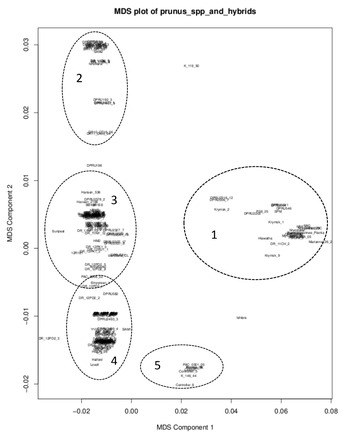

Figure 2. A 2-dimensional plotting of MDS coordinates using SNP data. Gr. 1 - composed of samples with plum background, Gr. 2 - almond and wild almond species, Gr. 3 - peach-almond hybrids, Gr. 4 - peach and wild peach species, and Gr. 5 - hybrids with peach x plum parentage.

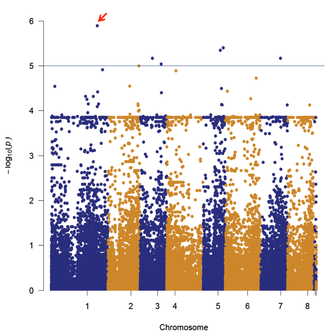

Figure 3. Manhattan plot - distribution of $\log_{10}$(p-values) showing the probability of association between SNP markers and crown gall. Red arrow indicates a SNP with significant association.

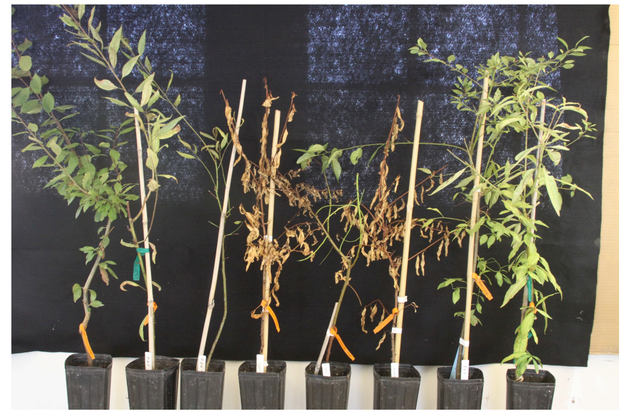

Figure 4. Disease testing of interspecific rootstock hybrids for *Phytophthora* crown and root rot. (L-R) Marianna 26-24, P-4-1, P-2-4, three dead project interspecific hybrids, L-1-2, and 'Nemaguard'.

Figure 5. Disease testing of interspecific rootstock hybrids for crown gall. (L-R) P-2-1, P-2-2, P-2-9, P-4-10, 198-3, and a water control.

Figure 6. Crown gall evaluation of 197-series hybrids with 197-113 (center) exhibiting resistance while 197-133 (right) is highly susceptible. A water inoculated control is shown at left.