Identification of Almond Self-Incompatibility and Self-Compatibility Sources Using Molecular Markers

Project No.: **10-HORT14-Dangl/Gradziel**

Project Leader: Jerry Dangl
Foundation Plant Services
UC Davis
One Shields Ave.
Davis, CA 95616
(530) 752-7540
gsdangl@ucdavis.edu

Project Cooperators and Personnel:
Thomas Gradziel, Plant Science, UC Davis
Judy Yang, Foundation Plant Services, UC Davis

Objectives:

1. Develop PCR (polymerase chain reaction)-based molecular markers effective in rapidly and unequivocally distinguishing among the different self-incompatibility and self-compatibility genotypes currently utilized in the UCD almond breeding program.
2. Begin to test the utility of these markers to identify/characterize the self-compatibility type in advanced UCD almond breeding lines to allow evaluation of associated differences in the final expression of self-compatibility and yield potential for different self-compatibility genes including different combinations of self-compatibility genes.
3. Make S gene-allele genotyping technology available to the California almond Industry.

Interpretive Summary:

Almond cultivars are self-incompatible. A given tree cannot pollinate itself, nor can it pollinate other trees of the same variety, since almond varieties are propagated through genetically identical cuttings. Thus, commercial production requires inter-planting of orchards with pollenizer cultivars that have similar bloom dates to and are cross-compatible with the primary commercial cultivar. Since the almond industry in California is centered on the high quality nut of ‘Nonpareil’, it is no surprise that nearly all production other than ‘Nonpareil’ comes from its primary pollenizers like ‘Carmel’, ‘Fritz’, and ‘Monterey’,(ABC, 2011). Accurate knowledge of self-incompatible type is necessary to ensure orchard plantings are cross-compatible and in the breeding and selection of future almond varieties.

Self-incompatibility in almond is controlled by many variants (alleles) of a single gene, the S-gene. A variety inherits one version of the S-gene (one S-allele) from each parent. Both of these S-alleles are active in the flower. A flower cannot be fertilized by pollen with the same S-allele as either one of the two S-alleles in the flower. Nonpareil flowers have the S-alleles S7 and S8, so any pollen with either S7 or S8 is cross-incompatible with Nonpareil and seed won’t be set. Since pollen is haploid and thus has only one S-allele, Nonpareil pollen is either S7 or S8. The S7 Nonpareil pollen will fail to fertilize any variety with S7 as one of its S-alleles and the S8 Nonpareil pollen fail to fertilize any variety with S8 as one of its S-alleles; this means no

pollen from Nonpareil can pollinate a Nonpareil flower. In the same way, different almond varieties with identical S-allele genotypes cannot pollinate one another and are said to belong to the same cross-incompatibility group (Kester et al. 1994). 'Butte' and 'Monterey', common pollenizers of 'Nonpareil', share the identical S-allele genotype, S1 and S8, and so do not pollinate each other. Both, however, produce an abundance of pollen with the S1 allele, which is compatible on 'Nonpareil'. Cross-incompatibility groups for California almond varieties have been established through controlled experimental crosses made in orchards and confirmed by years of grower observations and successful almond production throughout the state. Accurate assignment of cross-incompatibility groups by this method requires hundreds of replicated controlled crosses made over the course of several years, and is a substantial barrier to the breeding and introduction of new almond cultivars.

For more than a decade, researchers have been exploring the use of DNA marker analysis, or DNA fingerprinting, to study and assign S-allele genotypes in almond (Ushijima et al. 1998, Tamura et al. 2000). These methods can identify individual S-alleles directly using the polymerase chain reaction (PCR). More recently, this approach has been refined and the number of S-alleles that can be characterized has been increased ((Sánchez-Pérez et al. 2004, Barckley et al. 2006). However, these initial PCR-based procedures used for almond are both tedious and prone to error. A particular problem has been the failure of previous protocols to consistently and accurately discriminate between the Nonpareil S-alleles, S7 and S8. Because of the importance of 'Nonpareil' both as the primary commercial cultivar and as a major recurrent parent in breeding programs, distinguishing these S-alleles is very important to the California almond industry.

In this research, we have improved the consistency and accuracy of detecting and discriminating the S-allele, with specific attention given to varieties of importance to the California almond industry, particularly the S7 and S8 alleles of 'Nonpareil'. We have adapted previous protocols for use in highly automated genetic analyzers. We use the same genetic analyzer and similar methods for routine DNA fingerprinting of almond for varietal identification and pedigree analysis, this technology is available to the California almond industry through the Plant Identification Lab at Foundation Plant Services (http://fpms.ucdavis.edu/IDTesting.html). Integrating DNA fingerprinting of S-allele genotype with varietal identification and pedigree analysis provides a clear snapshot of new selections, particularly with respect to their potential compatibility with current elite varieties like 'Nonpareil'. To assist breeders and growers, DNA-based identification of S-allele genotype is now offered by the Plant Identification Lab on a case by case basis. Though DNA Fingerprinting of S-allele genotype can be a very useful tool, the correlation between the results of DNA analysis and actual fertilization and seed set in many cases will need to be confirmed by field testing.

Materials and Methods:

A set of 22 almond varieties was selected for study. As a group, these varieties hold the majority of known almond self-incompatible alleles (S-alleles) and all combinations of S-alleles seen in California almond varieties (**Table 1**). The set contains current varieties, including 'Nonpareil' and the major 'Nonpareil' pollenizers, several heritage varieties, the self-compatible variety 'Tuono' and two new selections submitted by California almond nurseries. Total genomic DNA was extracted from dried leaves using a commercial extraction kit (DNeasy® 96

Plant Kit, Qiagen, Valencia, CA). All of the samples were genotyped with a set of Simple Sequence Repeats (SSR) markers (Dangl et al. 2009) to confirm correct varietal identification.

Portions of the self-incompatibility gene were amplified by PCR using combinations of previously published DNA marker sequences designed to detect various groups of S-alleles and amplify fragments that are, based on the length of the fragment, diagnostic of a specific S-allele (Sánchez-Pérez et al. 2004, Tamura et al. 2000). The length, size in base-pairs (bp) of these amplified fragments was determined by capillary electrophoresis on a Genetic Analyzer (ABI Prism 3130xl, Applied Biosystems) using several experimental protocols, including a new size standard from Applied BioSystems, 1200LIZ. The fragments were also run on a 2% agarose gel to confirm amplification of S-allele fragments and to determine a rough estimate of the fragment size. Published full-length sequences of S-allele were obtained online from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). These sequences were used to localize published primer sequences to compare calculated size of a given S-allele to the size published in the original paper.

Results and Discussion:

The first issue to resolve was that, despite repeated attempts with an array of PCR protocols, we had previously failed to detect the S-alleles of 'Nonpareil' (S7 and S8). Others have remarked on the unreliability of the relevant primers. The second problem was that many of the S-allele-specific fragments amplified using published primers are longer than the ~ 500 base-pair (bp) maximum length that can be routinely detected by automated capillary electrophoresis. We planned to address both issues by analyzing the published full-length sequences of various S-alleles and redesigning primers to consistently detect S7 and S8 and to reduce the length of those S-alleles that were > 500 bp.

Our sequence analysis reveled inconsistencies in the published primer sequence for S8 and the predicted size of several S-alleles (Tamura et al. 2000). With this information, we obtained the correct primer sequence. After testing this sequence with different protocols, we could consistently detect both S7 and S8. Using these new protocols and the genetic analyzer, the lengths we obtain for S7 and S8 are very precise from one test to the next and much more accurate than published full-length sequences. With this new protocol, we have reliably identified S7 and S8 from 'Nonpareil' and several other varieties (see **Tables 2 & 3**).

We used two approaches to solve the problem of S-allele specific fragments generated by published protocols that are too long for consistent routine automated genotyping: reducing the lengths of the S-allele fragments and increasing the length of fragments detectable by the genetic analyzer. We tested a new pair of primers, "set B" (**Table 2**). Analysis of full-length S-allele sequences and published literature indicated this primer pair would amplify shorter fragments for many S-alleles, compared to the "set A" primers (**Tables 1 & 2**). The analyzer requires a "size standard" that is used to calculate the size of the test fragment. To increase the detectable fragment length, a longer size standard is necessary. The genetic analyzer also requires software modifications that instruct the genetic analyzer to "look for" both longer test fragments and longer size standard fragments. We continued our previous efforts with the 1200LIZ size standard from Applied BioSystems (see last year's report). As the name implies, 1200LIZ contains reference fragments as long as 1200 bp and there is an associated software package available online.

The EM-PC2consFD and EM-PC3consRD primer pair, “Set B” (Sutherland et al. 2004) worked very well after many trial-and-error adjustments to our protocols. Fragments of the correct diagnostic length were consistently generated for eight common almond S-alleles. We analyzed these fragments side-by-side with fragments of the same S-alleles generated using the ASIII, AmyC5R primer pair, “Set A”. As anticipated based on our sequence analysis, the fragments from primer “Set B” were shorter then those generated using ‘Set A”. Though the fragments lengths are different depending on the primer pair used, our results were consistent with expected results for a given S-allele, primer pair combination. Obtaining consistent results from both sets of primers provides a high degree of confidence that the system is working.

Our experiments with the LIZ1200 size standard were successful in increasing the length of fragments detectable by the genetic analyzer. We can detect and accurately analyze fragments as large as 1300 bp. However, the precision of fragment sizing decreases as the length of the fragment being measured increases beyond 1100 bp (**Table 3**). This variation does not pose a problem here, as the error margins are well within what is necessary for routine identification of almond S-alleles. We also intend to use both primer Sets “A” and “B” for future routine S-allele identification; the smaller fragment sizes obtained with “Set B” are more precise from one test to the next and, as previously noted, obtaining accurate results from both sets of primers provides a high degree of confidence in the S-allele determination.

Almond S-allele identification based on the protocols developed by this work has already been used to identify the S-allele genotypes of two almond varieties submitted by two California almond nurseries. This service is now available to the California almond industry though the Plant Identification Lab at Foundation Plant Services, UC Davis. (http://fpms.ucdavis.edu/IDTesting.html).

Table 1. Almond varieties and corresponding self-incompatiblity alleles with published and determined lengths in base-pairs.

Variety	S-allele	Expected sizes for each Primer Sets[a]	
		A or C	B
Aldrich	1, 7	1083, 220	750, 1719
Arbuckle	5, i[b]	600, 1230	270, 900
Carmel	5, 8	600, 207	270, 2418
Ferragnes	1, 3	1083, 1210	750, 900
Harriot	6, 14	860, 1300	550, 980
Jordanolo	7, 14	220, 1300	1719, 980
Kapareil	8, 13	207, 1400	2418, 1090
Kochi	8, 6	207, 860	2418, 550
Kutsch	8, 1	207, 1083	2418, 750
Languedoc	1, 5	1083, 600	750, 270
Marcona	11, 12	708, 1600	na[c], 1270
Milow	7, 13	220, 1400	1719, 1090
Mission	5, 1	600, 1083	270, 750
Monterey	1, 8	1083, 207	750, 2418
Nonpareil	7, 8	220, 207	1719, 2418
Padre	1, 18	1083, 647	750, 318
Ruby	1, 6	1083, 860	750, 550
Sonora	8, 13	207, 1400	2418, 1090
Sweetheart	3, 4	1210, 930	900, 600
Tuono	1, f	1083, 450	750, na[c]
California Nursery 1	1, 8	1083, 207	750, 2418
California Nursery 2	1, 8	1083, 207	750, 2418

[a] Primer Set as described in Table 2. Set C only amplifies S7(220bp) and S8(207bp)

[b] Si does not have a assigned European number

[c] Not enough information is published to determine expected sizes these S-alleles with this primer pair

Table 2. Primer combinations used to amplify Self-incompatibility gene fragments diagnostic for specific S-alleles.

Sets	Forward Primer	Reverse Primer	Target S-alleles
A	ASIII[a] (universal forward) & CEBASf[b] (Sf specific forward)	AmyC5R[a] (universal reverse)	All except S7, S8
B	EM-PC2consFD[c]	EM-PC3consRD[c]	All except Sf, S7, S8
C	Alsc1[a] (S7 specific forward) & Alsd2[a] (S8 specific forward)	AmyC5R	S7 and S8

[a]Tamura et al. 2000

[b]Sánchez-Pérez et al. 2004

[c]Sutherland et al. 2004

Table 3. Almond S-alleles that can now be consistently and reliably genotyped with capillary electrophoresis.

S-allele	Detected Size (BPs) +/- Standard Deviation		
	Primer Set A	**Primer Set B**	**Primer Set C**
Sf	480.75 +/-0.27		
Si	1213.90 +/-17.31	894.34 +/-0.17	
S1	1081.60 +/-1.85	742.15 +/-0.85	
S3	1198.35 +/-6.95	894.87 +/-0.59	
S5	598.31 +/-0.94	264.184 +/-0.22	
S6	856.24 +/-2.39	543.092 +/-0.42	
S7			219.24 +/-0.27
S8			204.82 +/-0.8
S11	702.51 +/-1.04	379.07 +/-.06	
S14	1275.5 +/-32.8	966.59 +/-1.75	
S18	643.93 +/-0.76	313.64 +/-0.18	

Research Effort Recent Publications:

None.

References Cited:

ABC (2011) Almond industry position report, 2010-2011 crop year. Available at http://www.almondboard.com/files/PDFs/2011.06.pdf

Barckley, K., S. Uratus, T.M. Gradziel and A.M. Dandekar. 2006. Multidimensional analysis of S-alleles from cross-incompatible groups of California almond cultivars. J. Amer. Soc. Hort. Sci. 131(5):632-636.

Dangl, Gerald S., Judy Yang, Thomas Gradziel and Deborah A. Golino. 2009. A practical method for almond cultivar identification and parental analysis using simple sequence repeat markers. Euphytica 168 (1): 41-48.

Sánchez-Pérez, R., F. Dicenta and P. Martínez-Gómez. 2004. Identification of S-alleles in almonds using multiplex PCR. Euphytica 138:263-269.

Sutherland, B.G., T.P. Ribbins and K.R. Tobutt. 2004. Primers amplifying a range of *Prunus* S-alleles. Plant Breeding 123:582-584.

Tamura, M., K. Ushijima, H. Sassa, H. Hirano, R. Tao, T.M. Gradziel and A.M. Dandekar. 2000. Identification of self-incompatibility genotypes of almonds by allele specific PCR analysis. Theor. Appl. Genet. 101:344-349.

Kester, DE, T.M. Gradziel and W.C. Micke. 1994. Identifying pollen incompatibility groups in California almond cultivars. Journal of the American Society for Horticultural Science 119:106–109

Ushijima, K., H. Sassa, R. Tao, H. Yamane, A.M. Dandekar, T.M. Gradziel and H. Hirano. 1998. Cloning and characterization of cDNAs encoding S-RNases from almond (Prunus dulcis): primary structural features and sequence diversity of the S-RNase in Rosaceae. Mol. Gen. Genet. 260:261-268.