Multidimensional Analysis of S-alleles from Cross-incompatible Groups of California Almond Cultivars

Kristi K. Barckley, Sandra L. Uratsu, Thomas M. Gradziel, and Abhaya M. Dandekar
Department of Plant Sciences, University of California, Davis, One Shields Avenue, Davis, CA 95616

ADDITIONAL INDEX WORDS. self-incompatibility, S-RNase, Prunus dulcis

Abstract. The California almond industry is the largest supplier of almonds [Prunus dulcis (Miller) D.A. Webb] in the United States and throughout the world. Self-incompatibility is a major issue in almond production as it greatly affects nut set. In this study, we determined full-length sequences for alleles S0 - S11, determined the genotypes of 44 California cultivars, and assigned the cultivars to cross-incompatibility groups (CIGs). Newly identified S-alleles led to an increase in the number of CIGs. A pairwise distance tree was constructed using the aligned amino acid sequences showing their similarity. Four pairs of alleles (S0 and S2, S0 and S11, S1 and S5, and S1 and S9) showed high sequence similarity. Because of its simplicity, reproducibility, and ease of analysis, PCR is the preferred method for genotyping S-alleles.

According to the Almond Board of California (2004), near 470 million kg of almonds were harvested in California in 2004. One-third was consumed domestically, while 328 million kg of almonds were exported, making California the number one supplier of almonds worldwide.

One of the major issues surrounding the almond industry is a genetic phenomenon known as self-incompatibility (SI). This naturally occurring process is the reason growers need pollinizers. Without cross-compatible cultivars, almonds would not be naturally occurring process is the reason growers need pollinizers.

The California almond industry is the largest supplier of almonds [Prunus dulcis (Miller) D.A. Webb] in the United States and throughout the world. Self-incompatibility is a major issue in almond production as it greatly affects nut set. In this study, we determined full-length sequences for alleles S0 - S11, determined the genotypes of 44 California cultivars, and assigned the cultivars to cross-incompatibility groups (CIGs). Newly identified S-alleles led to an increase in the number of CIGs. A pairwise distance tree was constructed using the aligned amino acid sequences showing their similarity. Four pairs of alleles (S0 and S2, S0 and S11, S1 and S5, and S1 and S9) showed high sequence similarity. Because of its simplicity, reproducibility, and ease of analysis, PCR is the preferred method for genotyping S-alleles.

Materials and Methods

Plant material. Young almond leaves were collected and used to extract DNA from 44 cultivars (Table 1). Styles were collected during the popcorn stage of flower development and used for RNA extraction. The styles were frozen at -80 °C until needed. Samples were collected from cultivars at the Univ. of California orchards in Davis and Winters, Calif., and from growers throughout California.

DNA and RNA isolation. DNA was isolated from 100 mg of young almond leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, Calif.). CDNA synthesis. CDNA was converted to cDNA using the Marathon CDNA Amplification Kit (Clontech, Mountain View, Calif.).

Cloning and sequencing of DNA. Amplification of the S-alleles was accomplished using primers AS1II (5' TATTTCAATTT-GTGCACAAATGG3'), AmyCSR (5' CAAAATACCACTCAT-GTAAACAC3'), AlsC1 (5' CAGACACTTAATCAATCCAG3') and AlsC2 (5' GTGTATTCATGTTGCAGAAC3') as described by Tamura et al. (2000). Polymerase chain reaction (PCR) was used to amplify S-alleles from genomic DNA using the program: 1 min 94 °C, 2 min 55 °C, and 2 min 65 °C, 20-25 cycles. The PCR products were cloned into a pcR2.1 TOPO vector (Invitrogen, Carlsbad, Calif.). DNA from positive clones was extracted using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, Calif.) and sequenced by Davis Sequencing (Davis, Calif.). Sequence results were analyzed using Sequencer (Gene Codes Corp., Ann Arbor, Mich.) and MacDNAsis programs (Hitachi Software, San Bruno, Calif.).
5´-AND 3´-RACE. RACE was performed on S-alleles, S=S, using the Marathon cDNA Amplification Kit (Clontech). The desired product was purified using Gene Clean (QBiogene, Irvine, Calif.) and cloned. Reverse primers of ASI and AmyCSR (universal primers) were used for 3´- and 5´-RACE, respectively. Procedures for cloning, sequencing and analysis were as described above.

INTRON I AMPLIFICATION. Primers flanking intron I (see Table 2 for primer sequences) were designed from sequences obtained from 5´-RACE procedures. Intron I was amplified from DNA of young almond leaves using PCR (1 min 94 °C, 2 min 55 °C, and 2 min 65 °C, 25 cycles). The PCR product was TOPO cloned and sequenced as described above.

PROTEIN COMPARISONS. SRNase sequences were obtained from GenBank. Apple (Malus × domestica Borkh.) sequences were from accession numbers AY187627, AY039702 and AY201748. European almond S alleles were represented by accession numbers AAF73756, AAG09286, AAF82613, AAG09287, AAK58577, AAK58579, AAK58580 corresponding to sequences from Ma and Oliviera (2001, 2002) for S and S=S, respectively. California almond S alleles were from accession numbers (BAA95317, AAD51787, AAD51788, and AAD51789) corresponding to S=S=S, respectively (Tamura et al., 1999; Ushijima et al., 1998). For a list of accession numbers corresponding to sequences determined in this work refer to Table 3.

Four S-alleles, S=S, isolated from European cultivars were compared with California S-alleles S=S and percent homolog calculated. Amino acid sequences were aligned using Clustal X (Thompson et al., 1997). Three amino acid sequences from apple were aligned with the almond S-alleles. The pairwise distance tree, based on the aligned amino acid sequences, using the neighbor-joining method was constructed using PAUP* version 4.0 (Swofford, 2002). Bootstrap analysis of 1000 replicates was also completed. Additionally a pairwise distance tree was constructed with almond S-alleles determined by Ma and Oliviera (2002).

Results

GENETIC ANALYSIS OF S-ALLELES. Six S-alleles (S=S) were amplified by PCR, cloned, and sequenced. In addition, RACE was performed for all S-alleles (S=S) to determine sequences at the 5´ and 3´ ends. Large differences in the size of the two introns

Table 1. California almond cultivars placed into cross-incompatibility groups (CIGs) based upon their S-alleles. The cultivars listed within the same row or CIG designation are incompatible pollinizers, but are compatible with cultivars listed in differing CIGs. Bolded cultivars indicate previously genotyped cultivars and the cultivars genotyped in this manuscript are unbolded. Genotypes were based on PCR of intron length polymorphisms. Successful cross-compatibility between CIGs has been confirmed by growers and nurseries throughout California. While the European alleles are not 100% homologous to the California alleles, the European genotypes listed have either been reported using the same cultivars or based on alignment of all the sequence data are the most similar to the California genotypes listed.

| CIG | Genotype  | Almond cultivars                                      | European genotype |
|-----|-----------|-------------------------------------------------------|-------------------|
| I   | S=S       | Nonpareil, Tardy Nonpareil, Grace, West Steyn, UCD F8:7-180, Galaxy | S,S             |
| II  | S=S       | Mission                                              | S,S             |
| III | S=S       | Thompson, Sauret no.2, Mono, Wood Colony, Durango, Le Grand, Wassum | S,S             |
| IV  | S=S       | Merced, Ne Plus Ultra, Rosetta, Price cluster, Aldrich, Pearl, Jenette, Sano | S,S             |
| V   | S=S       | Carmel, Sauret no. 1, Livingston, Folsom, Butte, Dottie Won, Plateau, Avalon, UCD D3-25, UCD F8:7-179, Blue Gum | S,S             |
| VI  | S=S       | Monterey, Butte, Dottie Won, Plateau, Avalon, UCD D3-25, UCD F8:7-179, Blue Gum | S,S             |
| VII | S=S       | Arbutkle                                             | S,S             |
| VIII| S=S       | Fritz, Ruby, Peerless                                | S,S             |
| IX  | S=S       | Padre                                                | S,S             |
| X   | S=S       | UCD 13-1 (Winters), UCD 36-52                        | S,S             |
| XI  | S=S       | Tokyo                                                | S,S             |
| XII | S=S       | Milow                                                | S,S             |
| XIII| S=S       | Jordanolo                                            | S,S             |
| XIV | S=S       | Kochi, UCD F8:8-160                                  | S,S             |
| XV  | S=S       | Solano, Sonora, Vesta, Kapareil, UCD F8:7-161        | S,S             |
| XVI | S=S       | UCD 25-75                                            | S,S             |
| XVII| S=S       | Harriot                                              | S,S             |
| XVIII| S=S    | Carrion                                               | S,S             |
| XIX | S=S       | Jeffries, UCD 3-6, Johlyn                            | S,S             |
| XX  | S=S       | Tuono                                                | S,S             |
| XXI | S=S       | UCD 54P455 (peach)                                   | ---             |

References for alleles include Bošković et al., 1997; Channuntapipat et al., 2001, 2003; Ma and Oliveira et al., 2001.
were observed (Table 4). These differences in intron size allow characterization of S-alleles by PCR.

A comparison of amino acid sequences was performed (Fig. 1) to determine exon differences. It should be noted that Sb was not included in this analysis because it is derived from peach and is not a true almond allele. All the alleles maintained the overall structure of five conserved regions, two introns, and a hypervariable region, but showed individual differences, especially within the hypervariable region. More single amino acid changes seemed to occur in conserved region C4 than regions C1, 2, 3, or 5. Although there are still some unknown sequences at the 5’ and 3’ end of the alleles, the most important regions of the sequences are evident from the analysis.

**Cross-incompatible groups (CIGs).** The ability to identify S-alleles using PCR allows rapid genotyping of cultivars and their assignment to a CIG (Table 1). The European designations of the alleles are shown. Cultivars within a CIG are cross-incompatible and thus not suitable for use as pollinizers. Thus, Table 1 groups the cultivars in such a way as to help determine potential pollinizers based on the S-alleles. Any new almond cultivars can easily be tested and added to the chart.

**Pairwise distance tree.** Using amino acid alignment, a pairwise distance tree was constructed (Fig. 2). The tree shows that the apple sequences are genetically different from the almond S-alleles. Analysis of the almond S-alleles indicates that the Sg allele was perhaps a progenitor S-allele from which all others were derived, as it is an outlier in Fig. 2. It is interesting to note that the Sb allele is the only one that does not seem to be closely related to another S-allele. According to the bootstrap analysis, Sb of ‘Mission’ and Sf of ‘Tuono’ are the only pair that is highly similar. The other pairings did not derive very high bootstrap percentages and therefore may not be as strongly similar as the tree suggests.

**Discussion**

**Genetic analysis of S-RNases.** We successfully identified 10 S-alleles (Si–Sj) in all known California almond cultivars. As shown by Tamura et al. (2000), a simple PCR confirms each allele based on intron size differences (Table 4). Each allele has two introns that vary in length, allowing for easy detection and confirmation. Additionally, all S-alleles have the same structural sequence motifs identified for the original four S-alleles (S^a, S^b, S^c, S^d) (Tamura et al., 2000; Ushijima et al., 1998). According to Matton et al. (1997), the HV region is required for allelic specificity and for the pollen rejection mechanism. The aligned amino acid sequences show the HV region, which is interrupted by intron II, as being highly variable among S-alleles. Therefore, amino acid changes within conserved regions should not have a negative effect on the ability to recognize “self” pollen. Sequences at the 5’ end have been analyzed for almost all of the S-alleles, resulting in the identification of the signal peptide, an additional intron (intron I), and the start of the mature protein. The structure as identified by the amino acid comparisons remains similar among the 10 analyzed S-alleles (Fig 1). The Sb allele cloned in this study is derived from ‘Tuono’ and has an amino acid sequence identical to the Sb allele. While our data show that the exons are the same, the introns differ in size (Table 4), indicating that Sb differs from Sb. However, if the HV region is responsible for allele specificity, then Sb should be recognized as Sb as the HV regions are the same. The presence of only one allele, or one self-compatible allele, however, does indicate self-compatibility. In a self-compatible system, pollen is not recognized as “self,” and therefore pollen tube growth is not terminated. The cultivar Tuono expresses the S^b allele, a self-compatible allele (Ma and Oliveira, 2001), which has previously been shown to have the genotype SiSi (Bošković et al., 1997; Sanchez-Perez et al., 2004). However, according to our data, ‘Tuono’ does not express an Sb allele.

Historically, the European S-alleles are designated with numbers, while the California S-alleles are designated with letters. When the European alleles listed in GenBank were compared with the lettered California alleles, the exon sequences showed almost complete homology. The intron I size of European almond S-alleles was determined by Ortega et al. (2005). When compared with California S-alleles, the intron I lengths were found to be significantly different. Intron II lengths also differed (Channuntapid et al., 2001; Tamura et al., 2000). Additionally, Ma and Oliveira (2002) identified Sb, S^b, Si, and Sj alleles in European almond cultivars. When the sequences were compared with the sequences from California almonds (data not shown), the majority of homology occurred within the conserved regions. The Si allele had the least amount of homology (67%) between the European and California almonds, while the Sj allele had the most homology (98%).
homology (81%). Additionally, the Sf allele, cloned from 'Tuono' by Ma and Oliveira (2001) differs in sequence from the Sf allele isolated in this study using the same cultivar. Therefore, it can be said that the alleles found in European and Californian cultivars are similar but not identical.

CROSS-INCOMPATIBLE GROUPS (CIGs). This study added 15 CIGs to the six previously identified by Kester et al. (1994). The original CIGs were determined through cross-pollinations and zymograms (Kester et al., 1994). The genotypes of 'Ruby', 'Aldrich', and 'Pearl' (cultivars also analyzed in this study) were previously determined using zymograms (Bošković et al., 1997, 1999, 2003; Channuntapipat et al., 2001). Due to the difference in methods, discrepancies have arisen between previous genotype designations (Bošković et al., 1997, 1999, 2003; Channuntapipat et al., 2001; Kester et al., 1994) and the genotypes/CIGs assigned in this research. Previously the genotype designations as determined by Bošković et al. (1997, 1999, 2003), Channuntapipat et al. (2001; Kester et al., 1994) and the genotypes/CIGs assigned in this research. Previously the genotype designations as determined by Bošković et al. (1997, 1999, 2003), Channuntapipat et al. (2001), and Kester et al. (1994), were 'Ruby' (Sf/Sf), 'Aldrich' (Sf/Sf), and 'Pearl' (Sf/Sf). The CIGs designated in this study ('Ruby' (Sf/Sf), 'Aldrich' (Sf/Sf), and 'Pearl' (Sf/Sf)) were a product of a multidimensional analysis using PCR to genotype the almond cultivars, followed by sequencing of the alleles (cDNA and gDNA) for specific confirmation. Our CIGs are consistent with the observations at the Univ. of California almond breeding program (directed by T. Gradziel) and several individuals in the California almond industry. Intron I amplification was used for additional confirmation. The PCR and
sequencing method gave clear reproducible results. These methods have been used to determine S-alleles in such systems as sweet cherry (Prunus avium L.), sour cherry (Prunus cerasus L.), and apricot (Prunus armeniaca L.) (Hauck et al., 2002; Sonneveld et al., 2001; Tao et al., 2002; Wunsch and Hormaza, 2004). Because of its simplicity, reproducibility, and ease of analysis, PCR is the preferred method for genotyping S-alleles.

**Pairwise distance tree.** Ushijima et al. (1998) compared S-RNases from different Rosaceae species in terms of their evolution. Our pairwise distance tree breaks down the larger picture by examining the relationships between individual S-alleles. Four pairs of alleles (Sf and Ss, S7 and S8, Sf and Ss, and Sf and Sf) have very few nucleotide substitutions, signifying that they are closely related. Ma and Oliveira (2002) formulated a phylogenetic tree showing the evolutionary relationships between S-alleles in various Rosaceae species. The relationships between S-alleles as determined by Ma and Oliveira (2002) do not correlate with our results. Additionally, Ma and Oliveira (2002) added additional sequences from other species. When the almond S-allele sequences were aligned and put into a pairwise distance tree using the same method for constructing our tree, a different output resulted (data not shown). The tree differed from the published tree, which was expected, but it also differed from our pairwise distance tree. The differences between the European S-allele tree we constructed and the tree in this manuscript are partly due to the differences in sequences used. The pairwise distance tree constructed from our amino acid alignment was based on full-length sequences whereas the sequences listed by Ma and Oliveira (2002) are only partial sequences and are not 100% homologous with the sequences aligned in this manuscript. Due to these variances, different clustering of alleles was expected.

**Conclusion**

This research expanded the number of known S-alleles in California almond cultivars and also increased the number of CIGs. PCR-based genotyping of almond cultivars remains the preferred method to determine their cross-compatibility. Until future research can expand our knowledge of the self-incompatibility system, genotyping of almond cultivars is the ideal way to determine the compatibility of cultivars.

**Literature Cited**

Almond Board of California. 2004. Almond almanac. Almond Board of California, Modesto, Calif.

Bošković, R., K.R. Tobutt, I. Batlle, and H. Duval. 1997. Correlation of ribonuclease zymograms and incompatibility of genotypes in almond. Euphytica 97:167–176.

Bošković, R., K.R. Tobutt, H. Duval, I. Batlle, F. Vicenta, and F.J. Vargas. 1999. A styra ribonuclease assay to detect self-compatible seedlings in almond progenies. Theor. Appl. Genet. 99:800–810.

Bošković, R., K.R. Tobutt, I. Batlle, H. Duval, P. Martinez-Gomez, and T.M. Gradziel. 2003. Styral ribonuclease alleles in almond: correlation with and prediction of incompatibility genotypes. Plant Breeding 122:70–76.

Channuntapipat, C., M. Sedgley, and G. Collins. 2001. Sequences of cDNAs and genomic DNAs encoding the S1, S7, S8, and SF alleles from almond, Prunus dulcis. Theor. Appl. Genet. 103:1115–1122.

Channuntapipat, C., M. Wirthsohn, S.A. Ramesh, I. Batlle, P. Arus, M. Sedgley, and G. Collins. 2003. Identification of incompatibility genotypes in almond (Prunus dulcis Mill.) using specific primers based on the introns of the S-alleles. Plant Breeding 122:164–168.

de Nettancourt, D. 1977. Incompatibility and incongruity in wild and cultivated plants. Springer-Verlag, New York.

Dicenta, F. and J.E. Garcia. 1993. Inheritance of self-compatibility in almond. Heredity 70:313–317.

Hauck, N., H. Yamane, R. Tao, and A.F. Iezzoni. 2002. Self-compatibility and incompatibility in tetraploid sour cherry (Prunus cerasus L.). Sexual Plant Reproduction 15:39–46.

Igoer, T.R., J.R. Gohlke, B. Xu, and T.H. Kao. 1991. Primary structural features of the self-incompatibility protein in Solanaceae. Sexual Plant Reproduction 4:81–87.

Kester, D.E., T.M. Gradziel, and W.C. Micke. 1994. Identifying pollen incompatibility groups in California almond cultivars. J. Amer. Soc. Hort. Sci. 119:106–109.

Lewis, D. and K. Crowe. 1954. The induction of self-fertility in tree fruits. J. Hort. Sci. 29:220–225.

Ma, R.-C. and M.M. Oliveira. 2001. Molecular cloning of the self-incompatibility genes S1 and S3 from almond (Prunus dulcis cv. Ferragnes). Sexual Plant Reproduction 14:163–167.

Ma, R.-C. and M.M. Oliveira. 2002. Evolutionary analysis of S-RNase genes from Rosaceae species. Mol. Genet. Genomics 267:71–78.

Matton, D.P., O. Maes, G. Laublin, Q. Xike, C. Bertrand, D. Morse, and M. Cappadocia. 1997. Hypervariable domains of self-incompatibility RNases mediate allele-specific pollen recognition. Plant Cell 9:1757–1766.

McCubbin, A.G. and T.H. Kao. 2000. Molecular recognition and response in pollen and pistil interactions. Annu. Rev. Cell. Dev. Biol. 16:333–364.

Ortega, E., B.G. Sutherland, F. Vicenta, R. Bošković, and K.R. Tobutt. 2005. Determination of incompatibility genotypes in almond using first and second intron consensus primers: Detection of new S alleles and correction of reported S genotypes. Plant Breeding 124:188–196.

Sanchez-Perez, R., F. Vicenta, and P. Martinez-Gomez. 2004. Identification of S-alleles in almond using multiplex PCR. Euphytica 138:263–269.

Silva, N.F. and D.R. Goring. 2001. Mechanisms of self-incompatibility in flowering plants. Cell. Mol. Life Sci. 58:1988–2007.

Sonneveld, T., T.P. Robbins, R. Bošković, and K.R. Tobutt. 2001. Cloning of six cherry self-incompatibility alleles and development of allele-specific PCR detection. Theor. Appl. Genet. 102:1046–1055.

Stone, J.L. 2002. Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. Orty. Rev. Biol. 77:17–32.

Stone, S.L. and D.R. Goring. 2001. The molecular biology of self-incompatibility systems in flowering plants. Plant Cell Tissue Organ Cult. 67:93–114.

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Mass.

Tamura, M., T.M. Gradziel, and A.M. Dandekar. 1999. Cloning of genomic DNA sequences encoding almond (Prunus dulcis) S-RNase genes. Plant Physiol. 120:1206.

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

Tao, R., H. Yamane, H. Sassa, H. Mori, T.M. Gradziel, and A.M. Dandekar. 1997. Identification of styra RNase associated with gametophytic self-incompatibility in almond (Prunus dulcis). Plant Cell Physiol. 38:304–311.

Tao, R., T. Habu, H. Yamane, and A. Sugiuira. 2002. Characterization and cDNA cloning for Sf-RNase, a molecular marker for self-compatibility, in japanese apricot (Prunus mume). J. Jpn. Soc. Hort. Sci. 71:595–600.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.

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-RNases in Rosaceae. Mol. Gen. Genet. 260:261–268.

Wunsch, A. and J.I. Hormaza. 2004. S-allele identification by PCR analysis in sweet cherry cultivars. Plant Breeding 123:327–331.