American Journal of Botany 84(8): 912–917. 1997.

SELF-INCOMPATIBILITY RNASES FROM THREE PLANT FAMILIES: HOMOLOGY OR CONVERGENCE?[1]

ADAM D. RICHMAN,[2] WIM BROOTHAERTS,[3] AND JOSHUA R. KOHN[2,4]

[2]University of California at San Diego, Department of Biology 0116, 9500 Gilman Drive, La Jolla, California 92093-0116; and [3]Fruitteeltcentrum, Katholieke Universiteit Leuven, W. de Croylaan 42, B-3001 Leuven, Belgium

In the Rosaceae, Scrophulariaceae, and Solanaceae, the stylar product of the self-incompatibility (S-) locus is an RNase. Using protein sequence data from 34 RNase genes (three fungal RNases, seven angiosperm non-S RNases, 11 Rosaceae S-alleles, three Scrophulariaceae S-alleles, and ten Solanaceae S-alleles) we reconstructed the genealogy of angiosperm RNases using the neighbor joining method and two distance metrics in order to assess whether use of S-RNases in these families is the result of homology or convergence. Four monophyletic groups of angiosperm RNases were found: the S-RNases of each of the three families and a group comprising most of the angiosperm non-S RNases. The S-RNases of the Scrophulariaceae and Solanaceae were found to be homologous but strong inference concerning the homology or convergence of S-RNases from the Rosaceae with those of the other families was not possible because of uncertain placement of both the root and two of the angiosperm non-S RNases. The most recent common ancestor of the Rosaceae and both the Scrophulariaceae and Solanaceae is shared by ~80% of dicot families. If the -RNases of the Rosaceae are homologous to those of the Scrophulariaceae and Solanaceae, then many other dicot families might be expected to share RNases as the mechanism of gametophytic self-incompatibility.

Key words: Rosaceae; S-alleles; S-locus; Scrophulariaceae; self-incompatibility; Solanaceae.

Whitehouse (1951) proposed that all physiological self-incompatibility mechanisms in flowering plants arose from a single system present in the most ancestral angiosperm and that incompatibility was a key feature leading the angiosperms to diversify and dominate terrestrial habitats (see also de Nettancourt, 1977). More recently, the isolation of different classes of molecules as the products of the self-incompatibility (S-) locus in the Brassicaceae (Nasrallah et al., 1985), Papaveraceae (Foote et al., 1994), Solanaceae (Anderson et al., 1986), and Poaceae (Li et al., 1994) strongly implies multiple evolutionary origins of incompatibility (Uyenoyama, 1995; Weller, Donoghue, and Charlesworth, 1995). Nevertheless, the number of times incompatibility evolved and which families, if any, share incompatibility by homology remain open questions. In the Solanaceae, the stylar S-gene product is an RNase and evidence is mounting that the incompatibility reaction involves destruction of self-pollen RNA (McClure et al., 1990; Huang et al., 1994; Lee, Huang, and Kao, 1994). The recent discovery that RNases are the stylar gene products of the S-locus in two additional plant families, the Rosaceae (Sassa, Hirano, and Ikehashi, 1992; Broothaerts et al., 1995) and the Scrophulariaceae (Xue et al., 1996), begs the question of whether these families share this trait due to common ancestry or whether there has been convergent evolution of the use of RNases for self-pollen rejection.

Two conflicting viewpoints have been published concerning the homology of S-RNases among the three angiosperm families known to use them. Sassa et al. (1996) performed a genealogical analysis of angiosperm RNase sequences and found three distinct groups: S-alleles of the Rosaceae, S-alleles of the Solanaceae, and angiosperm non-S RNases. Sassa et al. (1996) concluded that use of RNases in incompatibility in the Solanaceae and Rosaceae represents convergence. However, without reference to an outgroup, it remained possible that the S-alleles from these two families could form sister groups, a result that would support homology. Xue et al. (1996) isolated and sequenced three S-RNases from *Antirrhinum hispanicum,* a member of the Scrophulariaceae. Gene genealogical analysis found that S-RNases from each of the three angiosperm families formed separate monophyletic groups. Rooting their genealogy using the fungal T2 type RNases, Xue et al. (1996) found that angiosperm non-S RNases formed a sister group to a clade comprising the S-RNases from all three families, a topology consistent with a single origin of the incompatibility systems in these three families. Xue et al. (1996) were careful to point out, however, that alternative topologies that implied that the S-RNases of the Rosaceae were of independent origin from those of the Solanaceae and Scrophulariaceae were possible. Because the findings of these studies conflicted and because neither study attempted a statistical analysis of the certainty of genealogical inference, we undertake here an analysis of the problem using S-RNase sequences from all three plant families, non-S RNase sequences from a variety of plants, and fungal T2 type RNases.

The Scrophulariaceae and Solanaceae belong to the subclass Asteridae and are more closely related to one another than either is to the Rosaceae, a member of the Rosidae subclass (Cronquist, 1981; Chase et al., 1993). Descendants of the most recent common ancestor of the Rosidae and Asteridae comprise ~ 80% of all dicot families (Cronquist, 1981; Chase et al., 1993). Therefore, if the S-RNases from all three families are homologous, a

[1] Manuscript received 10 July 1996; revision accepted 6 January 1997.

This work was supported by NSF awards DEB 93-06473, DEB-9404386, and DEB 95-27834 and a Hellman Fellowship from the University of California, San Diego.

[4] Author for correspondence: 619-534-8233 (phone), 619-534-7108 (FAX), jkohn@ucsd.edu (email).

large number of angiosperm families could share this form of gametophytic self-incompatibility through common ancestry. On the other hand, if there have been one or more parallel conversions of RNases to incompatibility function, one must ask why RNases have repeatedly been adopted for this function? One suggestion is that RNases used in defense against invasion of floral organs by fungal or other pathogens have been converted to function in incompatibility (Lee, Singh, and Kao, 1992). If so, then characterization of RNase defense genes could lead to discovery of the antecedents of the S-locus.

MATERIALS AND METHODS

Thirty-five amino acid sequences were used in this analysis: three fungal T2 type RNases, seven non-S angiosperm RNases, 11 S-alleles from apple and Japanese pear (Rosaceae), three S-alleles from *Antirrhinum hispanicum* (Scrophulariaceae), and ten S-alleles from the Solanaceae, which are representative of the range of sequence diversity found among alleles in this family (Richman, Uyenoyama, and Kohn, 1996). In addition, we included RNase X2, a non-S RNase isolated from styles of *Petunia* (Solanaceae), which was previously reported to be most closely related to the *Nicotiana alata* S6 allele (Lee, Singh, and Kao, 1992; see Fig. 1 for names and sources of all sequences). Amino acid alignment of sequences (Fig. 1) was performed by Clustal W with the default settings and then adjusted by eye. Because of difficulty in alignment of terminal regions of divergent sequences, all analyses were performed on sequences beginning with the first amino acid of conserved region 1 (Ioerger et al., 1991) and ending with the amino acid following the last conserved cysteine residue (Fig. 1). This resulted in omission of one N-terminal amino acid and ~ 16 C-terminal amino acids for functional S-proteins from the Solanaceae and varying amounts of truncation in other sequences. The neighbor joining method was used for genealogical reconstruction. Protein distances among pairs of sequences were produced in two ways, either using percentage amino acid divergence or using the PAM 001 Dayhoff matrix implemented by the PROTDIST program in PHYLIP (Felsenstein, 1993). The PAM 001 Dayhoff matrix weights amino acid transitions according to their empirical frequency. For each distance metric, bootstrap analysis was performed by randomly drawing amino acid positions, with replacement, to produce 100 replicate protein distance matrices upon which neighbor joining was performed.

RESULTS

Several groups of sequences formed monophyletic clades with strong support among bootstrap replicates in analyses using both distance measures (Fig. 2). These include the three fungal RNases, the S-RNases from each family, and a group of five non-S plant RNases (LE, LX, NE, RNS1, and RNS3). In both analyses the non-S RNase RNS2 was sister to this group, although this position received only weak bootstrap support (55%). The only non-S plant RNases not in this clade were RNase MC, whose genealogical placement was uncertain with no single placement found among a majority of bootstrap replicates, and RNase X2, which was sister to the *Nicotiana alata* S6 (see also Lee, Singh, and Kao, 1992). In both analyses, S-alleles from the Scrophulariaceae and the Solanaceae formed sister groups. However, the relative positions of the non-S RNases (except X2), the S-alleles from the Rosaceae, and the position of the root were poorly supported (Fig. 2: note the short branch lengths and poor bootstrap support of basal nodes in both analyses).

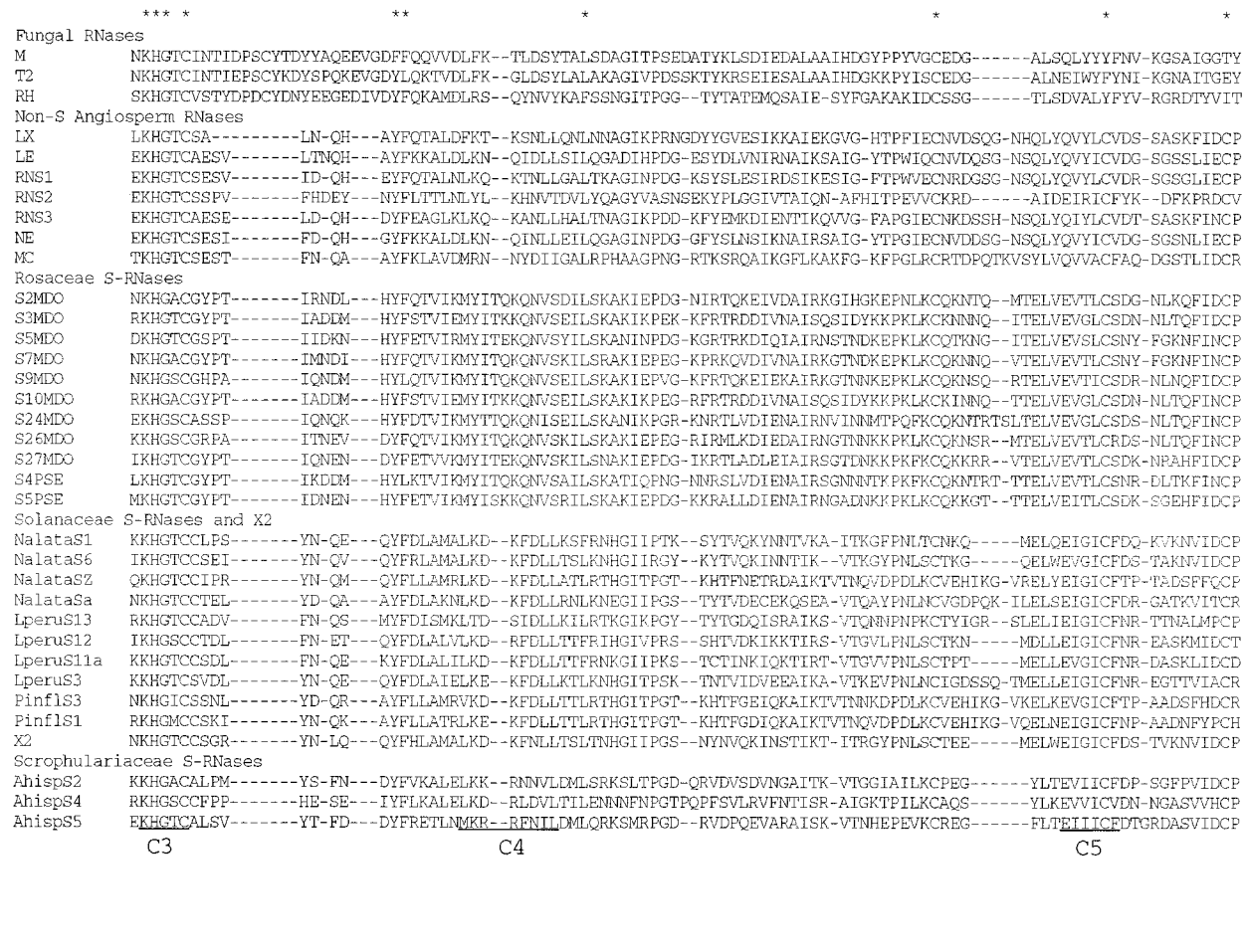
Lack of bootstrap support for nodes near the root could result from variation in root position, variation in the placement of RNase MC and RNS2, and variation in the relative positions among the four well-supported clades (the three families of S-alleles and the group of five non-S RNases). To examine the extent of uncertainty due to root position, we analyzed the data using percentage amino acid differences in the absence of the fungal sequences in order to obtain unrooted networks of ingroup (angiosperm) sequences. We obtained strong support (95%) for the grouping of alleles from the Scrophulariaceae and Solanaceae so that alleles from the Rosaceae grouped with all non-S RNases except X2. However, the placement of RNase MC with respect to the other non-S RNases and S-alleles from the Rosaceae remained uncertain. RNase MC joined the network at a node basal to the Rosaceae S-RNases (33% of bootstrap replicates) with approximately the same bootstrap frequency that it joined at a node basal to the other non-S RNases (31% of replicates). Thus, it appears that the uncertainty in basal nodes of outgroup analyses (Fig. 2) results from both poor resolution of the root position and the uncertain placement of RNase MC.

DISCUSSION

Genealogical analysis of protein sequences from plant RNases fails to unequivocally support either homology or convergence as explanations for the use of S-RNases in three angiosperm families. Analyses using either distance metric support the hypothesis (Xue et al., 1996) that the S-RNases from the Scrophulariaceae and Solanaceae are homologous. Neither the root nor any of the non-S RNases separate the alleles from these two families and topologies that contradict the inference of homology occur in only a small minority of bootstrap replicates (Fig. 2). However, whether the Rosaceae S-RNases are homologous or convergent with respect to the S-RNases from the other two families cannot be determined with certainty. With the exception of the uncertain placement of RNase MC, the topology produced using the PAM Dayhoff matrix (Fig. 2, right side) would support the inference that all S-RNases are homologous. That is, if all non-S RNases other than X2 form a clade sister to all S-RNases, then homology would be inferred. However, if the placement of RNase MC in the PAM Dayhoff analysis is accepted, then either RNase MC represents an ancient conversion of an S-allele to a different function (as RNase X2 apparently represents a more recent conversion) or the S-RNases from the Rosaceae represent a parallel evolutionary conversion of RNases to function in self-pollen rejection. The topology produced by the analysis of mean percentage amino acid differences (Fig. 2, left side) makes the S-alleles from the Rosaceae sister to all angiosperm RNases. Under this topology, the most reasonable interpretation would be that the S-locus of the S Rosaceae results from convergence rather than homology to the S-locus of the Scrophulariaceae and Solanaceae. The alternative interpretation, that all S-loci are homologous, would force the inference that all known angiosperm non-S RNases represent reversions from the S-locus.

Perhaps the most surprising result is the fact that most non-S plant RNases (RNS1, RNS3, LE, LX, and NE),


Fungal RNases
M EASCCFNSPGGSLLQTQF-WDYD-----PSDG-PSDSWTIHGLWPDNCDGSYQEY--CDDSREYSNITSILEAQDRTTELLSYMKEYWPD-YEGAD----EDE--SFWEHEW
T2 ADSCCFNSPGGALLQTQF-WDTN-----PPSG-PSDSWTIHGLWPDNCDGSYGQF--CDKTREYSNITAILQEQGR-TELLSYMKKYWPN-YEGDD----EE----FWEHEW
RH SDTCC--SPEYGLVVLNMQWA-------PGYG-PDNAFTLHGLWPDKCSGAYAPSGGCDSNR--ASSSIASVIKSKDSSLYNSMLTYWPS-NQGN-----NN---VFWSHEW
Non-S Angiosperm RNases
LX FDFFYFVQQWPASYCDTRRSC-CY----PTTGKPDEDFSIHGLWPNYKDGKWPQN--CD--------RESSLDESEFSDLISTMEKNWPS-LACPSS---DGL--KFWSHEW
LE FDFFYFVQQWPGSYCDTKQSC-CY----PTTGKPAADFGIHGLWPNNNDGTYPSN--CD--------PNSPYDQSQISDLISSMQQNWPT-LACPSG---SGS--TFWSHEW
RNS1 FDFFYFVQQWPGSYCDTQKKC-CY----PNSGKPAADFGIHGLWPNYKDGTYPSN--CD--------ASKPFDSSTISDLLTSMKKSWPT-LACPSG---SGE--AFWEHEW
RNS2 FDYFALSLQWPGTYCRGTRHC-CSKNACCRGSDAPTQFTIHGLWPDYNDGSWPSC--CY---------RSDFKEKEISTLMDGLEKYWPS-LSCGSPSSCNGGKGSFWGHEW
RNS3 FDFFYFVLQWPGAYCDSRHSC-CY----PQTGKPAADFGIHGLWPNYKTGGWPQN--CN--------PDSRFDDLRVSDLMSDLQREWPT-LSCPSN---DGM--KFWTHEW
NE FDFFYFVQQWPGSYCDTKQSC-CY----PKTGKPASDFGIHGLWPNNNDGSYPSN--CD--------SNSPYDQSQVSDLISRMQQNWPT-LACPSG---TGS--AFWSHEW
MC FDSFWFVQQWPPAVCSFQKSGSC-----PGSG-L-RTFTIHGLWPQGS-GTSLTN--CP--------QGSPFDITKISHLQSQLNTLWPNVLRA-----NNQ--QFWSHEW
Rosaceae S-RNases
S2MDO YDYFQFTQQYQPAACNSNPTP-CKD---P----PDKLFTVHGLWPSNMNRSELFN--CS--------SSNVT-YAKIQNIRTQLEMIWPNVFNRK-----NHL--GFWNREW
S3MDO FDYFQFTQQYQPAVCSSNPTP-CKD---P----PDKLFTVHGLWPSNVNGSDPKK--CK--------TTILN-PQTITNLTAQLEIIWPNVLNRK-----AHA--RFWRKQW
S5MDO ----------D---P----PDKLFTVHGLWPSNFNGPDPEN--CK--------VKPTASQTIDTSLKPQLEIIWPNVFNRA-----DHE--SFWQKQW
S7MDO YDYFQFTQQYQLAACNSKPIP-CKD---P----PDKLFTVHGLWPSDSNGHDPVN--CS--------SSNVT-YAKIQNIRTQLEMIWPNVFNRK-----NHL--GFWNREW
S9MDO YDYFQFTQQYQPAVCHFNPTP-CRD---P----PDKLFTVHGLWPSNSSGNDPIY--CK--------NTTMN-STKIANLTARLEIIWPNVLDRT-----DHI--TFWNKQW
S10MDO FDYFQFTQQYQPAVCNSNPTP-CKD---P----PDKLFTVHGLWPSNVNGSDPKK--CK--------ATILN-PQTITNLKAQLEIIWPNVLNRR-----AHV--RFWRKQW
S24MDO FDYYQFTQQYQPAVCNSNPTP-CKD---P----TDKLFTVHGLWPSNSNGNDPKY--CN--------AQQY---QTMKILEPQLVIIWPNVLNRN-----DHE--GFWRKQW
S26MDO YDYFQFTQQYQPAVCNSNPTP-CKD---P----PDKLFTVHGLWPSNSTGRDPKY--CN--------PSNVT-SHMLKNIQAQLEIIWPNVLNRT-----DHI--GFWERQW
S27MDO FDYFQFTQQYQPAVCNFNPTP-CKD---P----TDKLFTVHGLWPSNNVGGDPES--CK--------IRNHR--TRAKALEPQLEIIWPNVLDRT-----NHT--AFWRRQW
S4PSE FDYFQFTQQYQPAVCNSNPTP-CND---P----TDKLFTVHGLWPSNRNGPDPEK--CK--------TTTMN-SQKIGNMTAQLEIIWPNVLNRS-----DHV--GFWEREW
S5PSE YDYFQFTQQYQLAVCNSNRTP-CKD---P----PDKLFTVHGLWPSSMAGPDPSN--CP--------IRNIR--KREKLLEPQLAIIWPNVFDRT-----KNK--LFWDKEW
Solanaceae S-RNases and X2
NalataS1 FEYMQLVLTWPTAFCN--VMN-CER--------TPTNFTIHGLWPDNV-STELNY--CD--------RQKKFKLFEDDKKQNDLDDRWPD-LTLDRDDCKNGQ--GFWSYEY
NalataS6 FEYMQLVLQWPTAFCH--TTP-CKN--------IPSNFTIHGLWPDNV-STTLNF--CG--------KEDDYNIIMDGPEKNGLYVRWPD-LIREKADCMKTQ--NFWRREY
NalataSZ FDYMQLVLTWPASFCY-PKNF-CSR---I----APKNFTIHGLWPDKV-RGRLQF--CT--------SEKYVNFAQDSPILDDLDHHWME-LKYHRDFGLENQ--FLWRGQY
NalataSa ----------IPKNFTIHGLWPDEQ-HGMLND--CG--------ETFTKLEPREKK---ELTIRWPD-LKRSRSDAQDVE--SFWEYEY
LperuS13 FDHWQLVLTWPAGFCK-TKD--CPRKD------IPNNFTIHGVWPDHT-SFVMYD--CD--------PLKKYKTIDDTNILTELDARWPQ-LTSTKIIGLQFQ--RFWEYEY
LperuS12 --QLQLVLRWPTSFCN-GKN--CKR--------TPKDFTIHGLWPDSE-AGELNF--CN--------PRASYTIVRHGTFEK-RNKHWPD-LMRSKDNSMDNQ--EFWKHEY
LperuS11a ----------HGLWPDIK-GTILNN--CN--------PDAKYASVTGGKFVK-RNKHWPD-LILTEAASLNSQ--GFWAYQF
LperuS3 FDYLQLVLQWPRSFCK--TRY-CPN---P----VPRNFTIHGLWPDKQ-RIMPIN--CP--------AKESYKSITDSKKIKLLEQHWPD-LTSN----QGSA--EFWRYQY
PinflS3 FDYIQLVLTWPASFCYRPKNI-CRR--------IPNNFTIHGLWPEKE-HFRLEF--CD--------GDKFVSFSLKDRIVNDLERHWVQ-MKFDEKFAKIKQ--PLWTHEY
PinflS1 FEYLQLVLTWPASFCFRPKNI-CKR---P----A-KNFTIHGLWPEIT-GFRLEF--CT--------GSPKYETFKDNNIVDYLERHWVQ-MKFDENYAKYHQ--PLWSYEY
X2 YEYMQLVLQWPTAFCH-ASPT-CKV--------TPNNFTIHGLWPDNV-STTLNY--CKS-------KTGKYNNIKDPTIKNELYKRWPD-LTTSETDCLGNQ--NFWKREY
Scrophulariaceae S-RNases
AhispS2 FDYFKLVLQWPNSYCSLKTTH-CPR--TR----LPSQFTIHGLWPDNK-SWPLSN--CRD-------TSADVLKITDKGLIQDLAVHWPD-LTRR-QRKVPGQ--KFWVTQW
AhispS4 CDYLKLVLQWPKSFCLINSRK-CQR--NP----LPSNFTIHGLWPDNY-TRQAPQS-C---------TTNNFQRFTDTDIEQRMEESWPD-LKQQ-SIAGLSY--NFWQDQW
AhispS5 FEILKLVLQWPNSYCSLKTST-CRR--NP----LPLKFTIHGLWPDNY-SWPLSD--CG--------YDFTLPDITDKSLLKRLDRNWPD-LTKR-KNIRKPD-KTFWLTQW
C1 C2 HVa HVb


Fungal RNases
M NKHGTCINTIDPSCYTDYYAQEEVGDFFQQVVDLFK--TLDSYTALSDAGITPSEDATYKLSDIEDALAAIHDGYPPYVGCEDG------ALSQLYYYFNV-KGSAIGGTY
T2 NKHGTCINTIEPSCYKDYSPQKEVGDYLQKTVDLFK--GLDSYLALAKAGIVPDSSKTYKRSEIESALAAIHDGKKPYISCEDG------ALNEIWYFYNI-KGNAITGEY
RH SKHGTCVSTYDPDCYDNYEEGEDIVDYFQKAMDLRS--QYNVYKAFSSNGITPGG--TYTATEMQSAIE-SYFGAKAKIDCSSG------TLSDVALYFYV-RGRDTYVIT
Non-S Angiosperm RNases
LX LKHGTCSA---------LN-QH---AYFQTALDFKT--KSNLLQNLNNAGIKPRNGDYYGVESIKKAIEKGVG-HTPFIECNVDSQG-NHQLYQVYLCVDS-SASKFIDCP
LE EKHGTCAESV-------LTNQH---AYFKKALDLKN--QIDLLSILQGADIHPDG-ESYDLVNIRNAIKSAIG-YTPWIQCNVDQSG-NSQLYQVYICVDG-SGSSLIECP
RNS1 EKHGTCSESV-------ID-QH---EYFQTALNLKQ--KTNLLGALTKAGINPDG-KSYSLESIRDSIKESIG-FTPWVECNRDGSG-NSQLYQVYLCVDR-SGSGLIECP
RNS2 EKHGTCSSPV-------FHDEY---NYFLTTLNLYL--KHNVTDVLYQAGYVASNSEKYPLGGIVTAIQN-AFHITPEVVCKRD------AIDEIRICFYK--DFKPRDCV
RNS3 EKHGTCAESE-------LD-QH---DYFEAGLKLKQ--KANLLHALTNAGIKPDD-KFYEMKDIENTIKQVVG-FAPGIECNKDSSH-NSQLYQIYLCVDT-SASKFINCP
NE EKHGTCSESI-------FD-QH---GYFKKALDLKN--QINLLEILQGAGINPDG-GFYSLNSIKNAIRSAIG-YTPGIECNVDDSG-NSQLYQVYICVDG-SGSNLIECP
MC TKHGTCSEST-------FN-QA---AYFKLAVDMRN--NYDIIGALRPHAAGPNG-RTKSRQAIKGFLKAKFG-KFPGLRCRTDPQTKVSYLVQVVACFAQ-DGSTLIDCR
Rosaceae S-RNases
S2MDO NKHGACGYPT-------IRNDL---HYFQTVIKMYITQKQNVSDILSKAKIEPDG-NIRTQKEIVDAIRKGIHGKEPNLKCQKNTQ--MTELVEVTLCSDG-NLKQFIDCP
S3MDO RKHGTCGYPT-------IADDM---HYFSTVIEMYITKKQNVSEILSKAKIKPEK-KFRTRDDIVNAISQSIDYKKPKLKCKNNNQ--ITELVEVGLCSDN-NLTQFIDCP
S5MDO DKHGTCGSPT-------IIDKN---HYFETVIRMYITEKQNVSYILSKANINPDG-KGRTRKDIQIAIRNSTNDKEPKLKCQTKNG--ITELVEVSLCSNY-FGKNFINCP
S7MDO NKHGACGYPT-------IMNDI---HYFQTVIKMYITQKQNVSKILSRAKIEPEG-KPRKQVDIVNAIRKGTNDKEPKLKCQKNNQ--VTELVEVTLCSNY-FGKNFINCP
S9MDO NKHGSCGHPA-------IQNDM---HYLQTVIKMYITQKQNVSEILSKAKIEPVG-KFRTQKEIEKAIRKGTNNKEPKLKCQKNSQ--RTELVEVTICSDR-NLNQFIDCP
S10MDO RKHGACGYPT-------IADDM---HYFSTVIEMYITKKQNVSEILSKAKIKPEG-RFRTRDDIVNAISQSIDYKKPKLKCKINNQ--TTELVEVGLCSDN-NLTQFINCP
S24MDO EKHGSCASSP-------IQNQK---HYFDTVIKMYTTQKQNISEILSKANIKPGR-KNRTLVDIENAIRNVINNMTPQFKCQKNTRTSLTELVEVGLCSDS-NLTQFINCP
S26MDO KKHGSCGRPA-------ITNEV---DYFQTVIKMYITQKQNVSKILSKAKIEPEG-RIRMLKDIEDAIRNGTNNKKPKLKCQKNSR--MTELVEVTLCRDS-NLTQFINCP
S27MDO IKHGTCGYPT-------IQNEN---DYFETVVKMYITEKQNVSKILSNAKIEPDG-IKRTLADLEIAIRSGTDNKKPKFKCQKKRR--VTELVEVTLCSDK-NPAHFIDCP
S4PSE LKHGTCGYPT-------IKDDM---HYLKTVIKMYITQKQNVSAILSKATIQPNG-NNRSLVDIENAIRSGNNNTKPKFKCQKNTRT-TTELVEVTLCSNR-DLTKFINCP
S5PSE MKHGTCGYPT-------IDNEN---HYFETVIKMYISKKQNVSRILSKAKIEPDG-KKRALLDIENAIRNGADNKKPKLKCQKKGT--TTELVEITLCSDK-SGEHFIDCP
Solanaceae S-RNases and X2
NalataS1 KKHGTCCLPS-------YN-QE---QYFDLAMALKD--KFDLLKSFRNHGIIPTK--SYTVQKYNNTVKA-ITKGFPNLTCNKQ-----MELQEIGICFDQ-KVKNVIDCP
NalataS6 IKHGTCCSEI-------YN-QV---QYFRLAMALKD--KFDLLTSLKNHGIIRGY--KYTVQKINNTIK--VTKGYPNLSCTKG-----QELWEVGICFDS-TAKNVIDCP
NalataSZ QKHGTCCIPR-------YN-QM---QYFLLAMRLKD--KFDLLATLRTHGITPGT--KHTFNETRDAIKTVTNQVDPDLKCVEHIKG-VRELYEIGICFTP-TADSFFQCP
NalataSa NKHGTCCTEL-------YD-QA---AYFDLAKNLKD--KFDLLRNLKNEGIIPGS--TYTVDECEKQSEA-VTQAYPNLNCVGDPQK-ILELSEIGICFDR-GATKVITCR
LperuS13 RKHGTCCADV-------FN-QS---MYFDISMKLTD--SIDLLKILRTKGIKPGY--TYTGDQISRAIKS-VTQNNPNPKCTYIGR--SLELIEIGICFNR-TTNALMPCP
LperuS12 IKHGSCCTDL-------FN-ET---QYFDLALVLKD--RFDLLTTFRIHGIVPRS--SHTVDKIKKTIRS-VTGVLPNLSCTKN-----MDLLEIGICFNR-EASKMIDCT
LperuS11a KKHGTCCSDL-------FN-QE---KYFDLALILKD--KFDLLTTFRNKGIIPKS--TCTINKIQKTIRT-VTGVVPNLSCTPT-----MELLEVGICFNR-DASKLIDCD
LperuS3 KKHGTCSVDL-------YN-QE---QYFDLAIELKE--KFDLLKTLKNHGITPSK--TNTVIDVEEAIKA-VTKEVPNLNCIGDSSQ-TMELLEIGICFNR-EGTTVIACR
PinflS3 NKHGICSSNL-------YD-QR---AYFLLAMRVKD--KFDLLTTLRTHGITPGT--KHTFGEIQKAIKTVTNNKDPDLKCVEHIKG-VKELKEVGICFTP-AADSFHDCR
PinflS1 RKHGMCCSKI-------YN-QK---AYFLLATRLKE--KFDLLTTLRTHGITPGT--KHTFGDIQKAIKTVTNQVDPDLKCVEHIKG-VQELNEIGICFNP-AADNFYPCH
X2 NKHGTCCSGR-------YN-LQ---QYFHLAMALKD--KFNLLTSLTNHGIIPGS--NYNVQKINSTIKT-ITRGYPNLSCTEE-----MELWEIGICFDS-TVKNVIDCP
Scrophulariaceae S-RNases
AhispS2 KKHGACALPM-------YS-FN---DYFVKALELKK--RNNVLDMLSRKSLTPGD-QRVDVSDVNGAITK-VTGGIAILKCPEG------YLTEVIICFDP-SGFPVIDCP
AhispS4 RKHGSCCFPP-------HE-SE---IYFLKALELKD--RLDVLTILENNNFNPGTPQPFSVLRVFNTISR-AIGKTPILKCAQS------YLKEVVICVDN-NGASVVHCP
AhispS5 EKHGTCALSV-------YT-FD---DYFRETLNMKR--RFNILDMLQRKSMRPGD--RVDPQEVARAISK-VTNHEPEVKCREG------FLTEIIICFDTGRDASVIDCP
C3 C4 C5

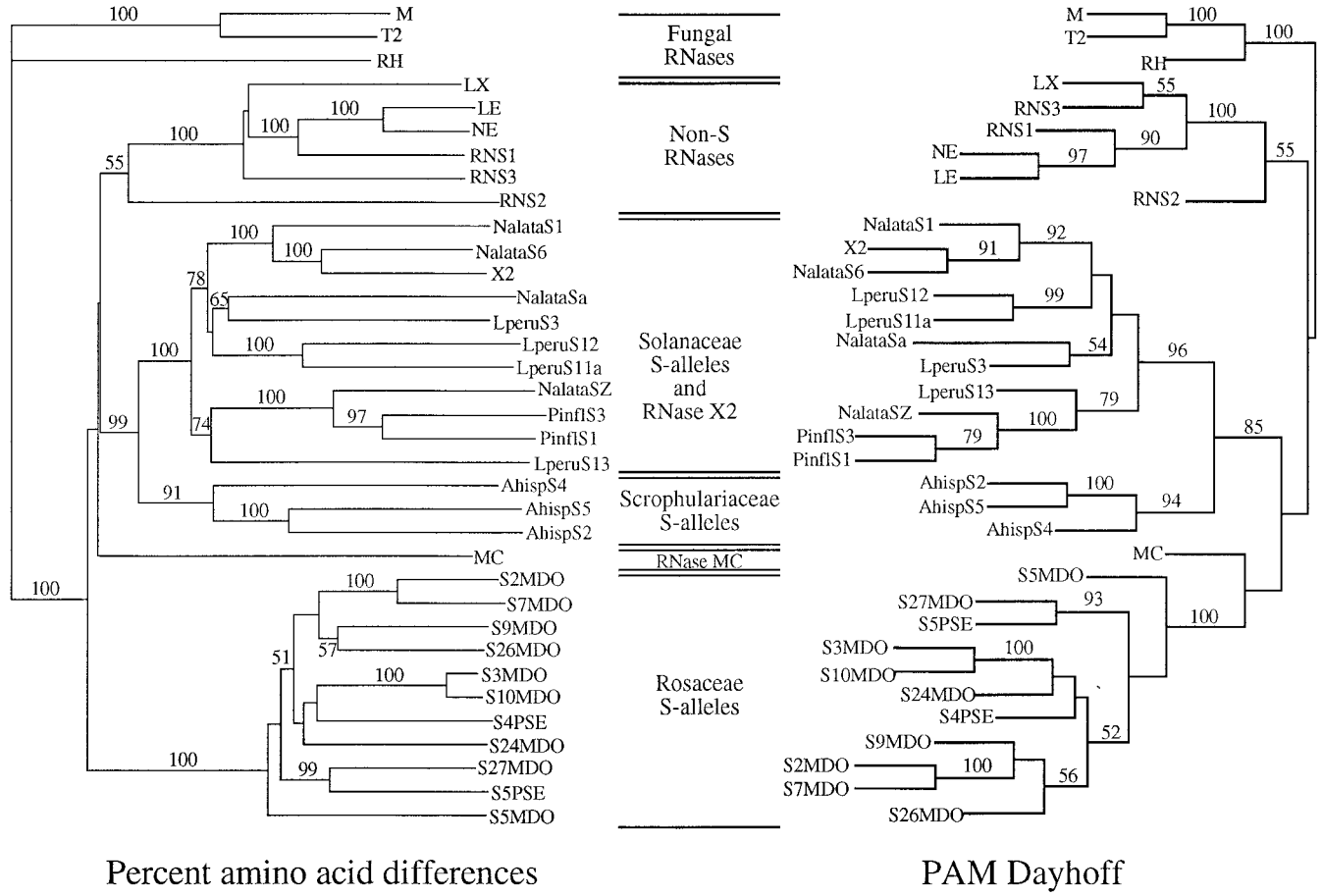


Fig. 2. Neighbor joining genealogies using distances calculated as mean percentage amino acid differences (left side) or using the PAM 001 Dayhoff matrix (right side). Numbers indicate the percentage of 100 bootstrap replicates in which a group was found (numbers <50% not shown).

with the exception of RNase X2 and perhaps MC and RNS2, form a group that is no more divergent than are the S-alleles from individual dicot families. These non-S RNases were drawn from a diverse group of taxa (LE, LX, and NE from Solanaceae, MC from Cucurbitaceae, RNS1-3 from Brassicaceae), which share the same most recent common ancestor as do the Solanaceae and Rosaceae (Chase et al., 1993). The non-S sequences may in fact have diverged well before the most recent common ancestor of the taxa from which they were sampled since there is little reason to assume that they represent sequences from the same locus. Nevertheless, they are far more similar to one another than are the S-alleles from different families.

There are two possible explanations for the sequence similarity of the non-S RNases. First, the known plant non-S RNases could be a biased sample of plant RNases. Several of these loci are known to be upregulated under phosphate limitation or in senescent or floral (soon to senesce) tissue (Ide et al., 1991; Jost et al., 1991; Löffler, Glund, and Irie, 1993; Green, 1994). Stress-related and scavenging RNases may be more closely related to one another than are RNase loci involved in other functions such as defense. It is interesting to note that RNase MC, the most divergent of the group of non-S RNases (except for X2), is found in ungerminated seed tissue and therefore may not be associated with senescence or stress functions (although colyledons senesce fairly soon after germination and MC could function in this regard). If a more diverse array of non-S RNases were known, gene-

←

Fig. 1. Amino acid alignment of 35 RNase protein sequences. The five conserved (C1-C5) and two hypervariable (HVa, HVb) regions identified for S-alleles from the Solanaceae (Ioerger et al., 1991) are underlined. Asterisks mark amino acids that are invariant among angiosperm RNases. Fungal RNases: M, *Aspergillis saitoi* (Watanabe et al., 1990); T2, *Aspergillus oryzae* (Kawata, Sakiyama, and Tamaoki, 1988); Rh, *Rhizopus niveus* (Kurihara et al., 1992). Angiosperm non-S RNases: LX and LE, *Lycopersicon esculentum* (Löffler, Glund, and Irie, 1993); RNS1-3, *Arabidopsis thaliana* (Taylor and Green, 1991); NE, *Nicotiana alata* (GenBank accession number U13256); MC, *Momordica charantia* (Ide et al., 1991); X2, *Petunia inflata* (Lee, Singh, and Kao, 1992). Rosaceae S-alleles: S2, 3, 5, 7, 9, 10, 24, 26, 27MDO, *Malus domestica* (Broothaerts et al., 1995; Janssens et al., 1995; W. Broothaerts, unpublished data); S4, 5PSE, *Pyrus serotina* (Sassa et al., 1996). Solanaceae S-alleles: NalataS1, 6, a, Z, *Nicotiana alata* (Kheyr-Pour et al., 1990); LperuS3, 11a, 12, 13 *Lycopersicon peruvianum* (Chung et al., 1993; Royo, Kowyama, and Clarke, 1994); PinfIS1, 3 *Petunia inflata* (Ai et al., 1990). Scrophulariaceae S-alleles: AhispS2, 4, 5, *Antirrhinum hispanicum* (Xue et al., 1996).

alogical analysis could find different non-S RNases sister to the S-loci of different families, refuting the hypothesis of homology. Even the S-RNases from the Solanaceae and Scrophulariaceae could prove nonhomologous if independent antecedents of the S-locus in each family were detected. Finally, the S-RNases of the Solanaceae and Scrophulariaceae might appear homologous but actually result from parallel conversions of the same RNase to function in incompatibility, a potential problem common to character reconstructions using parsimony.

On the other hand, the current sample of non-S RNases may not represent a biased sample. The fact that most sequences have diverged relatively little may reflect the influence of stabilizing selection on these genes while diversifying selection acts on the S-locus (Clark and Kao, 1991; Richman and Kohn, 1996; Richman, Uyenoyama, and Kohn, 1996). Under the hypothesis that all S-RNases are homologous, the large distance between the S-alleles from each family could represent the effects of diversifying selection and allelic turnover following phyletic divergence. Since the Scrophulariaceae and Solanaceae share a more recent common ancestor with each other than with the Rosaceae, the S-alleles from these families are more similar. Under this hypothesis, complete turnover has occurred since these families diverged, such that alleles from each family are more closely related to one another than to alleles from other families, and no interdigitation of alleles from different families occurs in genealogical reconstructions (Fig. 2; see also Xue et al., 1996).

Two types of additional data could greatly increase the power of this analysis and might significantly alter the conclusions reached. The first would be inclusion of a wider range of plant non-S RNase sequences. Many more RNase loci likely remain to be discovered in plants (Green, 1994). As discussed above, there may exist undiscovered non-S sequences sister to the S-RNases from different families. On the other hand, if newly discovered non-S RNases continue to cluster into a single group of non-S RNases, the inference of homologous origins of all S-RNases may be strengthened.

The second piece of additional data that could help improve the strength of genealogical inference would be the use of less distant outgroup sequences. At present RNase sequences from higher plants are available only from higher dicots. No sequences from nonangiosperm plants or from monocots are available. Use of distant (fungal) outgroup sequences may cause the uncertainty in root placement in the present analyses. Root placement is critical to the interpretation of the number of origins of S-RNases. If an RNase sequence from a group sister to the higher dicots were found that was similar to the group of angiosperm non-S RNases, it is quite possible that a strongly supported rooting making the non-S RNases sister to all S-RNases could be found. Such a topology would support the inference of homology of all S-RNases.

In conclusion, bootstrap analyses using outgroup rooting support the conclusion of Xue et al. (1996) that the use of RNases in the gametophytic self-incompatibility systems of the Scrophulariaceae and Solanaceae results from a single evolutionary origin. The inference of Sassa et al. (1996), that use of S-RNases by the Rosaceae and Solanaceae represents an example of convergent evolution, may be premature. The S-RNases of the Rosaceae may be homologous to those of the Solanaceae and Scrophulariaceae but uncertainties concerning root placement and the genealogical positions of some non-S RNases preclude strong inference on this point. Use of additional non-S RNase sequences from angiosperms and use of less distant outgroup sequences with which to root the genealogy should allow stronger evolutionary inferences to be drawn concerning homology or convergence. If the S-RNases from all three families are homologous, we might expect to find additional examples of the use of RNases in incompatibility among the 80% of dicot families that share the same most recent common ancestor.

LITERATURE CITED

AI, Y., A. SINGH, C. E. COLEMAN, T. R. IOERGER, A. KHEYR-POUR, AND T.-H. KAO. 1990. Self-incompatibility in *Petunia inflata*: isolation and characterization of cDNAs encoding three S-allele-associated proteins. *Sexual Plant Reproduction* 3: 130–138.

ANDERSON, M. A., E. C. CORNISH, S.-L. MAU, E. G. WILLIAMS, R. HOGGART, A. ATKINSON, I. BONIG, B. GREGO, R. SIMPSON, P. J. ROCHE, J. D. HALEY, J. D. PENSCHOW, H. D. NIALL, G. W. TREGEAR, J. P. COGHLAN, R. J. CRAWFORD, AND A. E. CLARKE. 1986. Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. *Nature* 321: 38–44.

BROOTHAERTS, W., G. A. JANSSENS, P. PROOST, AND W. F. BROEKAERT. 1995. cDNA cloning and molecular analysis of two self-incompatibility alleles from apple. *Plant Molecular Biology* 27: 499–511.

CHASE, M.W., ET AL. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. *Annals of the Missouri Botanical Garden* 80: 528–580.

CHUNG, I. K., K., NAKATA, H. TANAKA, T. ITO, H. HORIUCHI, A. OHTA, AND M. TAKAGI. 1993. Identification of cDNA clones coding for the style specific S11a-glycoprotein associated with gametophytic self-incompatibility in Tomato (*Lycopersicon peruvianum*). *Bioscience, Biotechnology, and Biochememistry* 57: 1172–1176.

CLARK, A. G., AND T.-H. KAO. 1991. Excess nonsynonymous substitution at shared polymorphic sites among self-incompatibility alleles of Solanaceae. *Proceedings of the National Academy of Sciences, USA* 88: 9823–9827.

CRONQUIST, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, New York, NY.

DE NETTANCOURT, D. 1977. Incompatibility in angiosperms. Springer-Verlag, Berlin.

FELSENSTEIN, J. 1993. PHYLIP (Phylogeny Inference Package). Department of Genetics, University of Washington, Seattle, distributed by the author.

FOOTE, H., J. P. RIDE, V. E. FRANKLIN-TONG, E. A. WALKER, M. J. LAWRENCE, AND F. C. FRANKLIN. 1994. Cloning and expression of a novel class of self-incompatibility (S-) gene from *Papaver rhoeas* L. *Proceedings of the National Academy of Sciences, USA* 91: 2265–2269.

GREEN, P. 1994. The ribonucleases of higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 45: 421–445.

HUANG, S., H. S. LEE, B. KARUNANANDAA, AND T.-H. KAO. 1994. Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. *Plant Cell* 6: 1021–1028.

IDE, H., M. KIMURA, M. ARAI, AND G. FUNATSU. 1991. The complete amino acid sequence of ribonuclease from the seeds of bitter gourd (*Momordica charantia*). *FEBS Letters* 284: 161–164.

IOERGER, 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.

JANSSENS, G. A., I. J. GODERIS, W. F. BROEKAERT, AND W. BROOTHAERTS. 1995. A molecular method for S-allele identification in apple based on allele-specific PCR. *Theoretical and Applied Genetics* 91: 691–698.

JOST, W., H. BAK, K. GLUND, P. TERPSTRA, AND J. J. BEINTEME. 1991.

Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells. *European Journal of Biochemistry* 198: 1–6.

KAWATA, Y., F. SAKIYAMA, AND H. TAMAOKI. 1988. Amino-acid sequence of ribonuclease T_2 from *Aspergillus oryzae*. *European Journal of Biochemistry* 176: 683–697.

KHEYR-POUR, A., S. B. BINTRIM, T. R. IOERGER, R. REMY, S. A. HAMMOND, AND T.-H. KAO. 1990. Sequence diversity of pistil S-proteins associated with gametophytic self-incompatibility in *Nicotiana alata*. *Sexual Plant Reproduction* 3: 88–97.

KURIHARA, H., Y. MITSUI, K. OHGI, M. IRIE, H. MIZUNO, AND K. T. NAKAMURA. 1992. Crystal and molecular structure of RNase Rh, a new class of microbial ribonuclease from *Rhizopus niveus*. *FEBS Letters* 306: 189–192.

LEE, H. S., S. S. HUANG, AND T.-H. KAO. 1994. S-proteins control rejection of self-incompatible pollen in *Petunia inflata*. *Nature* 367: 560–563.

———, A. SINGH, AND T.-H. KAO. 1992. RNase X2, a pistil-specific ribonuclease from *Petunia inflata,* shares sequence similarity with solanaceous S proteins. *Plant Molecular Biology* 20: 1131–1141.

LI, X., J. NIELD, D. HAYMAN, AND P. LANGRIDGE. 1994. Cloning of a putative self-incompatibility gene from the pollen of the grass *Phalaris coerulescens*. *Plant Cell* 6: 1923–1932.

LÖFFLER, A., K. GLUND, AND M. IRIE. 1993. Amino acid sequence of an intracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells. *European Journal of Biochemistry* 214: 627–633.

MCCLURE, B. A., J. E. GRAY, M. A. ANDERSON, AND A. E. CLARKE. 1990. Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. *Nature* 347: 757–760.

NASRALLAH, J. B., T.-H. KAO, M. L. GOLDBERG, AND M. E. NASRALLAH. 1985. A cDNA clone encoding an S-locus-specific glycoprotein from *Brassica oleracea*. *Nature* 318: 263–267.

RICHMAN, A. D., AND J. R. KOHN. 1996. Learning from rejection: the evolutionary biology of single-locus self incompatibility. Trends in Ecology and Evolution 11: 497–502.

RICHMAN, A. D., M. K. UYENOYAMA, AND J. R. KOHN. 1996. Contrasting patterns of allelic diversity and gene genealogy at the self-incompatibility locus in two species of Solanaceae. *Science* 273: 1212–1216.

ROYO, J., Y. KOWYAMA, AND A. E. CLARKE. 1994. Cloning and nucleotide sequence of two S-RNases from *Lycopersicon peruvianum* (L.) Mill. *Plant Physiology* 105: 751–752.

SASSA, H., H. HIRANO, AND H. IKEHASHI. 1992. Self-incompatibility-related RNases in styles of Japanese pear (*Pyrus serotina* Rehd.). *Plant and Cell Physiology* 33: 811–814.

———, T. NISHIO, Y. KOWYAMA, H. HIRANO, T. KOBA, AND H. IKEHASHI. 1996. Self-incompatibility (S) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease subfamily. *Molecular and General Genetics* 250: 547–557.

TAYLOR, C. B., AND P. J. GREEN. 1991. Genes with homology to fungal and S-gene RNases are expressed in *Arabidopsis thaliana*. *Plant Physiolology* 96: 980–984.

UYENOYAMA, M. K. 1995. A generalized least squares estimate for the origin of sporophytic self-incompatibility. *Genetics* 139: 975–992.

WATANABE, H., A. NAITOH, Y. SUYAMA, N. INOKUCHI, H. SHIMADA, T. KOYAMA, K. OHGI, AND M. IRIE. 1990. Primary structure of a base non-specific and adenylic acid preferential ribonuclease from *Aspergillus saitoi*. *Journal of Biochemistry* 108: 303–310.

WELLER, S. G., M. J. DONOGHUE, AND D. CHARLESWORTH. 1995. The evolution of self-incompatibility in flowering plants: a phylogenetic approach. *In* P. C. Hoch and A. G. Stephenson [eds.], Experimental and molecular approaches to plant biosystematics, 355–382. Missouri Botanical Garden, St. Louis, MO.

WHITEHOUSE, H. L. K. 1951. Multiple-allelomorph incompatibility of pollen and style in the evolution of the angiosperms. *Annals of Botany* 14: 198–216.

XUE, Y., R. CARPENTER, H. G. DICKINSON, AND E. S. COEN. 1996. Origin of allelic diversity in *Antirrhinum* S locus RNases. *Plant Cell* 8: 805–814.