Plant Molecular Biology **28**: 847–858, 1995.

The *S*-locus of *Nicotiana alata*: genomic organization and sequence analysis of two *S*-RNase alleles

Daniel P. Matton[1], Shiao-Lim Mau, Shigehisa Okamoto[2], Adrienne E. Clarke* and Ed Newbigin
*Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville, Victoria 3052, Australia (*author for correspondence); Present address: [1] Institut de Recherche en Biologie Végétale de Montréal, Université de Montréal, 4101 rue Sherbrooke est, Montréal, H1X 2B2, Canada; [2] Biological Laboratory, College of Liberal Arts, Kagoshima University, 1-21-30 Kourimoto, Kagoshima, Japan*

Received 25 October 1994; accepted in revised form 2 May 1995

Key words: *Nicotiana alata*, repetitive DNA, self-incompatibility, *S*-RNase gene

Abstract

Genomic clones encoding the S_2- and S_6-RNases of *Nicotiana alata* Link and Otto, which are the allelic stylar products of the self-incompatibility (*S*) locus, were isolated and sequenced. Analysis of genomic DNA by pulsed-field gel electrophoresis and Southern blotting indicates the presence of only a single *S*-RNase gene in the *N. alata* genome. The sequences of the open-reading frames in the genomic and corresponding cDNA clones were identical. The organization of the genes was similar to that of other S-RNase genes from solanaceous plants. No sequence similarity was found between the DNA flanking the S_2- and S_6-RNase genes, despite extensive similarities between the coding regions. The DNA flanking the S_6-RNase gene contained sequences that were moderately abundant in the genome. These repeat sequences are also present in other members of the Nicotianae.

Introduction

Self-incompatability is a mechanism used by some species of flowering plants to enforce outbreeding. In solanaceous plants such as *Nicotiana alata*, self-incompatibility is controlled by a single genetic locus (*S*-locus) with many alleles. The only known product of this locus in solanaceous plants is an extracellular glycoprotein with ribonuclease activity (*S*-RNase) that is expressed at high levels within the mature style [1, 23] and at much lower levels within developing pollen grains [7, 11]. Expression of *S*-RNases within the style is necessary for self-incompatibility [20, 26], and the ribonuclease activity of the protein is required for rejection of incompatible pollen [15, 19]. A current model for the action of *S*-RNases in self-incompatibility is based on their entry into incompatible pollen tubes during growth through the style. This results in degradation of RNA including ribosomal (r) RNA [13, 24]. As mature pollen grains and pollen tubes do not appear to

The DNA sequences reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers U08860 (S_2-RNase gene) and U08861 (S_6-RNase gene).

express rRNA genes, the loss of the rRNA made during pollen development would effectively destroy the ability of the pollen tube to synthesize protein and lead to arrest of cell growth [22].

What is not made clear by this model is the nature of the recognition event that occurs between the components of the *S*-locus expressed in pollen and the style, although the possibility that both are *S*-RNases has been raised [11]. However, mutational studies on self-incompatible plants suggest that the *S*-locus is made up of two distinct parts, one controlling the self-incompatibility phenotype of pollen and the other part, the phenotype of the style [21]. As part of a study to understand the structure of the *S*-locus in solanaceous plants, we isolated and sequenced genomic clones for the S_2- and S_6-RNases of *N. alata*. The structure of the genes was similar to that of genes for *S*-RNases cloned from other solanaceous plants and, like these genes, the flanking regions of the S_6-RNase gene were rich in repeated sequences [8, 16, 30].

Materials and methods

Plant materials

Nicotiana alata plants (genotypes S_2S_3, S_2S_6, S_3S_3 and S_6S_6) used in this study were maintained as described by Anderson *et al.* [2]. Other *Nicotiana* plants were obtained from the following sources: School of Botany, University of Melbourne, Parkville, Vic., Australia (*N. forgetiana*, *N. tabacum*); Australian Tropical Field Crops Genetic Resource Centre, Biloela, Qld, Australia (*N. longiflora*, *N. sylvestris*); CSIRO Division of Plant Industry, Canberra, A.C.T., Australia (*N. benthamiana*); the late Dr K.K. Pandey, DSIR Division of Grasslands Research, Palmerston North, New Zealand (*N. glauca*); Dr P. Heslop-Harrison, John Innes Institute, Norwich, UK (*N. tomentosiformis*). All plants were grown and propagated in a manner similar to that described for *N. alata*.

Construction of genomic DNA libraries

Genomic DNA was isolated from the leaves of *N. alata* plants as described in Bernatzky and Tanksley [4]. A genomic library from plants of the S_2S_3 genotype was prepared by digesting 200 μg of leaf DNA to completion with *Eco* RI (Promega, Madison, WI). The digested DNA was fractioned on a linear sucrose gradient (10–40%) by centrifugation for 17 h at 100 000 × *g* in a SW41 rotor (Beckman, Somerset, NJ) and a fraction containing DNA of 2–5 kb was ligated into the *Eco* RI site of λgt10 (Promega). The ligated DNA was packaged into phage particles (Promega) and 5×10^6 plaques from the resulting genomic library were hybridized with a ^{32}P-labelled cDNA clone for S_2-RNase [1]. Four positive plaques were purified and characterized by restriction mapping. The *Eco* RI insert from one of the genomic clones was ligated into pGEM-3 (Promega). The resulting plasmid was named pGS2.

A genomic library from plants of the S_6S_6 genotype was prepared by partially digesting 300 μg of leaf DNA with *Sau* 3A (Promega). Digested DNA was fractionated on a linear sucrose gradient as described above and fractions containing DNA with a molecular weight of 10–23 kb were combined and ligated into the *Bam* HI site of Lambda dash (Stratagene, La Jolla, CA). The ligated DNA was packaged into phage particles and 2.5×10^5 plaques from the resulting library were hybridized to a labelled cDNA probe for S_6-RNase [2]. Six positive plaques were purified and characterized by restriction mapping. One clone (pGS6) contained an insert of ca. 15 kb and a 5.1 kb *Hin*d III fragment from this clone was ligated into pBluescript IIKS (Stratagene) to form pGS6-4 (see Fig. 1).

DNA sequence analysis

The sequences of pGS2 and pGS6-4 were determined by the dideoxy chain termination method [32]. For this purpose, the *Eco* RI and *Hin*d III inserts of pGS2 and pGS6-4, respectively, were digested with restriction enzymes and the sub-

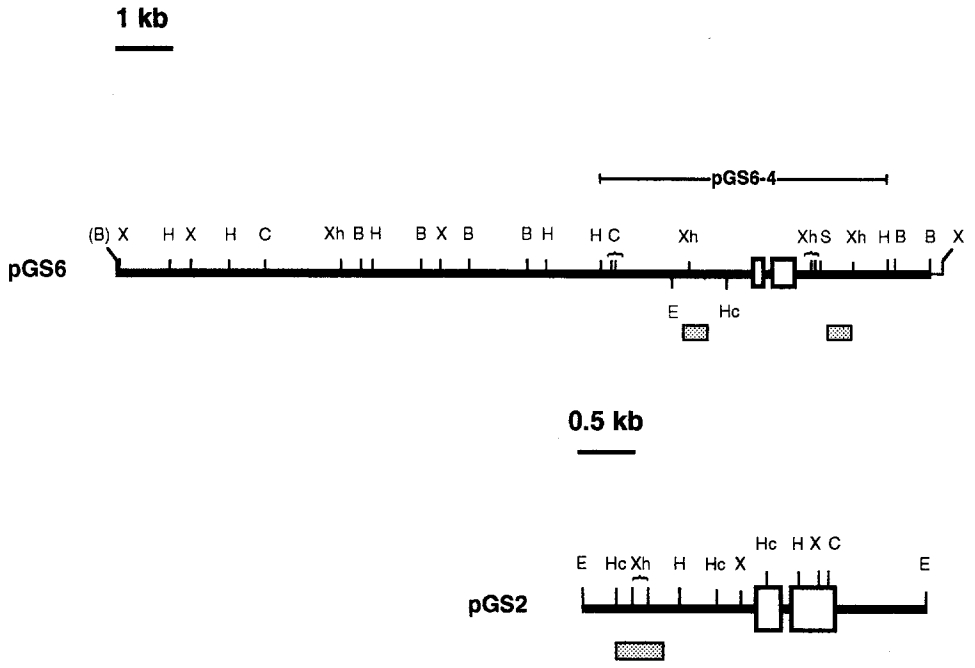


Fig. 1. Restriction maps of genomic clones for two *S*-RNase alleles from *N. alata*. pGS6 contains the S_6-RNase allele and pGS2 contains the S_2-RNase allele. pGS6-4 is a *Hin*d III subclone of pGS6. Open boxes represent the position of exons. Dotted boxes represent short stretches of DNA described in the text that have ca. 70% similarity within pGS6-4, and between pGS6-4 and pGS2. Abbreviations for restriction enzymes are: B, *Bam* HI; C, *Cla* I; E, *Eco* RI; H, *Hin*d III; Hc, *Hinc* II; S, *Sac* II; X, *Xba* I; Xh, *Xho* I. Abbreviations shown in parentheses indicate sites lost during cloning of genomic DNA. All enzyme sites shown above the line are given in full, only those sites relevant to experiments presented in Fig. 4 are indicated below the line.

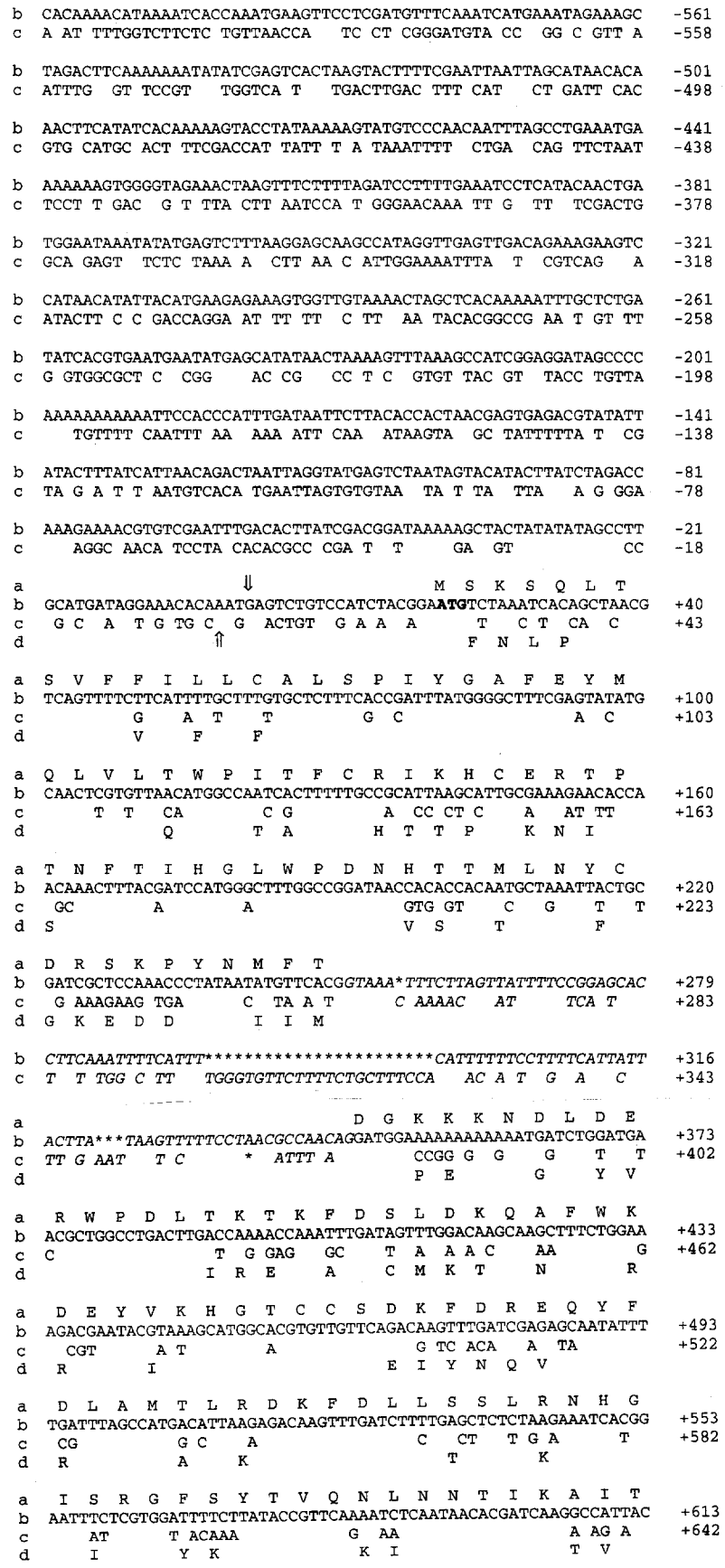
fragments ligated into pBluescript IIKS. Further sequence data was obtained from a series of exonuclease deletions of pGS2 and pGS6-4 made using the Erase-a-base system (Promega).

Primer extension

Total RNA was prepared from mature styles by the method of Chomczynski and Sacchi [6]. The synthetic primers SLM1 (5′-TGTTAACACGAGTTGCATATACTCGAAAGC-3′) and SLM2 (5′-AAATATGACGAAAACTGACGTGAGTGGTAA-3′) were used to determine the start of transcription of the S_2-RNase and S_6-RNase gene, respectively. SLM1 is complementary to the sequence of the S_2-RNase gene between nucleotides +86 and +115, and SLM2 is complementary to the sequence of the S_6-RNase gene between nucleotides +32 and +61 (see Fig. 2). A primer extension reaction using labelled primers (SLM1 or SLM2) and total RNA (10 μg) from style was done as described [3] except that the primers were annealed to RNA overnight at 50 °C in the presence of 50% formamide.

Analysis of DNA flanking the S_6-RNase gene

Total cellular DNA was isolated from *N. alata* leaves (genotype S_6S_6) as described by Saghai-Maroof *et al.* [31] and 10 μg of this DNA was digested with the restriction enzymes, *Hin*d III or *Eco* RI (Promega), according to the manufacturers' recommended protocol. The DNA was fractioned on a 0.8% (w/v) agarose gel run in 1× TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA) and transferred onto a Hybond-N membrane (Amersham, Buckinghamshire, UK) as described [3]. Specific DNA fragments, produced by digesting the genomic clone pGS6 with restriction enzymes (see Fig. 4), were purified from aga-


b CACAAAACATAAAATCACCAAATGAAGTTCCTCGATGTTTCAAATCATGAAATAGAAAGC -561
c A AT TTTGGTCTTCTC TGTTAACCA TC CT CGGGATGTA CC GG C GTT A -558
b TAGACTTCAAAAAAATATATCGAGTCACTAAGTACTTTTCGAATTAATTAGCATAACACA -501
c ATTTG GT TCCGT TGGTCA T TGACTTGAC TTT CAT CT GATT CAC -498
b AACTTCATATCACAAAAAGTACCTATAAAAAGTATGTCCCAACAATTTAGCCTGAAATGA -441
c GTG CATGC ACT TTCGACCAT TATT T A TAAATTTT CTGA CAG TTCTAAT -438
b AAAAAAGTGGGGTAGAAACTAAGTTTCTTTTAGATCCTTTTGAAATCCTCATACAACTGA -381
c TCCT T GAC G T TTA CTT AATCCA T GGGAACAAA TT G TT TCGACTG -378
b TGGAATAAATATATGAGTCTTTAAGGAGCAAGCCATAGGTTGAGTTGACAGAAAGAAGTC -321
c GCA GAGT TCTC TAAA A CTT AA C ATTGGAAAATTTA T CGTCAG A -318
b CATAACATATTACATGAAGAGAAAGTGGTTGTAAAACTAGCTCACAAAAATTTGCTCTGA -261
c ATACTT C C GACCAGGA AT TT TT C TT AA TACACGGCCG AA T GT TT -258
b TATCACGTGAATGAATATGAGCATATAACTAAAAGTTTAAAGCCATCGGAGGATAGCCCC -201
c G GTGGCGCT C CGG AC CG CC T C GTGT TAC GT TACC TGTTA -198
b AAAAAAAAAAATTCCACCCATTTGATAATTCTTACACCACTAACGAGTGAGACGTATATT -141
c TGTTTT CAATTT AA AAA ATT CAA ATAAGTA GC TATTTTTA T CG -138
b ATACTTTATCATTAACAGACTAATTAGGTATGAGTCTAATAGTACATACTTATCTAGACC -81
c TA G A T T AATGTCACA TGAATTAGTGTGTAA TA T TA TTA A G GGA -78
b AAAGAAAACGTGTCGAATTTGACACTTATCGACGGATAAAAAGCTACTATATATAGCCTT -21
c AGGC AACA TCCTA CACACGCC CGA T T GA GT CC -18
a ⇓ M S K S Q L T
b GCATGATAGGAAACACAAATGAGTCTGTCCATCTACGGA**ATG**TCTAAATCACAGCTAACG +40
c G C A T G TG C G ACTGT G A A A T C T CA C +43
d ⇑ F N L P
a S V F F I L L C A L S P I Y G A F E Y M
b TCAGTTTTCTTCATTTTGCTTTGTGCTCTTTCACCGATTTATGGGGCTTTCGAGTATATG +100
c G A T T G C A C +103
d V F F
a Q L V L T W P I T F C R I K H C E R T P
b CAACTCGTGTTAACATGGCCAATCACTTTTTGCCGCATTAAGCATTGCGAAAGAACACCA +160
c T T CA C G A CC CT C A AT TT +163
d Q T A H T T P K N I
a T N F T I H G L W P D N H T T M L N Y C
b ACAAACTTTACGATCCATGGGCTTTGGCCGGATAACCACACCACAATGCTAAATTACTGC +220
c GC A A GTG GT C G T T +223
d S V S T F
a D R S K P Y N M F T
b GATCGCTCCAAACCCTATAATATGTTCAC*GGTAAA**TTTCTTAGTTATTTTCCGGAGCAC* +279
c G AAAGAAG TGA C TA A T *C AAAAC AT TCA T* +283
d G K E D D I I M
b *CTTCAAATTTTCATTT**********CATTTTTTCCTTTTCATTATT* +316
c *T T TGG C TT TGGGTGTTCTTTTCTGCTTTCCA AC A T G A C* +343
a D G K K K N D L D E
b *ACTTA***TAAGTTTTTCCTAACGCCAACAG*GATGGAAAAAAAAAAAAATGATCTGGATGA +373
c *TT G AAT T C * ATTT A* CCGG G G G T T +402
d P E G Y V
a R W P D L T K T K F D S L D K Q A F W K
b ACGCTGGCCTGACTTGACCAAAACCAAATTTGATAGTTTGGACAAGCAAGCTTTCTGGAA +433
c C T G GAG GC T A A A C AA G +462
d I R E A C M K T N R
a D E Y V K H G T C C S D K F D R E Q Y F
b AGACGAATACGTAAAGCATGGCACGTGTTGTTCAGACAAGTTTGATCGAGAGCAATATTT +493
c CGT A T A G TC ACA A TA +522
d R I E I Y N Q V
a D L A M T L R D K F D L L S S L R N H G
b TGATTTAGCCATGACATTAAGAGACAAGTTTGATCTTTTGAGCTCTCTAAGAAATCACGG +553
c CG G C A C CT T G A T +582
d R A K T K
a I S R G F S Y T V Q N L N N T I K A I T
b AATTTCTCGTGGATTTTCTTATACCGTTCAAAATCTCAATAACACGATCAAGGCCATTAC +613
c AT T ACAAA G AA A AG A +642
d I Y K K I T V

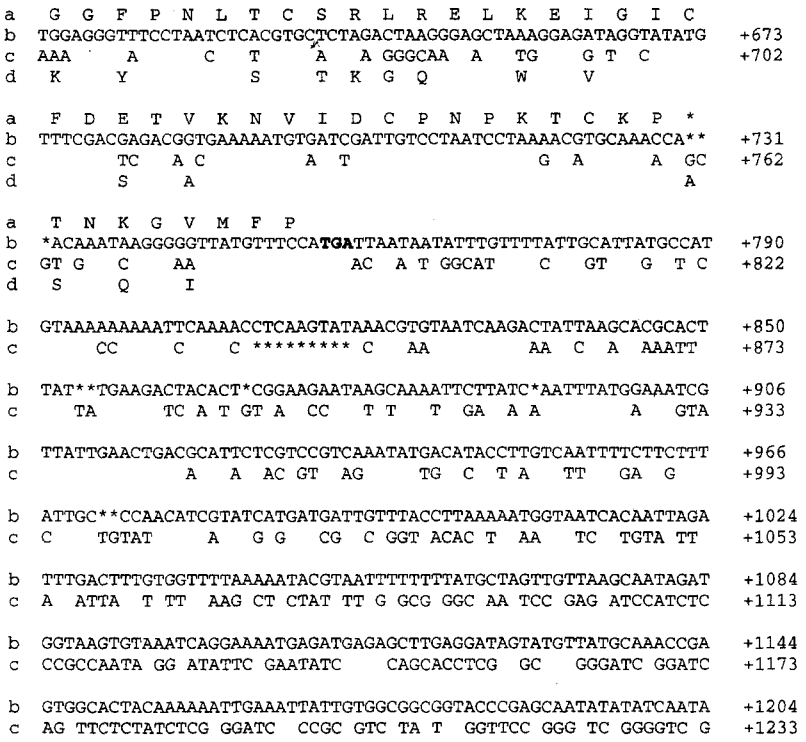


Fig. 2. Amino acid and nucleotide sequence of genomic clones for two S-RNase alleles from *Nicotiana alata*. Line a shows the deduced amino acid sequence and line b the DNA sequence of a part of the genomic clone for S_2-RNase. Line c shows the deduced amino acid sequence and line d the DNA sequence of a part of the genomic clone for S_6-RNase. For the S_6-RNase sequences (lines c and d), only those amino acids or nucleotides that differ from those in the S_2-RNase sequence (lines a and b) are shown. Gaps (indicated by *) have been introduced into the sequences to increase the level of identity. The initiator methionine (ATG) and terminator codons are shown in bold type and the intron sequences are italicized. The start site of transcription (determined by primer extension) for each allele is indicated by an arrow, and numbering of each sequence is relative to this nucleotide which is counted as + 1.

rose gels using SpinBind cartridges (FMC, Rockland, ME) and radiolabelled with random primers (Prime-a-gene, Promega). Hybridization of the radiolabelled pGS6 fragments to *N. alata* genomic DNA was done in 6 × SSC, 5 × Denhart's solution, 0.5% SDS and 125 μg/ml of denatured herring sperm DNA at 65 °C for 16 h. After hybridization, the membranes were washed twice (30 min each time) at 60 °C in 0.1 × SSC, 0.1% SDS and exposed to film.

High-molecular-weight DNA preparation

Nuclear DNA was prepared from leaves of *N. alata* plants (genotype S_6S_6) essentially as described by Creusot *et al.* [9]. Briefly, young leaves (25 g) were deveined, frozen in liquid nitrogen, and ground to a powder in a mortar and pestle. The frozen powder was allowed to thaw at room temperature for 10 min and 30–50 ml of extraction buffer (10 mM citrate buffer pH 6, 5 mM EDTA, 0.6 M mannitol, 0.1% w/v bovine serum albumin (fraction V, Sigma, St Louis) 0.1% (w/v) Triton X-100) was added. The mixture was homogenized gently, allowed to thaw at room temperature for 15 min, and the volume of buffer was adjusted to 150 ml. The homogenate was filtered through a single layer of sterile Miracloth and centrifuged at 600 × *g* for 10 min at 4 °C. The crude nuclear pellet was washed with extraction buffer, recentrifuged, and resuspended in 5 ml of extraction buffer. This suspension of nuclei was warmed to 45 °C for 1 min, mixed with ca. 4 ml

of 2% low-melting agarose in 0.125 M EDTA (45 °C), and cast into a mold (BioRad, Hercules, CA). The plugs were left for 15 min at 4 °C to solidify, washed for 48 h in three changes of ESP buffer (0.5 M EDTA pH 8, 1% sodium *N*-lauroyl sarcosinate, 1 mg/ml proteinase K (Boehringer-Mannheim)) at 50 °C, and subsequently washed for 24 h in three changes of TE_{50} buffer (10 mM Tris-HCl pH 7.5, 50 mM EDTA) at 50 °C. The plugs were stored at 4 °C in TE_{50} until used.

Pulsed-field gel electrophoresis

Agarose plugs containing high-molecular-weight DNA were digested overnight with restriction enzymes (40–60 units/plug) as described [9] except that spermidine was omitted from the preincubation buffer. Digestions were stopped by adding EDTA to a final concentration of 0.05 M. Prior to electrophoresis, the plugs were rinsed twice (30 min each time) in TE_{50} and equilibrated in 1× TBE (0.09 M Tris-borate, 0.002 M EDTA) for 15 min. The products of DNA digestion were resolved by pulsed-field gel electrophoresis at 14 °C using a BioRad CHEF DR II apparatus. To resolve DNA molecules of between ca. 23 and 500 kb, 1% (w/v) agarose gels (SeaKem LE, FMC) were run in 0.5× TBE at 200 V with the switching interval increasing from 10 s to 50 s over 24 h. For DNA molecules larger than 500 kb, a different agarose was used (FastLane, FMC) and the switching time was increased from 60 s to 110 s over 24 h. Lambda concatamers and 5 kb DNA ladder were obtained from BioRad. DNA was transferred to nylon membranes, hybridized to radiolabelled S_6-RNase cDNA, and washed as described above except that the temperature of the washes was 55 °C.

Results

Isolation and sequencing of genes encoding the S_2- *and* S_6-*RNases of* Nicotiana alata

The S_2-RNase cDNA was used as a probe to detect hybridizing sequences in the genomic DNA of *N. alata* genotypes S_2S_2, S_2S_3 and S_3S_3. When hybridized and washed under stringent conditions, the probe detected a single 3.1 kb *Eco* RI fragment in DNA from S_2S_2 and S_2S_3 genotypes, but no hybridization was detected in DNA from S_3S_3 plants (data not shown). A 3.1 kb *Eco* RI fragment that hybridized to S_2-RNase cDNA was cloned from DNA of *N. alata* plants (genotype S_2S_3) as described in the Materials and methods and a clone containing the *Eco* RI insert (pGS2) was sequenced on both strands. Similarly, the S_6-RNase cDNA probe hybridized strongly to a single 5.1 kb *Hin*d III fragment in plants carrying the S_6-allele [2]. A large genomic fragment (pGS6) that hybridized to S_6-RNase cDNA was cloned from *N. alata* plants (genotype S_6S_6), and a 5.1 kb *Hin*d III fragment (pGS6-4) from this genomic clone was completely sequenced on both strands. Alignment of the sequences of pGS2 and S_2-RNase cDNA, or pGS6-4 and S_6-RNase cDNA, showed that both genomic isolates contain the entire sequence of the corresponding cDNA clone. The restriction maps of pGS2 and pGS6 (the genomic clone from which pGS6-4 was derived) are shown in Fig. 1. No common restriction sites were found. Comparison of part of the sequences of pGS6-4 and pGS2 is shown in Fig. 2. The two DNA sequences were aligned for maximal similarity by the placement of gaps in either sequence.

The S_2- and S_6-RNase genes contain a single intron at the same position with respect to the amino acid sequence. The introns are small and neither the precise size (94 bp and 120 bp in the S_2- and S_6-RNase genes, respectively) nor DNA sequence is conserved except around the splice junctions where they conform to the consensus sequence described for other plant genes [14]. Overall, the sequence similarity between the two exons of the S_2- and S_6-RNase genes is about 76% and between the introns is ca. 40%.

In marked contrast to the high degree of similarity within the exons, the 5′-flanking regions of the two *S*-RNase genes have no long stretches of related sequence. However, there are small areas of close identity, particularly from about −70 to the start site of translation, a region that includes

a putative 'TATA' box (located at ca. −30) and the start site of transcription of both genes (see Fig. 2). The sequence conservation is more extensive in the region of the *S*-RNase genes that follows the termination codon (TGA) and covers a stretch of about 220 nucleotides that contains several putative poly(A) addition signals and the known 3′ end points for S_2-RNase transcripts [11, 25]. A stretch of 400 bp with ca. 70% identity was found about 1 kb upstream of the open-reading frame in both pGS6-4 and pGS2 (see Fig. 1) and was also present 600 bp downstream of the open-reading frame in pGS6-4. Whether this sequence is in a comparable position downstream of the S_2-RNase is unknown as the clone pGS2 has only a short stretch of DNA 3′ of the open-reading frame. This sequence is not flanked by direct repeats and has no similarity to the known sequences of either transposons or retrotransposons (data not shown).

N. alata *has a single copy of the* S-*RNase gene*

High-molecular-weight DNA, isolated from *N. alata* leaves (S_6S_6 genotype) was embedded in agarose, cut with restriction enzymes, separated on pulsed-field gels, blotted onto nylon membranes and hybridized to labelled S_6-RNase cDNA. Figure 3 shows an example of these results and Table 1 summarizes the sizes of DNA fragments detected for a number of restriction enzymes. A single hybridizing fragment was detected by the S_6-RNase cDNA with most restriction enzymes, although in some cases, the probe recognised multiple bands with similar intensities. In two instances (*Bcl* I and *Bgl* II), the two hybridizing fragments resulted from cleavage of the gene at sites within the second exon. In one case (*Sal* I), three fragments were detected that could have arisen as a result of partial cleavage caused by methylation at CpG dinucleotides. *Sal* I is a methylation-sensitive enzyme that contains CpG in its 6 bp recognition site.

The flanking regions of the S_6-RNase gene contain repetitive sequences. The report that the flanking regions of an *S*-RNase gene from *Petunia inflata* contain repetitive sequences [8] prompted us to examine the sequences surrounding the S_6-RNase gene of *N. alata*. This was done by probing DNA blots of *N. alata* genomic DNA (genotype S_6S_6) with a series of restriction fragments obtained from the genomic clone, pGS6

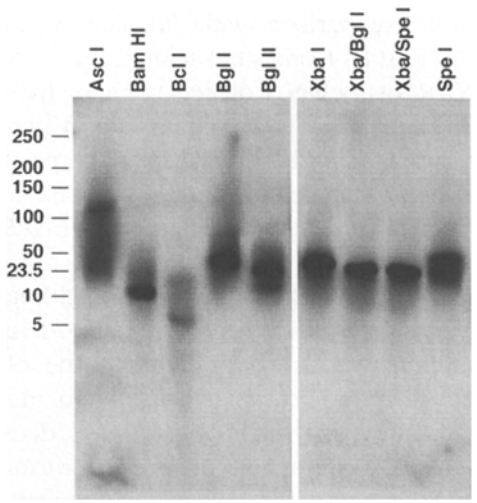


Fig. 3. Pulsed-field gel blots of DNA from *Nicotiana alata* leaves (genotype S_6S_6) hybridized to the S_6-RNase cDNA. DNA from *N. alata* (genotype S_6S_6) was digested with the enzymes indicated, fractionated by pulsed-field gel electrophoresis, transferred to a nylon membrane and probed with labelled S_6-RNase cDNA. Lambda and 5 kb DNA concatamers (BioRad) served as size standards.

Table 1. Summary of fragment sizes detected with the S_6-RNase cDNA probe after digestion of *N. alata* DNA with 10 different restriction enzymes.

Enzyme	Approximate size (kb)
Asc I	100
Bcl I	13.4
Bgl I	40
Bgl II	20, 10
Eag I	770
Nar I	200
Sal I	580, 370, 290
Sfi I	100
Spe I	40
Xba I	35
Spe I/*Bgl* I	30
Spe I/*Xba* I	25
Xba I/*Bgl* I	25

and, for comparison with labelled S_6-RNase cDNA (Fig. 4). Consistent with previous results [2], S_6-RNase cDNA detects a major hybridizing fragment of 5.1 kb and 6.4 kb in genomic DNA cut with *Hin*d III and *Eco* RI, respectively. Fragment 9, which contains the S_6-RNase coding region, hybridized predominantly to *Eco* RI or *Hin*d III fragments of a size similar to those detected using the S_6-RNase cDNA, although a number of other fragments were also detected, some of which also hybridized to the cDNA probe. All other probes hybridized to multiple DNA fragments, with no two probes detecting the same set of genomic sequences. For example, probes 2, 3, 4 and 10 gave hybridization patterns consisting of a smear with few distinct bands. The patterns are however, distinguishable, indicating that each probe contains a different repeat that is dispersed throughout the genome. In contrast, probes 1, 5, 6, 7, 8 and 11 hybridized strongly to a limited; although large, number of DNA fragments, indicating that different members of the family of elements detected by each probe share the same sequence organization. Coleman and Kao [8] observed a gradient of repetitive DNA in the DNA immediately flanking the S_1-RNase gene from *P. inflata* with the more abundant elements further from the coding region. We observed a similar gradient surrounding the S_6-RNase gene of *N. alata*. This is reflected in the length of time the filters hybridized to different probes were exposed to film; filters hybridized with probes 1–3 and probe 4 were exposed to film at room temperature for 4.5 h and 15 h, respectively, whereas filters hybridized with probes 5–11 were exposed to film for 2–3 days at −70 °C.

N. alata *repetitive elements are also found in other* Nicotiana *spp.*

We surveyed species from the three sub-genera of *Nicotiana* for the presence of sequences hybridizing to repetitive DNA surrounding the S_6-RNase

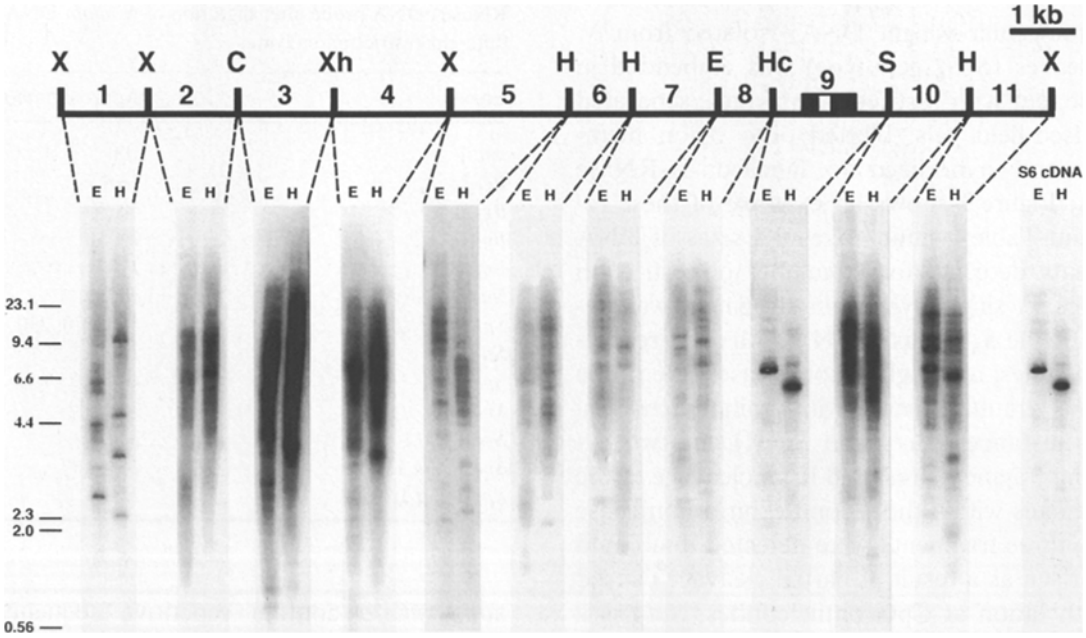


Fig. 4. DNA gel blot analysis of genomic DNA from *N. alata* probed with fragments from pGS6. Genomic DNA (10 μg) was digested with either *Eco* RI (E) or *Hin*d III (H), separated on agarose gels and transferred to nylon membranes. Membranes were probed with radiolabelled DNA fragments obtained from pGS6 (upper line), or S_6-RNase cDNA (S_6 cDNA). Each probe derived from pGS6 is numbered. Other features of pGS6 and the abbreviations for restriction enzymes are described in the legend to Fig. 1. Numbers on the left indicate molecular weight (in kb).

gene. The plants surveyed included four South American species, *N. alata*, *N. forgetiana*, *N. longiflora* and *N. sylvestris* and one Australian species, *N. benthamiana* from the sub-genus *Petuniodes*; two species, *N. tabacum* and *N. tomentosiformis*, from the sub-genus *Tabacum* and a single species, *N. glauca*, from the sub-genus *Rustica*. Probe 5 was used in Southern hybridizations of genomic DNA digested with *Eco* RI from each of these species (Fig. 5). Probe 5 hybridized strongly to DNA isolated from all species except *N. tomentosiformis* which hybridized only weakly. The hybridization was most intense to DNA from *N. forgetiana*, *N. glauca*, *N. sylvestris* and *N. tabacum*. Interestingly, a similar pattern of hybridization was detected in both *N. sylvestris* and *N. tabacum*, consistent with the known contribution of *N. sylvestris* to the genome of *N. tabacum* [17]. Comparable results were obtained with other probes derived from the genomic clone, pGS6 (data not shown).

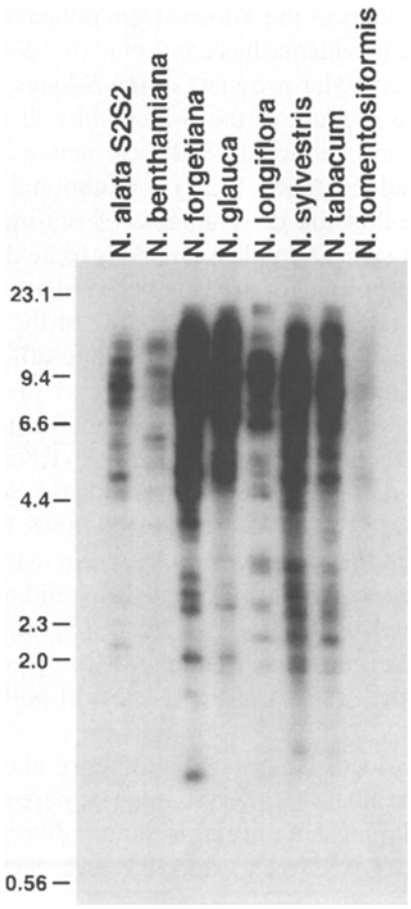


Fig. 5. DNA gel blot analysis of genomic DNA prepared from different *Nicotiana* species probed with fragment 5 (see Fig. 4). Genomic DNA (10 μg) from the indicated species was digested with *Eco* RI, fractionated, transferred to a nylon membrane as described in Materials and methods, and hybridized with radiolabelled fragment 5. Numbers on the left indicate molecular weight (in kb).

Discussion

The structure of the two genomic clones encoding *S*-RNase alleles of *N. alata* reported here is similar to that of *S*-RNase genes from *P. inflata*, *Solanum tuberosum*, and *Solanum chacoense* [8, 16, 30]. Each gene contains a single, small intron and the coding regions are surrounded by repetitive sequences. Apart from common features such as the TATA box, the upstream flanking region of *S*-RNase genes lacks obvious short, conserved sequences that could act as sites for interaction with specific transcription factors to regulate *S*-RNase gene expression. A number of sequence motifs have been identified in the flanking regions of two *S*-RNase alleles (S_1 and S_2) from *S. tuberosum* [16, 18] including two heptamers (ACATAAT and ATTATGT) that could potentially form an inverted repeat. The heptamers in the *S*-RNase genes from *S. tuberosum* are separated by 27 or 29 bp depending on the allele, and are also present and separated by a similar number of nucleotides in the S_1- and S_3-RNase genes from *P. inflata* [8] and the S_{11}-RNase gene from *S. chacoense* [30]. The heptamers were not found in the sequenced regions of the S_2-RNase gene from *N. alata* but were present in the S_6-RNase gene where they are separated by 3 kb; the second heptamer is found beginning at position + 744, 12 bp upstream of the termination codon (Fig. 2). A second motif reported by Kaufmann *et al.* [16] to be present in the S_1- and S_2-RNase genes from *S. tuberosum*, AGGAATA-15 bp-CTCACACT, was not found in either of the genes from *N. alata*, nor is it present in either of the sequenced genes from *P. inflata* [8]. The finding that a sequence element (the 56/59 box) required for the expres-

sion of two pollen-expressed genes from tomato, LAT 56 and LAT 59 [34] and also present in the upstream regions of two *S. tuberosum S*-RNase genes [18], was of interest in light of a recent report by Dodds *et al.* [11] that the S_2- and S_6-RNase genes presented here are expressed in pollen during anther development. However, no significant similarity to the 56/59 box was found in either *S*-RNase gene from *N. alata*. The functional significance of any of these motifs in regulating expression of *S*-RNase genes will only be clarified by a thorough study of the promoter region.

One of the features of all *S*-RNase genes characterized to date is the diversity of the repetitive sequences found flanking the gene. Each of the 10 fragments produced from DNA flanking the S_6-RNase coding region detected repetitive sequences when used to probe total genomic DNA from *N. alata*. Previously, Bernatzky *et al.* [5] showed that a 1 kb fragment upstream of the S_2-RNase gene of *N. alata* contained a dispersed, repeated sequence that was also present in mitochondrial DNA from *N. alata* and *Lycopersicon esculentum*. Similarly, Coleman and Kao [8] found repetitive elements surrounding both the S_1- and S_3-alleles of *P. inflata*, and Saba-El-Leil *et al.* [30] found repetitive sequences surrounding the S_{11} allele of *S. chacoense*. Plant nuclear genomes generally contain a very high content of repetitive DNA and it is variation in the amount of repetitive DNA that is predominantly responsible for the range of total DNA content found among species of higher plants [12, 29]. Using DNA/DNA reassociation kinetics, Zimmerman and Goldberg [35] found that 80% of the single-copy sequences present in the genome of *N. tabacum* were interspersed with moderately repetitive DNA. Although differing in DNA content [28], it is likely that the *N. alata* and tobacco genomes have a similar overall organization. Since the *N. alata* genome contains a single copy of the S_6-RNase gene, the repetitive DNA surrounding this gene should be typical of most genes from this species. Indeed, the interspersion of repetitive and single-copy DNA is a general characteristic of most eukaryotic genes [10].

Repetitive elements can be used to distinguish closely related species and those present around the S_6-RNase gene are found in other species of the genus *Nicotiana*, although their abundance varies. The repeat found on probe 5 hybridizes strongly with DNA from *N. sylvestris* and *N. tabacum*, but weakly with DNA from *N. tomentosiformis* (Fig. 5). It is likely that this sequence was contributed to the *N. tabacum* genome by *N. sylvestris*, one of the progenitors of tobacco. A different repeat that also hybridizes to *N. sylvestris* but not *N. tomentosiformis* DNA, has been used in a molecular cytogenetic approach to study the origin of the *N. tabacum* genome [17].

Self-incompatibility requires interaction of allelic products of the *S*-locus from pollen and style and recent evidence has confirmed that *S*-RNases are indeed stylar products of the *S*-locus [19, 20, 26]. The product of the *S*-locus in pollen is not known and, although *S*-RNase genes are also expressed in pollen [7, 11], mutational studies indicate that the determinants of self-incompatibility in style and pollen are likely to be different. Several approaches are being used in attempts to identify unequivocally the product of the *S*-locus in pollen. One approach is to clone other genes present at the *S*-locus [22]. As part of this approach, we have identified a number of large genomic fragments containing the S_6-RNase gene from *N. alata* that would be suitable for cloning in YAC or cosmid vectors (Fig. 3 and Table 1). However, the repetitive DNA present in the genus *Nicotiana* in general [35] and surrounding genes encoding *S*-RNases in particular ([5], this paper) will make map-based cloning of the gene(s) encoding the product of the *S*-locus in pollen difficult.

The *S*-locus has a large number of alleles [22] and this allelic richness is also apparent at the molecular level in the extraordinary divergence of sequences used to encode *S*-RNases, even within a single species [33]. As noted above, the flanking regions of different *S*-RNase genes are even more divergent in sequence than are the sequences of the *S*-RNases themselves. Despite this divergence, all the *S*-RNases studied share the same pattern of gene expression and are assumed to be regulated in the same way. Progress in under-

standing sequences involved in regulating *S*-RNase gene expression in *N. alata* has been hampered by difficulties encountered in using the 5′-flanking regions of the S_2- and S_6-RNase genes to express *S*-RNase coding regions or the GUS reporter gene in transgenic plants [27]. Recently Lee *et al.* [20] have overcome this constraint and expressed sense and antisense *S*-RNase constructs in transgenic petunias using a fragment of ca. 2 kb from the 5′-flanking region of the S_3-RNase gene of *P. inflata* as the promoter. Also, Kirch *et al.* [18] reported that a 5 kb fragment from the upstream region of the S_2-RNase gene of *S. tuberosum* effectively directs GUS expression in a developmentally regulated manner in transgenic tobacco. These observations open the way to further experiments to define the sequences that regulate expression of *S*-RNase genes.

Acknowledgements

The authors would like to thank Bruce McGinness for his expert care and handling of the plants and an anonymous reviewer for their comments on this manuscript. D.P.M. is the recipient of a Fellowship from the Natural Sciences and Engineering Research Council of Canada, S.O. was supported by a Fellowship from the Japanese Society for the Promotion of Science. The authors acknowledge the continuing support of a Special Research Centre grant from the Australian Research Council.

References

1. Anderson MA, Cornish EC, Mau S-L, Williams EG, Hoggart R, Atkinson A, Bönig I, Greg B, Simpson R, Roche PJ, Haley JD, Penschow JD, Niall HD, Tregear GW, Coghlan JP, Crawford RJ, Clarke AE: Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature: 321: 38–44 (1986).
2. Anderson MA, McFadden GI, Bernatzky R, Atkinson A, Orpin T, Dedman H, Tregear G, Fernley R, Clarke AE: Sequence variability of three alleles of the self-incompatibility gene of *Nicotiana alata*. Plant Cell 1: 483–491 (1989).
3. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current Protocols in Molecular Biology. Wiley, New York (1987).
4. Bernatzky R, Tanksley SD: Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203: 8–14 (1986).
5. Bernatzky R, Mau S-L, Clarke AE: A nuclear sequence associated with self-incompatibility in *Nicotiana alata* has homology with mitochondrial DNA. Theor Appl Genet 77: 320–324 (1989).
6. Chomczynski P, Sacchi N: Single-step method of RNA isolation by guanidinium thiocynate-phenol-chloroform extraction. Anal Biochem 162: 156–159 (1987).
7. Clark RK, Sims TL: The S-ribonuclease gene of *Petunia hybrida* is expressed in nonstylar tissue, including immature anthers. Plant Physiol 106: 25–36 (1994).
8. Coleman CE, Kao T-h: The flanking regions of two *Petunia inflata S* alleles are heterogeneous and contain repetitive sequences. Plant Mol Biol 18: 725–737 (1992).
9. Creusot F, Sevignac M, Dron M: Large-scale mapping using high-molecular-weight DNA from *P. vulgaris*. Plant Mol Biol Rep 10: 207–218 (1992).
10. Davidson EH, Galau GA, Angerere RC, Britten RJ: Comparative aspects of DNA sequence organization in Metazoa. Chromosoma, Berl 51: 253–259 (1975).
11. Dodds PN, Bönig I, Du H, Rödin J, Anderson MA, Newbigin E, Clarke AE: S-RNase gene of *Nicotiana alata* is expressed in developing pollen Plant Cell 5: 1771–1782 (1993).
12. Flavell RB: Repetitive DNA and chromosome evolution in plants. Phil Trans R Soc Lond Ser B 312: 227–242 (1986).
13. Gray JE, McClure BA, Bönig I, Anderson MA, Clarke AE: Action of the style product of the self-incompatability gene of *Nicotiana alata* (*S*-RNase) on *in vitro*-grown pollen tubes. Plant Cell 3: 271–83 (1991).
14. Hanley BA, Schuler MA: Plant intron sequences: evidence for distinct groups of introns. Nucl Acids Res 16: 7159–7176 (1988).
15. Huang S, Lee H-S, Karunanandaa B, Kao T-h: Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. Plant Cell 6: 1021–1028 (1994).
16. Kaufmann S, Salamini F, Thompson RD: Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. Mol Gen Genet 226: 457–466 (1991).
17. Kenton A, Parokonny AS, Gleba YY, Bennett MD: Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Gen Genet 240: 159–169 (1993).
18. Kirch H-H, Kaufmann H, Wemmer T, Liu J, Thompson RD: Molecular analysis of gametophytic self-incompatability. In: Dattée Y, Dumas C, Gallais A (eds) Reproductive Biology and Plant Breeding, pp. 163–175. Springer-Verlag, Berlin (1992).

19. Kowyama Y, Kunz C, Lewis I, Newbigin E, Clarke AE, Anderson MA: Self-compatibility in a *Lycopersicon peruvianum* variant (LA2157) is associated with a lack of style *S*-RNase activity. Theor Appl Genet 88: 859–864 (1994).
20. Lee H-S, Huang S, Kao T-h: S-proteins control rejection of incompatible pollen in *Petunia inflata*. Nature 367: 560–563 (1994).
21. Lewis D: Genetic control of specificity and activity of the *S*-antigen in plants. Proc R Soc Lond Ser B 151: 468–477 (1960).
22. Matton DP, Nass N, Clarke AE, Newbigin E: Self-incompatibility: how plants avoid illegitimate offspring. Proc Natl Acad Sci USA 91: 1992–1997 (1994).
23. McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE: Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342: 955–57 (1989).
24. McClure BA, Gray JE, Anderson MA, Clarke AE: Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature 347: 757–760 (1990).
25. McClure BA, Du H, Liu Y-H, Clarke AE: S-locus products of *Nicotiana alata* pistils are subject to organ-specific post-transcriptional processing but not post-translational processing. Plant Molec Biol 22: 177–181 (1993).
26. Murfett J, Atherton TL, Mou B, Gasser CS, McClure BA: S-RNases expressed in transgenic *Nicotiana* cause S-allele-specific pollen rejection. Nature 367: 563–566 (1994).
27. Murfett J, Ebert PR, Haring V, Clarke AE: An S-RNase promoter from Nicotiana alata functions in transgenic N. alata plants but not Nicotiana tabacum. Plant Molec Biol 28: 957–963 (this issue) (1995).
28. Narayan RKJ: Nuclear DNA changes, genome differentiation and evolution in *Nicotiana* (Solanaceae). Pl. Syst. Evol. 157: 161–180 (1987).
29. Narayan RKJ, Rees H: DNA, heterochromatin and phylogeny of *Nicotiana* amphidiploids. Chromosoma 47: 75–83 (1974).
30. Saba-El-Leil MK, Rivard S, Morse D, Cappadocia M: The S11 and S13 self incompatibility alleles in *Solanum chacoense* Bitt. are remarkably similar. Plant Molec Biol 24: 571–583 (1994).
31. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW: Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81: 8014–8018 (1984).
32. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
33. Tsai D-S, Lee H-S, Post LC, Kreiling KM, Kao T-h: Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other solanaceous S-proteins. Sex Plant Reprod 5: 256–263 (1992).
34. Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S: Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Devel 5: 496–507 (1991).
35. Zimmerman JL, Goldberg RB: DNA sequence organization in the genome of *Nicotiana tabacum*. Chromosoma 59: 227–252 (1977).