

# Sheltered load associated with S-alleles in *Solanum carolinense*

JL Stone

Department of Zoology, Duke University, Durham, NC 27708-0325, USA

Bud pollinations allowed me to examine the effects of homozygosity at loci in the area of suppressed recombination around the S-locus in *Solanum carolinense*, whose S-alleles show surprisingly low diversification rates. The total number of seeds produced was lower for incompatible than compatible pollinations, revealing that self-incompatibility was only somewhat overcome by bud pollination. However, low seed set in incompatible crosses was not due solely to the incompatibility response; crosses between distinct plants sharing the same alleles at the S-locus led to dramatically high seed abortion, nearly equal to that found upon selfing.

An excess of heterozygotes in the surviving progeny supports the supposition that these high abortion rates are due to sheltered load, that is, previously unexpressed load accumulated due to enforced heterozygosity and recombination suppression around the S-locus. Of the seven alleles examined in total, two showed evidence of severe load and five did not. The magnitude of load was consistent with terminal branch length in some, but not all, cases.

*Heredity* (2004) 92, 335–342, advance online publication, 4 February 2004; doi:10.1038/sj.hdy.6800425

**Keywords:** inbreeding depression; lineage-specific effects; recombination; self-incompatibility; S-locus; sheltered load

## Introduction

Genes within organisms evolve at varying rates depending on many factors, including rates of mutation and recombination, and the effects of selection. Some of the most spectacular examples of genes evolving at long time scales are those exhibiting 'trans-specific evolution': major-histocompatibility loci in mammals and self-incompatibility alleles in flowering plants. In these cases, a gene genealogy reveals that the lineages representing alleles are older than the species that bear them; the diversification of allelic lineages predates speciation (Klein, 1986; Ioerger *et al.*, 1990). The extremely long residence time of allelic lineages suggests that the structure of the gene genealogy may be useful for inferring demographic history of natural populations (Takahata, 1990).

Work on the genealogical structure of incompatibility alleles sequenced from natural populations has revealed a puzzling lack of correlation between allele number and number of trans-specific lineages (eg Richman *et al.*, 1996; Uyenoyama, 1997a,b). For example, *Physalis crassifolia* possesses 34 alleles residing in two lineages, while *Solanum carolinense* possesses 13 alleles scattered over multiple lineages within the Solanaceae. The lack of correlation between allele and lineage number is surprising because differences in historical population size should lead to a positive correlation between allele and lineage number, rather than a negative one, as is sometimes observed. Simulations of expected ratios of divergence times showed that *S. carolinense* departs from

expectations, showing longer than expected terminal branch lengths (Uyenoyama, 1997b).

Differences in the fixation rate of mutations may explain the lack of correlation between allele number and divergence times. In particular, lineage-specific load constraining diversification, combined with a difference in demographic history, could result in the observed pattern (Uyenoyama, 1997a). Lineage-specific load refers to deleterious recessive mutations associated with particular alleles at the S-locus, and is expected to accumulate due to two factors. First, the S-locus in the Solanaceae, by definition, is always expressed in heterozygous condition, as pollen is rejected if it bears either allele held by the maternal plant. Second, the S-locus is in a non-recombining region of the genome (Coleman and Kao, 1992). Therefore, deleterious recessive mutations that accumulate in the region surrounding the S-locus will not normally be exposed to selection. As these mutations are sheltered by being kept in heterozygous state, the lineage-specific load is called 'sheltered load'.

Sheltered load will be expressed when a mutation within the specificity-determining region of the S-locus leads to the formation of a new S-allele. The new allele will be compatible with any other in the population, leading to the formation of embryos containing both the new allele and its progenitor. Suddenly, the regions surrounding the S-locus will be combined in homozygous condition, and the lineage-specific load will be expressed. If this load is severe enough, selection will eliminate one of the two alleles in the population (Uyenoyama, 2003). Thus, the fixation of new mutations and the diversification of the lineage will be slowed. In this paper, I seek to determine the magnitude of sheltered load in *S. carolinense*, the species showing the longest terminal branches found to date.

## Materials and methods

### Basic approach

In the Solanaceae, self-incompatibility is mediated by an RNase that is expressed at high levels in the style at the time the flower opens, causing the degradation of pollen tubes bearing alleles matching either of the two alleles held by the maternal plant (McClure *et al*, 1990). The self-incompatibility reaction can be overcome, to some extent, by applying pollen to stigmas before the flower opens and the RNase is expressed at high levels ('bud pollination'; de Nettancourt, 1977). I performed bud pollinations to create progeny that are homozygous at the S-locus. If a flower is pollinated in the bud with pollen from an individual sharing its genotype at the S-locus, the null expectation is that half of the seeds in the fruit will be heterozygous at the S-locus and the other half will be divided equally between the two homozygote classes. An excess of heterozygotes would provide evidence for sheltered load. I performed self, incompatible outcross, and compatible outcross bud pollinations for individuals sharing genotypes at the S-locus, and measured seed production, abortion rates, germination rates, and segregation ratios.

### Plant material and pollinations

*S. carolinense* is a weedy perennial widely distributed in eastern North America. In fallow land or landscaped areas, it typically consists of small populations of less than 20 individuals, but in pastures and in row crops, populations often number in thousands. Numerous *S. carolinense* plants were collected by digging up the rhizomes of individuals separated by at least 3 m from a large population in Hillsboro, North Carolina. Genotypes at the S-locus were determined following Richman *et al* (1995). Three pairs of individuals who shared the same genotype at the S-locus were subjected to two tests. To verify that they were incompatible, reciprocal pollinations were performed on open flowers, and pollen tubes were stained with aniline dye and visualized using fluorescence microscopy (Martin, 1959). To verify that they were not ramets of a single genet, plants were screened for variation at four isozyme loci using three enzyme systems: peroxidase (PER), isocitric dehydrogenase (IDH), and malic acid (ME; Soltis and Soltis, 1989). In addition to working with the three pairs of individuals who shared genotypes at the S-locus, I also included two individuals with distinct genotypes. In order to increase the available sample of flowers, individuals were cloned by planting pieces of root stock in four separate pots for each plant.

The three classes of pollinations (outcross compatible, outcross incompatible, and self) were performed while flowers were 4–6 mm long, about 2 days before opening. I collected pollen from the poricidal anthers of open flowers using a battery-operated device that vibrates the anthers, causing pollen to shoot out into a clean glass vial (J Verne Coontz, Centre Hal, Penna.). A clean toothpick was used to coat the stigmatic surface with pollen. Individual flowers were pollinated by single sires, and outcross sires were used with approximately equal frequency across maternal plants. As seed production per fruit declines with distance from the base of the inflorescence, I pollinated only the basal four flowers out

of the 10 or more per inflorescence, rotating basal flowers across pollination treatments. I performed about 10 pollinations per treatment per clone. Flowers were marked and assigned consecutive numbers.

### Seed numbers and germination

I collected ripe fruits and allowed them to dry naturally. Upon examining the seeds, it became evident that some fruits contained two types of seeds: plump, rounded seeds resembling tomato seeds, and much smaller or flattened seeds. This second category was labeled 'aborted'. For each fruit (identified by consecutive number, not by treatment), good seeds and aborted seeds were tallied separately. In all, 20 good seeds from each fruit were planted without pretreatment into seedling mix in small flats. For each mother showing sufficient quantities of aborted seeds, a sample of 50 seeds designated as aborted was also planted simultaneously. Germinating seeds were exposed to supplemental lights on a 12-h period. Seedlings emerged in about 2 weeks and were counted and harvested after 3 weeks.

### Estimates of genetic load – seed numbers

Crosses using incompatible but distinct individuals allow the simultaneous estimation of sheltered load and non-S-linked load expressed upon selfing. A third parameter must also be estimated; the degree to which bud pollinations fail to overcome the self-incompatibility response. I estimate these parameters by comparing total and viable seed production from incompatible, compatible, and self pollinations. Table 1 shows relationships between parameters as reflected in predicted seed numbers. I assume that all mortality is detected, that is, that no embryos die before they are large enough to be seen. This assumption may lead to an overestimate of  $p$ , the proportion of incompatible tubes achieving fertilization, and to underestimates of  $h$  and  $d$ , proportional reduction in survivorship due to the sheltered and genome-wide components of load. I also assume that heterozygotes suffer no sheltered effect; that is that co-occurring alleles share no allele-specific load. This assumption seems appropriate for this case in which co-occurring alleles are distantly related, for example, coming from trans-specific clades, but would need to be modified for genotypes combining alleles that have a most recent common ancestor after the time of speciation. Parameters were estimated using linear regression of square-root-transformed seed counts averaged for each clone. Regressions were constrained to pass through the origin.

### Progeny genotypes

For 40 seedlings from incompatible pollinations, I screened progeny to see whether or not they were heterozygous at the S-locus. A direct test for heterozygosity, such as single-strand conformational polymorphism (SSCP) would have been ideal for this purpose, but S-alleles are so divergent from each other that I have not been able to create a single primer pair to uniformly amplify both alleles from genomic DNA. Instead, I designed allele-specific primers (Figure 1, Table 2) and performed PCR twice on each sample, once with each set of primers. A heterozygous individual

should give amplification products for both PCRs; a homozygote for only one. To design the primers, I first identified likely regions for each allele and then used Primers! for the Mac to select particular sequences (Resnick, 1996).

Total genomic DNA was extracted from about 0.02 g of dried leaf tissue using the Puregene DNA isolation kit of Gentra systems. PCR was performed in a 50 µl volume of 50 mM KCl, 10 mM Tris-HCl (pH 8.0), 1.5 mM MgCl<sub>2</sub>, 10 mM dNTPs, 50 ng of each primer, 2 U AmpliTaq Taq polymerase, and about 20 ng DNA. Two drops of mineral oil covered each reaction. Following an initial 2.5 min denaturation step at 94°C, amplification occurred over 40 cycles of 1 min at 94°C, 2 min at 55°C, and 1.5 min at 72°C. PCR products were run out in a volume of 15 µl on 2% agarose gels at 90 V.

Each primer pair was tested on a standard sample of 10 plants with known genotypes. In no case did primers amplify a product in a sample known to lack the allele. To address the potential problems with false negatives, I extracted DNA from two 'null samples' for each 10 experimental samples. These null samples were progeny from compatible pollinations and served as combined positive and negative controls. Nulls should amplify for one but not both alleles. In some cases, nulls did not amplify for either allele; in these cases, the PCR was repeated for nonamplified samples of that extraction, first by adding more Taq polymerase to the existing reaction, and then by adjusting the amount of template in a new reaction, until the blank null sample showed an amplification product. In no case did a null sample show PCR products for both alleles.

**Table 1** Predicted seed production for three types of crosses

	Compatible	Incompatible	Self
Total seeds/fruit	$N_1$	$N_1p$	$N_1p$
Viable seeds/fruit	$N_2$	$N_2p(1+h)/2$	$N_2pd(1+h)/2$

$N_1$  and  $N_2$  represent baseline production for outcross compatible pollinations.  $p$  = fertilization rate of incompatible (relative to compatible) pollen,  $h$  = survivorship of S-locus homozygotes relative to heterozygotes, and  $d$  = survivorship due to non-S-linked load of selfers relative to outcrossers

## Estimates of genetic load – genotype frequencies of progeny

I used maximum likelihood to estimate the degree of sheltered load averaged between co-occurring alleles. Assume that homozygotes and heterozygotes are formed in equal proportions at fertilization. The proportion of homozygous progeny at fertilization is  $N_{\text{homo}}/(N_{\text{homo}} + N_{\text{het}})$ . Subsequent to fertilization, in the presence of sheltered load, the frequency of homozygotes is reduced by the factor  $h$ . The maximum-likelihood estimate of  $h$  is  $x/(1-x)$ , where  $x$  is the proportion of homozygous progeny.

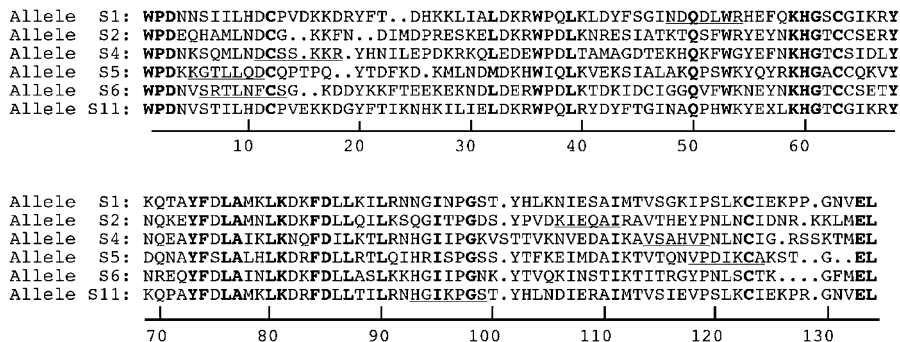
## Results

### Pollen tubes and isozymes

Of the four pairs of plants with the same genotype at the S-locus, only one pair had distinct genotypes at allozyme loci. Plants 15 and 20, having S-genotype  $S_4 S_6$ , clearly differed from each other at both the cathodal PER locus and at the ME locus, indicating that they are not merely different shoots from a single rootstock. Crosses between plants 15 and 20 produced many pollen tubes, which were immediately arrested within the style. These plants were verified as being both genetically distinct and mutually incompatible. Based on their difference at allozyme loci, we assume that plants 15 and 20 are unrelated, and that matings between them are random outcrosses. This assumption provides an upper bound to  $h$ , the survivorship of homozygotes relative to heterozygotes at the S-locus.

**Table 2** Primer sequences used for allele-specific PCR

Allele	Length	$T_m$	Sequence
$S_2$	261	60.1	tg gat ccc cgt gag agt aag t gat ggc ttg ttc gat ctt g
$S_4$	324+intron	59.3	ac tgc tct tcg aaa aag cgt
		59.7	aa gtt agg aac atg tgc act cac a
$S_5$	348+intron	61.6	aag ggt aca ctg ctg cag gac
		59.7	tgc aca ctt gat gtc agg aac
$S_6$	262+intron	59.8	g agc aga acg ttg aat ttc tgt t
		60.2	tcc atg ctt ttt caa aga tgc
$S_{11}$	199	62.7	ac aaa cgc tgg cct caa ctc
		57.2	ga acc tgg gtt tat tcc att



**Figure 1** Alignment of partial amino-acid sequences showing locations of allele-specific primers. Alignment from Richman *et al* (1995). Conserved sites are in boldface. Primer locations are underlined.

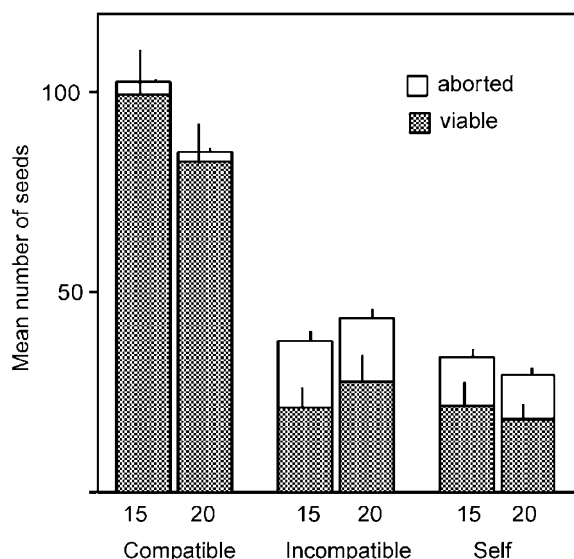
Other potential incompatible pairs (18 and 25, genotype  $S_2 S_{11}$ ; 9 and 11, genotype  $S_5 S_{13}$ ) failed to show any allozyme differences. Clones of these pairs also resembled each other, but were distinct from other clones for morphological and phenological traits (unpublished data). They were therefore treated as genets of a single ramet. For these cases, there is no outcross incompatible treatment, but it is still instructive to examine abortion rates upon selfing and proportion of heterozygote embryos surviving to the seedling stage.

### Seed numbers and germination rates

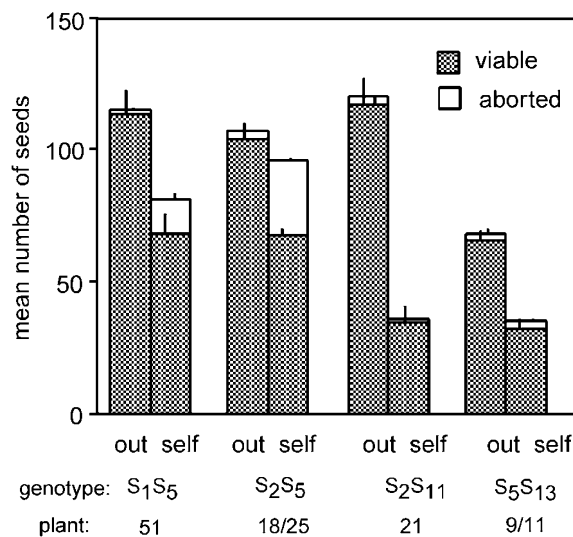
Fruits from compatible pollinations had many more seeds than did fruits from outcross incompatible pollinations, revealing that bud pollinations were only partially successful in overcoming the self-incompatibility reaction (Figure 2). Abortion rates for compatible crosses were low (Figure 2). In contrast, about one-third of seeds formed from outcross incompatible crosses were aborted, nearly indistinguishable from the abortion rates of selfed incompatible seeds (Figure 2).

Comparison of selfed versus outcrossed total seed numbers shows that some individuals produce seed more easily from incompatible pollinations than do others (Figure 3). A factor that may contribute to this observation is different timing of S-RNase expression among alleles (Clark *et al.*, 1990). Specifically, genotypes including  $S_{11}$  or the closely-related  $S_{13}$  show much reduced seed set of incompatible pollinations, which is consistent with early-acting activity of the S-RNase for these alleles. Genotypes lacking these alleles showed less of a decrement of total seeds upon selfing, but relatively higher abortion rates (Figure 3).

Seeds that had been classified as aborted uniformly failed to germinate. Seed germination varied widely across maternal plants (Table 3;  $P < 0.0001$  from mixed-model ANOVA on arcsin-transformed germination rates). Crosstype did not significantly affect germination for most plants, with the exception of plant 9/11, where



**Figure 2** Total numbers of seeds, divided into viable and aborted, following bud pollinations of cloned *S. carolinense* individuals 15 and 20 (genotype  $S_4 S_6$ ). Vertical lines indicate SEs.



**Figure 3** Total numbers of seeds, divided into viable and aborted, following bud pollinations of cloned *S. carolinense* individuals 51( $S_1 S_5$ ), 18/25 ( $S_2 S_5$ ), 21 ( $S_2 S_{11}$ ), and 9/11( $S_5 S_{13}$ ). Vertical lines indicate SEs.

**Table 3** Germination rates of seeds produced by self and outcross bud pollinations for six cloned plants of five genotypes at the S-locus

Genotype	Plant	Outcross	Self	Significance level
$S_1 S_5$	51	0.81	0.89	NS
$S_2 S_5$	18.25	0.82	0.80	NS
$S_2 S_{11}$	21	0.95	0.78	NS
$S_4 S_6$	15	0.59	0.52	NS
$S_4 S_6$	20	0.42	0.37	NS
$S_5 S_{13}$	9.11	0.36	0.27	$P < 0.01$

Significance was determined using a series of unpaired *t*-tests on arcsin-transformed germination rate. NS = nonsignificant.

selfed progeny germinated at a significantly lower rate than outcross progeny (Table 3). About one-quarter of the selfed progeny of plant 18/25 lacked chlorophyll and died soon after germination.

### Estimates of genetic load – seed numbers

Following the predictions from the top row in Table 1, a regression of total numbers of incompatible seeds on total number of compatible seeds gives the proportion of incompatible pollen tubes achieving fertilization,  $P = 0.61$  (Table 4). This value was then treated as a constant in the second set of regressions, where numbers of viable incompatible seeds were regressed on numbers of viable seeds from outcross compatible pollinations. The coefficient in the regressions involving viable seeds gives  $(1+h)/2$ , for outcross incompatible pollinations (Table 4). The average value of  $h$  for the two alleles ( $S_4$  and  $S_6$ ) is 0.56. In summary, seed production data suggest that almost half of incompatible pollinations did not succeed, and that for those that succeeded, almost half of the progeny died.

**Table 4** Linear regressions based on averages of square-root-transformed seed production for four clones of plant 15 and three clones of plant 20

Independent, dependent variables	Model d.f. Resid. d.f.	Model MS Resid. MS	F	P > F	R <sup>2</sup>	Coeff. (95% CI)
Total compatible, total incompat	1 6	260.87 1.60	163.2	0.0000	0.964	$p = 0.607$ (0.49–0.72)
Viable compatible, viable incompat	1 6	416.53 3.09	134.9	0.0000	0.957	$(1+h)/2 = 0.780$ (0.62–0.94)
Viable incompat, viable self (15)	1 3	68.86 1.43	48.25	0.006	0.942	$d_{15} = 0.923$ (0.50–1.3)
Viable incompat, viable self (20)	1 3	48.86 0.05	1019.3	0.001	0.998	$d_{20} = 0.784$ (0.68–0.89)

The first row gives a regression of total seeds from incompatible and self on compatible bud pollinations, for an estimate of  $p$  (see Table 1). The second row gives a regression of viable seeds from incompatible on outcross pollinations, for an estimate of  $(1+h)/2$ , where  $h$  = relative survivorship of S-locus homozygotes. The dependent variable for the second row was divided by the constant  $p$ . The third and fourth rows give regressions of viable seeds from self on incompatible pollinations for an estimate of  $d$  (relative survivorship of selfed progeny due to non-S-linked load) for plants 15 and 20.

Selfing appears to be slightly more detrimental than simple homozygosity around the S-locus (bottom two rows, Table 4). For plant 15, zygotes from selfed fruits survived at a rate of about 0.9 of those from incompatible fruits, with the confidence interval including 1, suggesting that inbreeding depression upon selfing for that plant may be largely attributable to the S-locus effects. For plant 20, the relative survivorship of selfed progeny ( $d$ ) was about 0.8. The confidence interval for this estimate did not include 1, therefore for this maternal family, selfing does confer fitness costs beyond those imposed by S-locus-linked loci.

For plants where outcross incompatible pollinations were not available, we can still gain a maximum estimate of sheltered load by imagining the case where all early-acting inbreeding depression is attributable to S-locus-linked effects (Table 5). The estimates of  $h_{\min}$  give the proportion of embryos killed by sheltered load, assum-

ing that there is no non-S-linked load. If  $h_{\min}$  is close to 1, we can say with confidence that sheltered load is negligible; as  $h_{\min}$  declines, the possibility increases that sheltering load is important. Even with the generous assumption of no early-acting inbreeding load except at the S-locus, the  $h_{\min}$  for plant 21 (genotype  $S_2 S_{11}$ ) was close to 1.0, indicating negligible sheltered load. Values of  $h_{\min}$  for all four genotypes are considerably higher than the mean  $h = 0.56$  found for the plants 15 and 20 (genotype  $S_4 S_6$ ), suggesting that the dramatically high estimate of sheltered load for alleles in that genotype may not be typical for other alleles in this species.

#### Progeny genotypes

Consistent amplification of alleles from genomic DNA was achieved for alleles  $S_2$ ,  $S_4$ ,  $S_5$ ,  $S_6$ , and  $S_{11}$ , allowing the analysis of progeny arrays for three crosses. The proportion of progeny which were heterozygous at the S-locus that survived from incompatible crosses differed among plants with different genotypes at the S-locus (Table 6). For the plant with genotype  $S_2 S_{11}$ , progeny met expectations of random mating. In contrast, for plants with genotypes  $S_4 S_6$  and  $S_2 S_5$ , one of the two homozygous classes was under-represented (Table 6).

#### Estimates of genetic load – genotype frequencies of progeny

The maximum-likelihood estimate of sheltered load differed dramatically across the three genotypes. Genotype  $S_2 S_{11}$  showed no load, with the ml estimate of  $h$  equal to 1.0. This is congruent with the value of  $h = 0.98$  estimated through regression on seed production. For genotype  $S_4 S_6$ , the ml estimate of  $h$  was 0.38, lower than the estimate of 0.56 obtained by regression on seed

**Table 5** Linear regressions based on averages of square-root-transformed seed production from self and outcross pollinations of four sets of cloned plants representing four genotypes at the S-locus

Genotype	Plant	$p$	$h_{\min}$
$S_1 S_5$	51	0.81	0.86
$S_2 S_5$	18.25	0.88	0.69
$S_2 S_{11}$	21	0.53	0.98
$S_5 S_{13}$	9.11	0.67	0.87

See text for explanations of regressions for  $p$ , the proportion of incompatible pollen tubes fertilizing ovules during bud pollinations, and  $h_{\min}$ , the relative survivorship of homozygotes at the S-locus, assuming that non-S-linked early-acting load is zero.  $h_{\min}$  provides a boundary on the possible severity of sheltered load for these genotypes.

**Table 6** Numbers of seedlings, resulting from incompatible bud pollinations, heterozygous or homozygous for alternate alleles

Genotype	Plant(s)	1st Homo	Hetero	2nd Homo	G-statistic	Significance
$S_2 S_5$	18	16 (11.5)	25 (23)	5 (11.5)	6.41	$P < 0.05$
$S_2 S_{11}$	21	8 (9.5)	19 (19)	11 (9.5)	0.47	NS
$S_4 S_6$	15 and 20	2 (10)	29 (20)	9 (10)	13.2	$P < 0.005$

1st homo refers to the first allele listed in the genotype. Values expected if there were no S-associated load are in parentheses.

numbers, although within the confidence limits of that estimate. Finally, for genotype  $S_2S_5$ , the ml estimate of  $h$  was 0.84, higher than the estimate of 0.69 provided by the regression on seed numbers, although again, not significantly so. The point estimates of  $h$  obtained by the two approaches differ by about 20% in either direction, but the rank order of the estimates is consistent, with genotype  $S_2S_{11}$  showing no load, genotype  $S_2S_5$  showing intermediate values, and genotype  $S_4S_6$  showing severe levels amounting to almost complete mortality for one of the homozygous classes.

## Discussion

The suppression of recombination around the S-locus, combined with its obligatory heterozygosity, may permit the accumulation of linked deleterious recessive alleles. When thrust into homozygous state, such as by the appearance of a new S-specificity together with its progenitor, the accumulated load will be expressed. Experiments presented here show that this sheltered load varies among S-alleles in horse nettle, *S. carolinense*. The load is expressed by seed abortion and, to a lesser extent, reduced germination. The existence of load was verified by genotypes of progeny arrays showing a deficiency of homozygous progeny. For two alleles,  $S_4$  and  $S_5$ , the load causes almost complete mortality for homozygous offspring, while for five others, the effect appears to be minor (Tables 5 and 6).

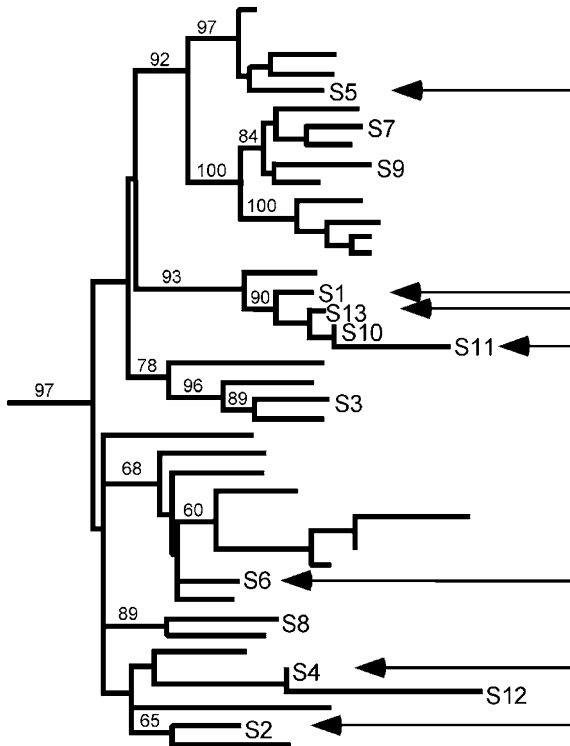
Different S-alleles may differ in the onset of timing of expression (Clark *et al*, 1990). In this study, for example, it appears that genotype  $S_2S_5$  had less effective rejection responses than did other genotypes in response to bud pollination. If two alleles within an individual had different levels of expression in the bud stage, a deficiency of homozygotes would be found for the allele that had the higher expression level. Thus homozygote deficiency itself is not a clear evidence of sheltered load. However, if the S-RNase for one allele was expressed so early that very few homozygotes of this class were formed, the number of heterozygotes would also be very low, in a 1:1 ratio with the less-abundant homozygote class. Therefore, heterozygote excess as observed in this study can be explained by sheltered load but not by differential expression times of alleles. Also, predictions of load from seed abortion rates were concordant with genotype frequencies, supporting the idea that sheltered load, and not different timing of expression, was responsible for the skewed genotype frequencies.

Just as homozygote deficiency does not provide unequivocal evidence for sheltered load, the lack of heterozygote excess does not, in itself, demonstrate the absence of sheltered load. The shared history of the two alleles of the genotype must be considered. If two alleles are recently diverged, the locations of their deleterious mutations will be shared, and thus homozygotes and heterozygotes will suffer equally from S-locus-linked load. In this study, however, none of the genotypes combined alleles that were at all similar; in fact, all pairs have been separated at least since speciation (Richman *et al*, 1996). Therefore, the lack of heterozygote excess for genotype  $S_2S_{11}$  can be interpreted as the lack of genetic load associated with these alleles, rather than as shared deleterious sites.

The concept of sheltered load was pioneered by Mather and de Winton (1941), who discovered a deficiency of homozygous progeny resulting from self-fertilization of the short-styled *Primula sinensis*. They posited that genetic load linked to the S-locus was 'sheltered' by the obligatory heterozygous condition of the short-styled morph in this distylous species. Although they did not discuss the extent to which recombination may be suppressed around the S-locus in *Primula*, the implication is that the suppressed recombination allowing maintenance of the S-locus itself extends to surrounding regions. Richards (1998) reviews more recent data and confirms the dearth of SS homozygotes, presumably due to sheltered load. Rasmuson (1980) also discusses the possibility that the linked loci may affect selection on the S-locus, although Leach *et al* (1986) suggest that the linkage would have to be absolute for the sheltered effect to persist. More recently, Glémin *et al* (2001) demonstrate that sheltered load does increase dramatically as recombination declines, but that load is maintained in the face of small amounts of recombination, on the order of  $r = 0.001$ . Recombination does appear to be completely suppressed around the S-locus in the Solanaceae (Coleman and Kao, 1992), but it is unclear how far this region of complete linkage extends. In incompatible *S. chacoense*, selfed progeny show severely distorted segregation ratios of markers linked to the S-locus (Rivard *et al*, 1996), suggesting the presence of allele-linked load.

Allele-specific load may also exist in families with completely different self-incompatibility systems. For example, Lawrence and Franklin-Tong (1994) documented the existence of S-linked selection affecting segregation ratios from controlled outcrosses in *Papaver*. A dramatic example of lethal loci linked to mating-system alleles is well documented in fungi (eg Fischer and Holton, 1957, p. 233). If, as expected, genomic regions placed into service as mating-type loci are frequently located in areas of suppressed recombination, sheltered load may be a commonplace force influencing diversification rates at these loci.

Genetic factors such as sheltered load do not act in isolation, but in concert with ecological factors such as population size to determine allelic diversity. In *S. carolinense*, precolonial population sizes were probably quite small, restricted to patches of open canopy in eastern deciduous forest. New alleles would be likely to encounter their progenitors frequently in these tiny populations, and only alleles with low levels of sheltered load would be likely to diversify. Alternatively, a species characterized by small population sizes might maintain S-alleles nonrandomly, with more distant alleles being selectively maintained. This is consistent with diversification patterns in *Physalis cinerascens* (Richman and Kohn, 1999), a fugitive species with few S-alleles. The analysis of tree characteristics shows that *P. cinerascens* has longer terminal branches than expected by chance, suggesting preferential loss of closely-related alleles (Richman and Kohn, 1999). The pattern is not universally supported by the data; however, a similar analysis on a rain-forest gap species, *Witheringia maculata*, suggests the maintenance of shorter than expected terminal branches, and therefore alleles more closely related than expected by chance (Richman and Kohn, 2000).



**Figure 4** Neighbor-joining tree showing phylogenetic relationships of S-RNases in *S. carolinense*. Alleles investigated in this study are indicated by arrows. Nonlabelled branches represent alleles from *Lycopersicon peruvianum*, *Nicotiana glauca*, *Petunia hybrida*, *Petunia inflata*, *Physalis crassifolia*, *S. chacoense*, and *S. tuberosum*. S-like RNases were used as an outgroup. Adapted from Richman *et al* (1996).

Sheltered load, as a product of random mutational processes, is expected to occur stochastically across S-alleles. In this study, I found two alleles that exhibited a high degree of sheltered load, with mortality rates from selfing almost entirely attributable to homozygosity around a single locus. Five alleles, in contrast, did not exhibit any detectable load. Three of the alleles with low load ( $S_1$ ,  $S_{11}$ , and  $S_{13}$ ) coexist in the single clade of *S. carolinense* S-alleles that exhibits diversification (Figure 4), suggesting that the lack of load around the S-locus has permitted the diversification of these particular alleles. One of the alleles showing high levels of load ( $S_5$ ) occurs on a long terminal branch, as would be expected if the load prevented the diversification of this lineage. However, phylogenetic patterns for the other three alleles do not provide support for the hypothesis that sheltered load affects diversification rates. Allele  $S_4$ , which exhibits high load, shares a branch (albeit a long one), with allele  $S_{12}$ . Alleles  $S_2$  and  $S_6$  do not show evidence of load yet occur on unbranched terminal lineages. Much more work will be necessary before we can evaluate the importance of sheltered load on diversification rates of S-alleles.

## Acknowledgements

I thank MK Uyenoyama for suggesting the problem and a mechanism for addressing it, A Richman for help with

genotyping plants in the base population, Y Lu for isozyme work, J Kohn, B McClure and T Holtsford, and anonymous reviewers for comments on the manuscript, and A Richman and J Kohn for sharing unpublished manuscripts. Financial support was provided by an NSF-Sloan Postdoctoral Fellowship in Molecular Evolution.

## References

- Clark KR, Okuley JJ, Collins PD, Sims TL (1990). Sequence variability and developmental expression of S-alleles in self-incompatible and pseudo-self-compatible *Petunia*. *Plant Cell* **2**: 815–826.
- Coleman CE, Kao T-H (1992). The flanking regions of two *Petunia inflata* S alleles are heterogeneous and contain repetitive sequences. *Plant Molecular Biology* **18**: 725–737.
- de Nettancourt D (1977). *Incompatibility in Angiosperms*. Springer-Verlag: Berlin.
- Fischer GW, Holton CS (1957). *Biology and Control of the Smut Fungi*. Ronald Press: New York.
- Glémin S, Bataillon T, Ronfort J, Mignot A, Olivieri I (2001). Inbreeding depression in small populations of self-incompatible plants. *Genetics* **159**: 1217–1229.
- Ioerger TR, Clark AG, Kao T-H (1990). Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc Natl Acad Sci, USA* **87**: 9732–9735.
- Klein J (1986). *Natural History of the Major Histocompatibility Complex*. Wiley: New York.
- Lawrence MJ, Franklin-Tong VE (1994). The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. IX. Evidence of an extra effect of selection acting on the S-locus. *Heredity* **72**: 353–364.
- Leach CR, Mayo O, Morris MM (1986). Linkage disequilibrium and gametophytic self-incompatibility. *Theor Appl Genet* **73**: 102–112.
- Martin FW (1959). Staining and observing pollen tubes in the style by means of fluorescence. *Stain Technol* **34**: 125–128.
- Mather K, de Winton D (1941). Adaptation and counter-adaptation of the breeding system in *Primula*. *Ann Bot* **5**: 297–311.
- McClure BA, Gray JE, Anderson MA, Clarke AE (1990). Self-incompatibility in *Nicotiana glauca* involves degradation of pollen rRNA. *Nature* **347**: 757–760.
- Rasmuson M (1980). Linkage to a self-sterility locus. Consequences for segregation and selection. *Heredity* **92**: 315–319.
- Resnick R (1996). *Primers! for the Mac.*. Apple Pi: Ashland, Mass.
- Richards AJ (1998). Lethal linkage and its role in the evolution of plant breeding systems. In: Owens SJ, Rudall PJ (eds) *Reproductive Biology*. Royal Botanic Gardens: Kew.
- Richman AD, Kohn JR (1999). Self-incompatibility alleles from *Physalis*: implications for historical inference from balanced genetic polymorphisms. *Proc Nat Acad Sci, USA* **96**: 168–172.
- Richman AD, Kohn JR (2000). Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Mol Biol* **42**: 169–179.
- Richman AD, Kao T-H, Schaeffer SW, Uyenoyama MK (1995). S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* **75**: 405–415.
- Richman AD, Uyenoyama MK, Kohn JR (1996). Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. *Science* **273**: 1212–1216.
- Rivard SR, Cappadocia M, Landry BS (1996). A comparison of RFLP maps based on anther culture derived, selfed, and hybrid progenies of *Solanum chacoense*. *Genome* **39**: 611–621.
- Soltis DE, Soltis PS (eds) 1989. *Isozymes in Plant Biology*. Dioscorides Press: Portland OR.

- Takahata N (1990). A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc Natl Acad Sci USA* **87**: 2419–2423.
- Uyenoyama MK (1997a). The evolution of breeding systems. In: Singh RS, Krimbas C (eds) *Evolutionary Genetics from Molecules to Morphology*. Cambridge University Press: New York.
- Uyenoyama MK (1997b). Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. *Genetics* **147**: 1389–1400.
- Uyenoyama MK (2003). Genealogy-dependent variation in viability among self-incompatibility genotypes. *Theor Popul Biol* **63**: 281–293.