Evolutionary Dynamics of Self-Incompatibility Alleles in Brassica

Marcy K. Uyenoyama

Department of Zoology, Duke University, Durham, North Carolina 27708-0325

Manuscript received December 30, 1999
Accepted for publication May 25, 2000

ABSTRACT

Self-incompatibility in Brassica entails the rejection of pollen grains that express specificities held in common with the seed parent. In Brassica, pollen specificity is encoded at the multipartite *S*-locus, a complex region comprising many expressed genes. A number of species within the Brassicaceae express sporophytic self-incompatibility, under which individual pollen grains bear specificities determined by one or both *S*-haplotypes of the pollen parent. Classical genetic and nucleotide-level analyses of the *S*-locus have revealed a dichotomy in sequence and function among *S*-haplotypes; in particular, all class I haplotypes show dominance over all class II haplotypes in determination of pollen specificity. Analysis of an evolutionary model that explicitly incorporates features of the Brassica system, including the class dichotomy, indicates that class II haplotypes may invade populations at lower rates and decline to extinction at higher rates than class I haplotypes. This analysis suggests convergence to an evolutionarily persistent state characterized by the maintenance in high frequency of a single class II haplotype together with many class I haplotypes, each in low frequency. This expectation appears to be consistent with empirical observations of high frequencies of relatively few distinct recessive haplotypes.

FLOWERING plants exhibit a variety of mechanisms that discourage self-fertilization. In a number of self-incompatibility (SI) systems, regulation of outcrossing occurs prezygotically under the control of genetic factors that segregate as a single locus in the classical sense (see de Nettancourt 1977). In heteromorphic SI systems, this *S*-locus determines floral morphology, which promotes fertilization between but not within morphs. In homomorphic systems, in which the mating classes have similar floral morphologies, the *S*-locus encodes specificities expressed by pollen and recognized on the stigma or in the style of the plant that receives it. Expression by pollen of a specificity held in common with the seed parent induces a self-incompatibility response, which prevents fertilization. Pollen specificity is determined sporophytically (by the *S*-locus genotype of the pollen parent) under sporophytic SI (SSI), and gametophytically [by the *S*-allele(s) carried in the pollen genome] under gametophytic SI (GSI).

Evolution of homomorphic SI: Genetic characterization of components controlling SSI in Brassica (Nasrallah *et al.* 1985) and GSI in the Solanaceae, Rosaceae, Scrophulariaceae (Anderson *et al.* 1986; Broothaerts *et al.* 1995; Xue *et al.* 1996), and the field poppy *Papaver rhoeas* (Foote *et al.* 1994) has revealed a remarkable evolutionary convergence of distantly related taxa to one-locus regulation of various physiological mechanisms of SI. Compared to other loci, a very large number of *S*-alleles, each encoding a distinct specificity, typically segregate at the *S*-locus, although empirical estimates may fall short of theoretical expectation under SI (Lawrence 2000). Sequence comparisons across species or genera indicate the maintenance of *S*-allele lineages over tens of millions of years (Ioerger *et al.* 1990; Dwyer *et al.* 1991; Hinata *et al.* 1995; Uyenoyama 1995). Both the high allelic polymorphism and the persistence of ancient lineages appear to reflect the strong balancing selection engendered by the transmission advantage accruing to pollen that expresses rare specificities (Wright 1939; Vekemans and Slatkin 1994).

Sporophytic self-incompatibility in Brassica: Classical genetic studies of the form of SSI expressed in Brassica revealed a complex pattern of dominance in both pollen and stigma expression (Bateman 1952; Thompson and Taylor 1966; Ockendon 1974). Ockendon's (1974) analysis of 16 cultivars of Brussels sprouts (*Brassica oleracea* var. *gemmifera*) identified 19 *S*-alleles and documented considerable inequality in frequency. The most recessive *S*-alleles were the most common, with S_2 and S_5 observed in high frequencies in all cultivars studied (Thompson and Taylor 1966; Ockendon 1974).

By stripping *B. oleracea* pollen grains of their coating and interposing extracts from pollen coatings of the same or different specificities between the stripped pollen and the stigmatic surface, Stephenson *et al.* (1997) demonstrated that proteins located in the pollen coating determine pollen specificity. Under sporophytic determination of specificity, pollen grains may bear proteins encoded by one or both *S*-alleles held by the pollen parent. The mechanism of dominance in the expression of pollen specificity in Brassica remains a prime outstanding question.

Address for correspondence: Department of Zoology, Box 90325, Duke University, Durham, NC 27708-0325. E-mail: marcy@duke.edu

Nucleotide-level analyses of the *S*-locus region have revealed considerable structural complexity and a high density of expressed genes (Boyes *et al.* 1997; Suzuki *et al.* 1999). *S*-haplotypes include the *SLG* (*S*-locus glycoprotein), *SRK* (*S*-locus receptor kinase), and *SCR* (*S*-locus cysteine-rich protein) genes (Schopfer *et al.* 1999). The role of *SLG* in SSI has not yet been resolved, with recent work challenging the view that *SLG* directly determines *S*-specificity (Kusaba and Nishio 1999; Okazaki *et al.* 1999; Kusaba *et al.* 2000). *SRK* encodes a membrane-bound receptor kinase, the extracellular domain of which shows greater sequence similarity to *SLG* of the same haplotype than to homologous extracellular domains of different haplotypes (Stein *et al.* 1991). Schopfer *et al.* (1999) have proposed that *SCR* encodes secreted proteins that become incorporated into the pollen coat, from where they translocate into stigma epidermal cells, the site of highly specific binding with the corresponding *SRK* receptor.

Dominance in expression of pollen specificity: Analysis of *SLG* and *SRK* revealed the existence of two genetic classes, defined by sequence differences (Chen and Nasrallah 1990) and correlated with functional differences, including dominance in pollen expression of class I over class II (Nasrallah *et al.* 1991). A phylogenetic analysis of *SLG*, *SRK*, and homologous sequences yielded generalized least-squares estimates of divergence times of ~25 million years (MY) within class I, 7 MY within class II, and 40 MY between class I and class II (Uyenoyama 1995), indicating that the rate of mutation within class exceeds the rate between classes.

Okazaki *et al.* (1999) used DNA gel-blot analysis to characterize *SLG* genes in the *B. oleracea* *S*-tester collection, originally established by K. F. Thompson at Cambridge and maintained and expanded at the National Vegetable Research Station at Wellesbourne, Warwick, United Kingdom. This study revealed that of the 43 *S*-tester lines examined, 40 carried class I alleles, with the remaining 3 bearing class II alleles S_2, S_5, and S_{15}, which had been shown to be recessive in pollen expression relative to all other alleles (Thompson and Taylor 1966). Okazaki *et al.* (1999) noted previous observations of both the class I/class II partition of *S*-alleles and higher numbers of class I alleles in other species in the Brassicaceae, *B. campestris* (18 class I, 6 class II; Nishio *et al.* 1996) and *Raphanus sativus* (13 class I, 5 class II; Sakamoto *et al.* 1998).

Evolutionary models of SSI: Bateman (1952) formulated a number of models of self-incompatibility. Cope (1962) analyzed the dynamics of genotypic change under the serial dominance scheme, characterized by an ordering of alleles such that for $i < j$, S_i shows complete dominance in both pollen and stigma expression over S_j. Thompson and Taylor's (1966) characterization of dominance relationships among *S*-alleles in kale departs in a number of respects from this model. Contrary to the serial dominance scheme, which assigns each allele to a distinct dominance level and assumes identical dominance hierarchies for pollen and stigma expression, Thompson and Taylor (1966) described the sharing of dominance levels by a number of *S*-alleles. Further, they reported that the relative dominance of a given pair of alleles may differ in stigma and pollen expression, with fewer instances observed of dominance in stigma than in pollen expression. Sampson's (1974) numerical iteration model improved the approximation to the Brassica system by permitting multiple *S*-alleles to share a given level of dominance and assuming codominant expression in stigma.

Schierup *et al.* (1997, 1998) and Vekemans *et al.* (1998) have begun the exploration of the stochastic evolution of allele number and frequency under SSI through numerical simulation models. Not designed to describe the Brassica system, these models incorporate the serial dominance scheme, with each allele assigned a distinct dominance level, and assume either identical dominance hierarchies in pollen and stigma expression or codominance in stigma. A new *S*-allele arising by mutation assumes either a random position in the dominance hierarchy or a position similar to that of its parent allele. More dominant *S*-alleles were found to invade with higher probability, resulting in a progressive increase in dominance over evolutionary time.

In this article, I present an evolutionary model that explicitly incorporates features of the Brassica system of SSI, including the class I/II dichotomy among *S*-alleles. This model specifies dominance in pollen expression of class I over class II, codominance in pollen expression within class, and codominance in stigma expression between all alleles, both within and between classes. My results indicate that dominance in pollen expression promotes the increase of rare class I *S*-alleles and diminishes their rate of random loss. While the increase of rare class I alleles is always favored, deterministic pressures oppose a return to higher levels of class II alleles once the number falls to one. A single class II allele maintained in high frequency together with many class I alleles appears to correspond to an evolutionarily persistent state. These findings suggest that the observation in *B. oleracea* of high frequencies of relatively few class II alleles reflects the dominance relationship in pollen expression between the classes.

EVOLUTIONARY MODEL

Fundamental recursions: As noted by Fisher (1941), differential fertilization by pollen constitutes the sole selective force inherent in the expression of self-incompatibility. Half of the gene pool of each generation is transmitted through egg and half through pollen. Under Mendelian segregation of mating-type alleles, each gene has an expectation of transmitting one descendant gene to the offspring generation through egg, while the expectation through pollen depends on access to

compatible mates. Let q_i denote the frequency of the ith mating-type allele; in the next generation it is

$$q_i' = \frac{1}{2}q_i + \frac{1}{2}\sum_j t_{ij}P_j, \quad (1)$$

in which t_{ij} denotes the rate of production of pollen bearing the ith allele by the jth phenotypic class ($q_i = \Sigma_j\, t_{ij}$) and P_j the fertilization success of the jth class.

This expression indicates that equilibrium entails, for each mating-type gene, equal rates of transmission through pollen and egg and, consequently, equal transmission of all mating-type genes:

$$\frac{\Sigma_j t_{ij}P_j}{\Sigma_j t_{ij}} = 1 \quad \text{for all } i. \quad (2)$$

Model formulation: With respect to pollen specificity, the model specifies dominance of all class I alleles over all class II alleles and codominance within class and, with respect to stigma phenotype, codominance both within and between classes. Representing class I alleles by superscripts and class II by subscripts, I classify genotypes into four groups: homozygotes bearing one class II allele (G_{ii}), heterozygotes bearing two class I alleles (G^{ij}), two class II alleles (G_{ij}), or one allele of each class (G^i_j). Inspection of the full set of recursions in genotypic frequencies indicates that zygote formation returns populations initiated on the symmetric surface (equal frequencies among genotypes within groups) to that surface. Under the assumption that populations initiated at arbitrary positions off the symmetric surface converge to it, the subsequent description of the population can be reduced to four variables, H, G, R, and D, representing the frequency of any particular genotype in the respective four groups. Under these assumptions, the frequencies of each of n_1 class I (a) and each of n_2 class II (b) alleles are

$$a = G(n_1 - 1)/2 + Dn_2/2 \quad (3)$$

$$b = H + R(n_2 - 1)/2 + Dn_1/2. \quad (4)$$

Genotypic frequencies necessarily sum to unity:

$$Hn_2 + G\binom{n_1}{2} + R\binom{n_2}{2} + Dn_1n_2 = 1. \quad (5)$$

Recursions in H, G, R, and D (given in the APPENDIX) confirm the general expression for the change in gene frequency (1),

$$a' = \frac{1}{2}a + \frac{1}{2}[G(n_1 - 1)P_G/2 + Dn_2P_D/2] \quad (6)$$

$$b' = \frac{1}{2}b + \frac{1}{2}[HP_H + R(n_2 - 1)P_R/2 + Dn_1P_D/2], \quad (7)$$

in which P_G, P_D, P_H, and P_R represent the rates of transmission through pollen produced by the four phenotypic groups [P_j in (1)]. The frequencies in the population of compatible stigmas and the load of compatible pollen those stigmas receive from all sources determine the rates of transmission through pollen,

$$P_G = \frac{Hn_2}{1 - y} + \frac{R\binom{n_2}{2}}{1 - 2y + R} + \frac{D(n_1 - 2)n_2}{1 - x - y} + \frac{G\binom{n_1 - 2}{2}}{1 - 2x + G} \quad (8)$$

$$P_D = \frac{Hn_2}{1 - y} + \frac{R\binom{n_2}{2}}{1 - 2y + R} + \frac{D(n_1 - 1)n_2}{1 - x - y} + \frac{G\binom{n_1 - 1}{2}}{1 - 2x + G} \quad (9)$$

$$P_H = \frac{H(n_2 - 1)}{1 - y} + \frac{R\binom{n_2 - 1}{2}}{1 - 2y + R} + \frac{Dn_1(n_2 - 1)}{1 - x - y} + \frac{G\binom{n_1}{2}}{1 - 2x + G} \quad (10)$$

$$P_R = \frac{H(n_2 - 2)}{1 - y} + \frac{R\binom{n_2 - 2}{2}}{1 - 2y + R} + \frac{Dn_1(n_2 - 2)}{1 - x - y} + \frac{G\binom{n_1}{2}}{1 - 2x + G}, \quad (11)$$

in which each denominator represents the frequency of compatible pollen received by the genotype indicated in the numerator and x and y denote the frequencies of pollen that express any given class I and any given class II specificity:

$$x = G(n_1 - 1) + Dn_2 \quad (12)$$

$$y = H + R(n_2 - 1). \quad (13)$$

RESULTS

Equilibrium state: *A necessary condition for equilibrium:* At equilibrium, the rates of transmission of each S-allele through egg and pollen are equal [see (2)]:

$$1 = \frac{G(n_1 - 1)P_G + Dn_2P_D}{G(n_1 - 1) + Dn_2} \quad (14)$$

$$1 = \frac{HP_H + R(n_2 - 1)P_R/2 + Dn_1P_D/2}{H + R(n_2 - 1)/2 + Dn_1/2}. \quad (15)$$

For populations in which at least two class I alleles exist ($n_1 \geq 2$), P_D exceeds P_G [compare (8) and (9)], which implies [through (14)] that at equilibrium the rate of transmission through pollen grains that express only a single class I specificity exceeds the rate through egg, which in turn exceeds the rate through pollen grains that express two class I specificities ($P_D > 1 > P_G$). Condition (14) indicates that P_D equals unity only if a single class I allele exists. Similarly, in populations that maintain at least two class II alleles ($n_2 \geq 2$), the rate of transmission of pollen grains that express only a single class II specificity exceeds that of grains that express two class II specificities ($P_H > P_R$). In populations that maintain at least two class I alleles (for which $P_D > 1$), (15) indicates that the rate of transmission through pollen expressing two class II specificities falls below the rate through egg ($1 > P_R$). At equilibrium, P_H may assume values both above and below unity.

Deterministic dynamics: *Fate of a rare class I allele:* Carriers of a rare class I allele may bear in addition either a common class I allele or a common class II allele. Let

ε_G and ε_D, respectively, represent the total frequencies of these two kinds of carriers. Linearized recursions in these variables appear in the APPENDIX.

For populations containing at least three common alleles, including one from each class, the frequency among offspring of the rare allele in the next generation is

$$\varepsilon'_G + \varepsilon'_D = \frac{1}{2}(\varepsilon_G + \varepsilon_D) + \frac{1}{2}(\varepsilon_G P_D + \varepsilon_D P_*) \qquad (16)$$

[compare (1)], in which P_D is given by (9) and P_* represents the transmission success of a pollen grain that, because it expresses only the rare specificity, is compatible with all common genotypes:

$$P_* = \frac{Hn_2}{1 - y} + \frac{G\binom{n_1}{2}}{1 - 2x + G} + \frac{R\binom{n_2}{2}}{1 - 2y + R} + \frac{Dn_1n_2}{1 - x - y}. \qquad (17)$$

P_* always exceeds unity [see (5)]. From (16), the rate of transmission through egg of the rare allele is unity (as is the case for all alleles), while the rate through pollen is an average of the rates associated with pollen grains that express a single common class I specificity (P_D) and no common specificities (P_*). Because P_D is greater than or equal to unity and P_* exceeds unity, rare class I alleles always increase at a geometric rate.

In the absence of common class II alleles ($n_2 = 0$), reproduction requires at least four common class I alleles ($n_1 \geq 4$). Pollen grains bearing a rare class I allele express both the rare specificity and one common class I specificity. This type of pollen is compatible with $\binom{n_1}{2}^{-1}$ common stigmas, each of which receives compatible pollen from $\binom{n_1}{2}^{-2}$ common genotypes; the relative rate of fertilization of pollen carrying the rare class I allele is the ratio. Consequently, rare class I alleles increase at a geometric rate:

$$\varepsilon'_G = \frac{1}{2}\varepsilon_G\left[1 + \frac{n_1 - 1}{n_1 - 3}\right]. \qquad (18)$$

Similarly, in the absence of common class I alleles ($n_1 = 0$), ε_G is zero and a minimum of four common class II alleles must reside in the population ($n_2 \geq 4$). Complete dominance of class I over class II implies that pollen produced by carriers of a rare class I allele would be compatible with all common stigmas. Each of $\binom{n_2}{2}$ common stigmas accepts pollen from $\binom{n_2}{2}^{-2}$ common genotypes. The high transmission success of pollen carrying the rare class I allele permits its invasion:

$$\varepsilon'_D = \frac{1}{2}\varepsilon_D\left[1 + \frac{n_2(n_2 - 1)}{(n_2 - 2)(n_2 - 3)}\right]. \qquad (19)$$

Nongeneric behavior arises in populations that contain a single common class I allele (S_1) and a single common class II allele (S_2). Only two common genotypes occur in this population, with S_1S_2 stigmas excluding pollen from both genotypes and S_2S_2 pollen incompatible with both stigmas. Equal frequencies of the two genotypes are maintained in the population through fertilization of S_2S_2 stigmas by pollen from S_1S_2. Unlike all other cases, a rare class I allele (S_3) introduced into this population becomes common immediately upon the formation of the genotype carrying it together with the common class II allele (S_2S_3), because pollen from only this genotype can fertilize the common genotype S_1S_2.

These results indicate that rare class I alleles are uniformly protected from loss, irrespective of the number or class of common alleles in the population.

Fate of a rare class II allele: Carriers of a rare class II allele bear in addition either a common class I allele or a common class II allele; let ε_D and ε_R represent their frequencies. The APPENDIX presents linearized recursions in these variables.

Class II alleles in the absence of common class I alleles ($n_1 = 0$) behave identically to class I alleles in the absence of class II. In particular, rare class II alleles increase at a geometric rate that depends on the number of compatible stigmas encountered by pollen carrying those alleles,

$$\varepsilon'_R = \frac{1}{2}\varepsilon_R\left[1 + \frac{n_2 - 1}{n_2 - 3}\right], \qquad (20)$$

in which the number of common class II alleles (n_2) must exceed four to permit reproduction [compare (18)].

In populations that lack common class II alleles ($n_2 = 0$), at least four class I alleles must exist ($n_1 \geq 4$) and pollen grains bearing a rare class II allele express only one common class I specificity. As before, this kind of pollen is compatible with $\binom{n_1}{2}^{-1}$ common stigmas, each of which receive compatible pollen from $\binom{n_1}{2}^{-2}$ common genotypes, ensuring the increase of the rare class II allele,

$$\varepsilon'_D = \frac{1}{2}\varepsilon_D\left[1 + \frac{n_1 - 1}{n_1 - 3}\right] \qquad (21)$$

[compare (19)].

In populations containing at least one common allele of each class, the frequency of a rare class II allele in the next generation is

$$\varepsilon'_D + \varepsilon'_R = \frac{1}{2}(\varepsilon_D + \varepsilon_R) + \frac{1}{2}(\varepsilon_D P_D + \varepsilon_R P_H) \qquad (22)$$

[compare (1)]. As is the case for any allele, the rate of transmission of the rare allele through egg is unity [compare (1) and (16)]. Pollen grains that bear the rare class II allele either express a single class I specificity and achieve fertilization at rate P_D or express a single common class II specificity and achieve fertilization at

rate P_H. Because P_D is greater than or equal to unity, with equality only if a single common class I allele exists, (22) indicates that P_H greater than unity provides a sufficient condition for the increase of a rare class II allele.

In the nongeneric case of a population containing a single common allele of each class, the introduction of a rare, additional class II allele always succeeds. The genotype carrying both common alleles (S_1S_2) rejects all pollen except that produced by an individual carrying the new class II allele in homozygous form. Immediately upon the formation of such extremely rare individuals, the new class II allele becomes common as a consequence of its very high rate of transmission through pollen on S_1S_2 stigmas.

In all other cases involving a single common class II allele ($n_1 \geq 2$, $n_2 = 1$), individuals carrying two common class II alleles are absent ($R = 0$) and P_D exceeds unity; consequently, (15) indicates that P_H is strictly less than unity. Analysis of the linearized recursions for this case shows that rare class II alleles fail to increase.

For populations with only a single common class I allele and two or more common class II alleles ($n_1 = 1$, $n_2 \geq 2$), P_D equals unity [see (14)] and P_H exceeds P_R [see (10) and (11)]. The equilibrium condition (15) that an average among P_D, P_H, and P_R equals unity ensures that P_H exceeds unity, satisfying the sufficient condition for increase of a rare class II allele [see (22)].

To study the fate of a rare class II allele in the remaining cases ($n_1 > 1$, $n_2 > 2$), I conducted numerical iterations of the full system of recursions, incorporating all possible genotypes in possibly asymmetric frequencies for populations containing every valid combination of 0 to 50 common class I and class II alleles. In every case, the population converged rapidly to the symmetric state characterized by equal frequencies of alleles within each class and equal frequencies within each of the four phenotypic groups (G_{ii}, G^{ij}, G_{ij}, and G^{i}_{j}). These numerical iterations confirmed the analytical finding that rare class II alleles fail to increase in populations containing one common class II allele and two or more common class I alleles and indicated their successful invasion in all other cases.

These results indicate that a rare class II allele always increases in frequency, with the exception of cases in which its loss would leave a single common class II allele together with two or more class I alleles ($n_1 \geq 2$; $n_2 = 1$). While this finding would suggest that exactly two common class II alleles cannot coexist (except in the nongeneric case with only a single common class I allele), an equilibrium state comprising two class II alleles in equal frequencies together with any number of class I alleles does in fact exist. Numerical iterations indicate convergence to this state in populations initiated at random points on the symmetric surface; however, any perturbation of the population from this surface results in the rapid extinction of one of the class II alleles.

Effect of dominance in pollen expression within class II: A rare class II allele that shows recessivity in pollen expression relative to common class II alleles showing codominance among themselves declines in frequency only if a single common class II allele exists and increases otherwise. This condition is identical to that for a rare codominant class II allele. Codominant or recessive expression of the rare specificity by pollen grains has equivalent effects because in both cases grains carrying the rare allele induce incompatibility only on stigmas that recognize the common specificity they express.

In contrast, pollen grains carrying a rare class II allele that shows dominance in pollen expression over common class II alleles induce incompatibility only on stigmas that also carry the rare allele. Because pollen produced by rare carriers encounters stigmas of rare carriers at negligible rates, the frequency of such an allele in the next generation becomes

$$\varepsilon'_D + \varepsilon'_R = \frac{1}{2}(\varepsilon_D + \varepsilon_R) + \frac{1}{2}(\varepsilon_D P_D + \varepsilon_R P_*) \qquad (23)$$

[compare (22)]. Because both P_D and P_* exceed unity, a class II allele that expresses dominance in pollen expression over other class II alleles increases when rare, even in populations that contain only a single common class II allele. In the latter case, invasion is expected to cause the replacement of the initially common allele, returning the number of class II alleles in the population to unity. A single class II allele maintained together with any number of class I alleles excludes rare class II alleles over which it shows dominance or codominance in pollen expression and is replaced by rare class II alleles against which it is recessive.

Stochastic dynamics: In populations of finite size, the effective population size and the relative rates of origin and extinction of class I and class II *S*-alleles determine the numbers maintained in a population. Results of the deterministic analysis concerning the fate of rare class II alleles suggest that the number of class II alleles is unlikely to return to high levels once it falls below three. Indeed, preliminary results from Monte Carlo simulations indicate that under equal mutation rates to new class I and class II alleles, the number of class II alleles rapidly declines to one while many class I alleles persist.

To explore the greater vulnerability of class II alleles relative to class I alleles, I compared the rates of increase of rare *S*-alleles. A population was initiated in state (n_1; n_2) (n_1 common class I and n_2 common class II alleles) with random allele frequencies, and the deterministic recursions iterated to equilibrium. Equilibrium genotypic frequencies were used to determine the dominant eigenvalue of the local stability matrix describing the initial increase of a class I or a class II allele.

A rare class I allele increases in the population at a higher rate than a rare class II allele, unless no common class II alleles initially exist. In the exceptional case, rare alleles of the two classes increase at the same rate; this

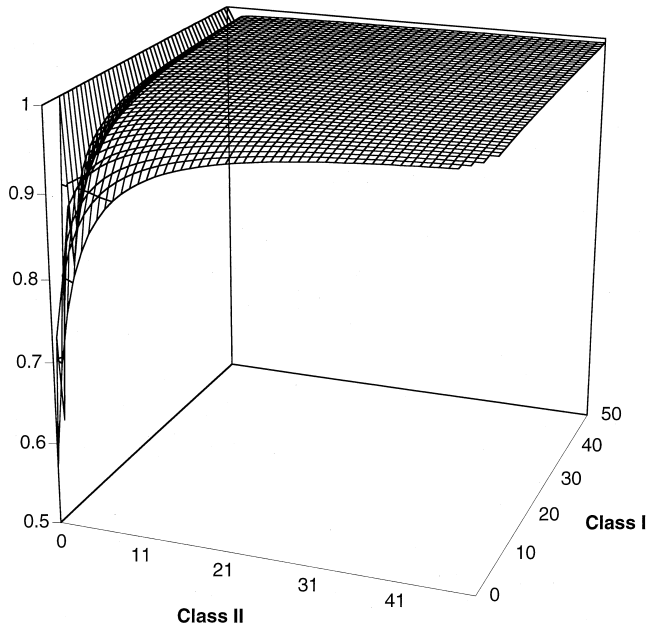

FIGURE 1.—Rate of increase of a rare class II allele relative to a rare class I allele in a population containing all combinations of class I and class II alleles between 0 and 50. The ridge at the boundary corresponding to 0 class II alleles [state ($n_1 \geq 4$; $n_2 = 0$)] indicates that in the absence of common class II alleles, rare alleles of both classes increase at the same rate. With the exception of this case, class I alleles introduced in low frequency always increase at higher rates.

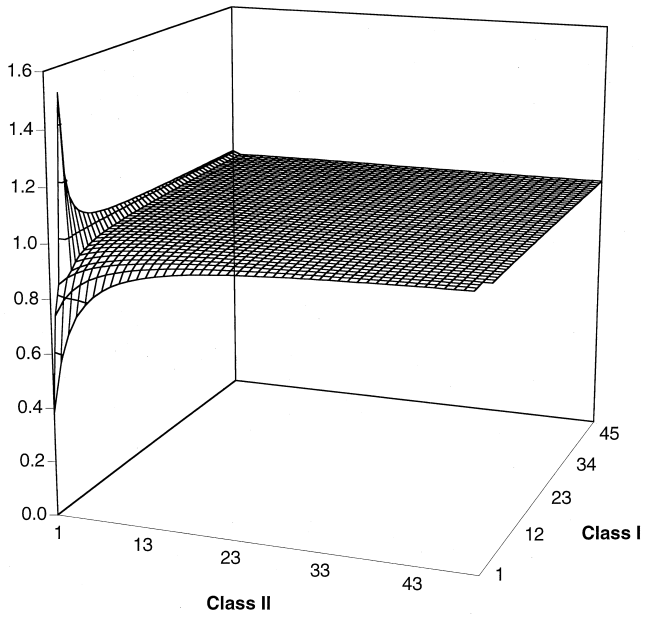

FIGURE 2.—Index of the relative rate of return of class II alleles that have become rare, corresponding to the rate of increase of a rare class II allele at state (n_1; $n_2 - 1$) relative to that of a rare class I allele at state ($n_1 - 1$; n_2). class II alleles appear more vulnerable to extinction except in the case of the last class II allele in a population maintaining multiple class I alleles.

rate depends only on the number of common (class I) alleles, because the dominance of the rare allele in pollen expression relative to the common alleles has no effect on the frequency with which pollen bearing the rare allele encounters incompatible stigmas. Figure 1 shows the relative rate of increase of a rare class II allele relative to a rare class I allele in a population with n_1 common class I alleles and n_2 common class II alleles. Given the appearance of a new *S*-allele, it invades at a higher rate if it belongs to class I.

I also examined the relative rate of return to higher frequencies of an *S*-allele, given that it has drifted to low frequencies. Upon the loss of one common *S*-allele, a population in state (n_1; n_2) passes either to state ($n_1 - 1$; n_2) or to state (n_1; $n_2 - 1$). I compared the rate of increase of a rare class II allele in state (n_1; $n_2 - 1$) to that of a rare class I in state ($n_1 - 1$; n_2). Figure 2 indicates that this index of the relative rate of return is less than unity for all valid values of n_1 and n_2 up to 50, except for the case in which a single class II remains [state (n_1; $n_2 = 1$)]. This comparison suggests that in all but one case, the loss of an *S*-allele, once it has become rare, is more likely if it belongs to class II. In the exceptional case, a single class II allele maintained in a population with multiple class I alleles increases from low frequencies at higher rates than any class I allele.

This comparison of deterministic rates of initial increase suggests that class II alleles are more prone to loss from all states, except the one that maintains a single class II allele. Preliminary stochastic numerical simulations indicate that over evolutionary time the number of class I alleles increases to a limit determined by the population size and the rate of mutation to new class I alleles, while the number of class II alleles declines to one per population. Figure 3 shows the expected frequencies, obtained from iteration of the deterministic recursions, of class I and class II alleles in a population in state (n_1; $n_2 = 1$).

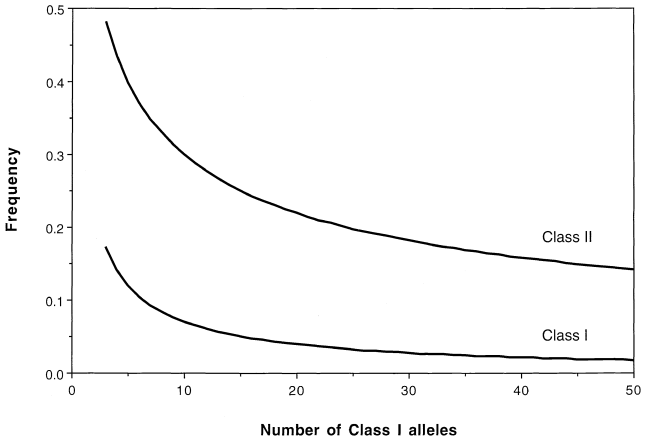

FIGURE 3.—Deterministic expected frequencies of a single class I or class II allele in a population in which multiple class I and one class II alleles segregate.

DISCUSSION

Fundamental selective force: Differential fertilization success constitutes the sole selective force in the deterministic model of self-incompatibility analyzed here. Equation (1) embodies the fundamental principle, which applies to all forms of mating incompatibilities under single-locus control: each gene transmits an average of one gene through female gametes, while the rate of transmission through male gametes (P_j) depends on the frequency of compatible genotypes among potential mates. A necessary condition for equilibrium is that the expected rate of transmission be identical among all alleles, with one gene on average transmitted through male as well as through female gametes [see (2)].

A particular class I allele occurring in a genome together with another class I allele will be borne by pollen grains that express two specificities (type *G*), while the same allele occurring with a class II allele will be borne by pollen grains that express only one (type *D*). Because grains that express fewer specificities induce incompatibility in fewer stigmas, type *D* pollen achieves greater fertilization success than type *G* pollen [$P_D > P_G$ in (11)]. The equilibrium condition that the average rate of transmission of all *S*-alleles through pollen converge to unity determines the equilibrium frequencies of carriers of class I alleles [(14) with $P_D > 1 > P_G$]. Similarly, pollen grains produced by an individual carrying a given class II allele in homozygous form (type *H*) express one specificity and achieve greater transmission than grains produced by an individual carrying the class II allele together with another class II allele (type *R*), which express two ($P_H > P_R$).

Rates of invasion: Differential fertilization success strongly influences the rate of increase of rare *S*-alleles. Pollen grains bearing a rare class I allele express one common specificity if produced by a type *G* genotype and no common specificity if by type *D*. Grains bearing a rare class II allele always express one common specificity: a class I specificity if produced by type *D* and class II by type *R*. Consequently, higher pollen transmission success accrues to rare alleles that belong to class I.

S-alleles of both classes increase when rare, with one exception: exactly two class II alleles cannot coexist in a population containing more than one common class I allele unless the frequencies of the class II alleles never depart from equality. Sampson (1974) noted similar behavior in a model that shares some features, but considered it of "minor importance" (p. 617). Nevertheless, this case applies to a major system of self-incompatibility, SSI in Brassica.

Rates of random extinction: While characterization of the stochastic dynamics requires a full treatment that incorporates mutation and genetic drift, my deterministic study of initial increase conditions suggests that class II alleles are more prone to random extinction. Class I alleles that become rare generally increase at higher rates, with the exception of a single class II allele segregating with multiple class I alleles (Figure 2). This analysis indicates that any but the last class II allele in a population is more prone to extinction than a class I allele.

Preliminary stochastic simulations indicate a progressive loss of class II alleles while the number of class I alleles increases to high levels. Once the number of class II alleles falls below three, deterministic forces oppose a return to higher levels. These considerations suggest that a single class II allele segregating together with many class I alleles represents an evolutionarily persistent state.

In summarizing the equilibrium frequency distribution of his model, Sampson (1974) described a "recessive effect" and a "small number effect": alleles with more recessive pollen expression and those that share a given dominance level with fewer alleles tend to occur in higher frequencies. In accordance with these qualitative trends, my analysis suggests that the maintenance in high frequency of the single class II allele constitutes an evolutionarily persistent state.

Figure 3 shows the frequencies of each class I and class II allele in a population containing a single class II allele and several class I alleles. The high deterministic rate of increase of a single rare class II allele in a population with more than one class I alleles protects it from loss (Figure 2). Further, the high equilibrium frequency of the single class II allele suggests that it is less likely than a class I allele to drift to extinction. Because a deterministic analysis can provide only a rough indication of the stochastic dynamics, these conjectures will be explored in a separate article.

Comparison to Brassica cultivars: In addition to the major trend of dominance in pollen expression of class I over class II alleles, cultivars of *B. oleracea* exhibit other dominance interactions within and between class. Thompson and Taylor (1966) observed dominance in stigma expression somewhat less commonly than in pollen, with given pairs of alleles showing different dominance relationships in pollen and stigma expression. They also described "nonlinear" (nontransitive) dominance relationships among alleles within the two major groups now identified as class I and class II.

Schierup *et al.* (1997, 1998) and Vekemans *et al.* (1998) conducted stochastic numerical simulations, assuming either identical dominance expression in pollen and stigma expression or complete codominance. They studied a form of SSI distinct from the Brassica system: a serial dominance scheme for pollen expression with a maximum of one allele at any dominance level (compare Cope 1962). Schierup *et al.* (1997) found that the probability that a new *S*-allele will increase when rare or that an existing *S*-allele will become extinct depends on position in the dominance hierarchy, with dominant alleles tending to invade and recessive alleles tending to become extinct. As a consequence, the population

proceeds up the dominance hierarchy through evolutionary time. Dominance among class I alleles within the Brassica system of SSI may evolve in a similar manner.

I explored the effect of dominance in pollen expression on the invasion of a novel class II allele into a population that maintains a single common class II allele, preserving the assumption of strict codominance in stigma expression between as well as within class. While codominant rare alleles always fail to invade the population, rare alleles that express dominance over the common class II allele succeed in replacing it (23).

The qualitative expectation of the maintenance in high frequencies of few class II alleles is consistent with empirical observations in cultivars of *B. oleracea* (Thompson and Taylor 1966; Ockendon 1974). Further, those studies indicated that the relative frequencies of class II alleles ($S_2 > S_5 > S_{15}$ in Table 2 of Ockendon 1974) conform to their relative dominance in pollen expression ($S_2 > S_5 > S_{15}$ in Figure 3 of Thompson and Taylor 1966). This correspondence may reflect the tendency, even within class II, of more dominant to replace less dominant *S*-alleles.

I thank M. Kusaba and M. J. Lawrence for providing manuscripts prior to publication, Ruth Shaw and two anonymous reviewers for constructive comments and suggestions, and D. Liu for assistance. U.S. Public Health Service grant GM 37841 provided support for this study.

LITERATURE CITED

Anderson, M. A., E. C. Cornish, S.-L. Mau, E. G. Williams, R. Hoggart *et al.*, 1986 Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature **321:** 38–44.

Bateman, A. J., 1952 Self-incompatibility in angiosperms: I. Theory. Heredity **6:** 285–310.

Boyes, D. C., M. E. Nasrallah, J. Vrebalov and J. B. Nasrallah, 1997 The self-incompatibility (*S*) haplotypes of Brassica contain highly divergent and rearranged sequences of ancient origin. Plant Cell **9:** 237–247.

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 Mol. Biol. **27:** 499–511.

Chen, C.-H., and J. B. Nasrallah, 1990 A new class of *S* sequences defined by a pollen recessive self-incompatibility allele of *Brassica oleracea*. Mol. Gen. Genet. **222:** 241–248.

Cope, F. W., 1962 The effects of incompatibility and compatibility on genotype proportions in populations of *Theobroma cacao* L. Heredity **17:** 183–195.

de Nettancourt, D., 1977 *Incompatibility in Angiosperms*. Springer-Verlag, Berlin.

Dwyer, K. G., M. A. Balent, J. B. Nasrallah and M. E. Nasrallah, 1991 DNA sequences of self-incompatibility genes from *Brassica campestris* and *B. oleracea*: polymorphism predating speciation. Plant Mol. Biol. **16:** 481–486.

Fisher, R. A., 1941 The theoretical consequences of polyploid inheritance for the mid style form of *Lythrum salicaria*. Ann. Eugen. **11:** 31–38.

Foote, H. C. C., J. P. Ride, V. E. Franklin-Tong, E. A. Walker, M. J. Lawrence *et al.*, 1994 Cloning and expression of a distinctive class of self-incompatibility (S) gene from *Papaver rhoeas* L. Proc. Natl. Acad. Sci. USA **91:** 2265–2269.

Hinata, K., M. Watanabe, S. Yamakawa, Y. Satta and A. Isogai, 1995 Evolutionary aspects of the *S*-related genes of the Brassica self-incompatibility system: synonymous and nonsynonymous base substitutions. Genetics **140:** 1099–1104.

Ioerger, T. R., A. G. Clark and T.-h. Kao, 1990 Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. Proc. Natl. Acad. Sci. USA **87:** 9732–9735.

Kusaba, M., and T. Nishio, 1999 Comparative analysis of *S* haplotypes with very similar *SLG* alleles in *Brassica rapa* and *B. oleracea*. Plant J. **17:** 83–92.

Kusaba, M., M. Matsushita, K. Okazaki, Y. Satta and T. Nishio, 2000 Sequence and structural diversity of the *S* locus genes from different lines with the same self-recognition specificities in *Brassica oleracea*. Genetics **154:** 413–420.

Lawrence, M. J., 2000 Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. Ann. Bot. **85a:** 221–226.

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.

Nasrallah, J. B., T. Nishio and M. E. Nasrallah, 1991 The self-incompatibility genes of Brassica: expression and use in genetic ablation of floral tissues. Annu. Rev. Plant Physiol. Plant Mol. Biol. **42:** 393–422.

Nishio, T., M. Kusaba, M. Watanabe and K. Hinata, 1996 Registration of *S*-alleles in *Brassica campestris* L. by the restriction fragment sizes of *SLG*s. Theor. Appl. Genet. **92:** 388–394.

Ockendon, D. J., 1974 Distribution of self-incompatibility alleles and breeding structure of open-pollinated cultivars of Brussels sprouts. Heredity **33:** 159–171.

Okazaki, K., M. Kusaba, D. J. Ockendon and T. Nishio, 1999 Characterization of *S* tester lines in *Brassica oleracea*: polymorphism of restriction fragment length of *SLG* homologues and isoelectric points of *S*-locus glycoproteins. Theor. Appl. Genet. **98:** 1329–1334.

Sakamoto, K., M. Kusaba and T. Nishio, 1998 Polymorphism of the its-locus glycoprotein gene (*SLG*) and the *S*-locus related gene (*SLR1*) in *Raphanus sativus* L. and self-incompatible ornamental plants in the Brassicaceae. Mol. Gen. Genet. **258:** 397–403.

Sampson, D. R., 1974 Equilibrium frequencies of sporophytic self-incompatibility alleles. Can. J. Genet. Cytol. **16:** 611–618.

Schierup, M. H., X. Vekemans and F. B. Christiansen, 1997 Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. Genetics **147:** 835–846.

Schierup, M. H., X. Vekemans and F. B. Christiansen, 1998 Allelic genealogies in sporophytic self-incompatibility systems in plants. Genetics **150:** 1187–1198.

Schopfer, C. R., M. E. Nasrallah and J. B. Nasrallah, 1999 The male determinant of self-incompatibility in *Brassica*. Science **286:** 1697–1700.

Stein, J. C., B. Howlett, D. C. Boyes, M. E. Nasrallah and J. B. Nasrallah, 1991 Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc. Natl. Acad. Sci. USA **88:** 8816–8820.

Stephenson, A. G., J. Doughty, S. Dixon, C. Elleman, S. Hiscock *et al.*, 1997 The male determinant of self-incompatibility in *Brassica oleracea* is located in the pollen coating. Plant J. **12:** 1351–1359.

Suzuki, G., N. Taki, T. Hirose, K. Fukui, T. Nishio *et al.*, 1999 Genomic organization of the *S* locus: identification and characterization of genes in *SLG/SRK* region of *S9* haplotype of *Brassica campestris* (syn. *rapa*). Genetics **153:** 391–400.

Thompson, K. F., and J. P. Taylor, 1966 Non-linear dominance relationships between *S* alleles. Heredity **21:** 345–362.

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

Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.

Vekemans, X., M. H. Schierup and F. B. Christiansen, 1998 Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. Evolution **52:** 19–29.

Wright, S., 1939 The distribution of self-sterility alleles in populations. Genetics **24:** 538–552.

Xue, Y. B., R. Carpenter, H. G. Dickinson and E. S. Coen, 1996 Origin of allelic diversity in *Antirrhinum S*-locus RNases. Plant Cell **8:** 805–814.

Communicating editor: R. G. Shaw

APPENDIX

Recursions under the assumption of symmetric frequencies within genotypic class: Genotypic frequencies in the next generation (denoted by primes) are

$$H' = \frac{1}{4}Dn_1\left[\frac{2H}{1-y} + \frac{R(n_2-1)}{1-2y+R} + \frac{D(n_1-1)}{1-x-y}\right] \tag{A1}$$

$$G' = \frac{1}{2}G(n_1-2)\left[\frac{G(n_1-3)}{1-2x+G} + \frac{Dn_2}{1-x-y}\right] + \frac{1}{2}Dn_2\left[\frac{G(n_1-2)}{1-2x+G} + \frac{Dn_2}{1-x-y}\right] \tag{A2}$$

$$R' = H\left[\frac{2H}{1-y} + \frac{R(n_2-2)}{1-2y+R} + \frac{Dn_1}{1-x-y}\right] + \frac{1}{2}R(n_2-2)\left[\frac{2H}{1-y} + \frac{R(n_2-3)}{1-2y+R} + \frac{Dn_1}{1-x-y}\right] + \frac{1}{2}Dn_1\left[\frac{2H}{I-y} + \frac{R(n_2-1)}{1-2y+R} + \frac{D(n_1-1)}{1-x-y}\right] \tag{A3}$$

$$D' = \frac{1}{2}D + \frac{1}{4}x\left[\frac{2H}{1-y} + \frac{R(n_2-1)}{1-2y+R}\right] + \frac{1}{4}[2H + R(n_2-1) + D(n_1-2)]\frac{G(n_1-1)}{1-2x+G}. \tag{A4}$$

Linearized recursions in the frequencies of carriers of a rare class I allele: Recursions linearized in the frequencies (ε_G and ε_D) of carriers of a rare class I allele in populations in which at least two common class I alleles and at least one common class II allele occur reduce to

$$\varepsilon'_G = \varepsilon_G\frac{1}{4}(n_1-1)\left[\frac{G(n_1-2)+Dn_2}{1-x} + \frac{G(n_1-2)}{1-2x+G} + \frac{Dn_2}{1-x-y}\right] + \varepsilon_D\frac{1}{4}n_1\left[\frac{G(n_1-1)+Dn_2}{1-y} + \frac{G(n_1-1)}{1-2x+G} + \frac{Dn_2}{1-x-y}\right] \tag{A5}$$

$$\varepsilon'_D = \varepsilon_G\frac{1}{4}n_2\left[\frac{2H+R(n_2-1)+D(n_1-1)}{1-x}\right] + \varepsilon_G\frac{1}{4}n_2\left[\frac{2H}{1-y} + \frac{R(n_2-1)}{1-2y+R} + \frac{D(n_1-1)}{1-x-y}\right] + \varepsilon_D\frac{1}{4}n_2\left[\frac{2H}{1-y} + \frac{R(n_2-1)}{1-2y+R} + \frac{Dn_1}{1-x-y}\right] + \varepsilon_D\frac{1}{4}\left[\frac{2H(n_2-1)+R(n_2-1)(n_2-2)+Dn_1n_2}{1-x}\right]. \tag{A6}$$

Linearized recursions in the frequencies of carriers of a rare class II allele: In populations containing at least three common *S*-alleles, including one from each class, linearized recursions in the frequencies of carriers of a rare class II allele are given by

$$\varepsilon'_D = \varepsilon_D\frac{1}{4}(n_1-1)\left[\frac{G(n_1-2)+Dn_2}{1-x} + \frac{G(n_1-2)}{1-2x+G} + \frac{Dn_2}{1-x-y}\right] + \varepsilon_R\frac{1}{4}n_1\left[\frac{G(n_1-1)+Dn_2}{1-y} + \frac{G(n_1-1)}{1-2x+G} + \frac{D(n_2-1)}{1-x-y}\right] \tag{A7}$$

$$\varepsilon'_R = \varepsilon_D\frac{1}{4}n_2\left[\frac{2H+R(n_2-1)+D(n_1-1)}{1-x}\right] + \varepsilon_D\frac{1}{4}n_2\left[\frac{2H}{1-y} + \frac{R(n_2-1)}{1-2y+R} + \frac{D(n_1-1)}{1-x-y}\right] + \varepsilon_R\frac{1}{4}\left[\frac{2H(n_2-1)+R(n_2-1)(n_2-2)+Dn_1n_2}{1-y}\right] + \varepsilon_R\frac{1}{4}(n_2-1)\left[\frac{2H}{1-y} + \frac{R(n_2-2)}{1-2y+\mathrm{R}} + \frac{Dn_1}{1-x-y}\right]. \tag{A8}$$