Vallejo-Marin, Mario and Uyenoyama, Marcy K. 2008.

On the evolutionary modification of self-incompatibility: Implications of partial clonality for allelic diversity and genealogical structure. Pages: 53-71
Chapter 3 in:
V. E. Franklin-Tong, ed. Self-Incompatibility in Flowering plants: Evolution, Diversity and Mechanisms. Springer-Verlag, Berlin.

http://dx.doi.org/10.1007/978-3-540-68486-2_3

On the Evolutionary Modification of Self-Incompatibility: Implications of Partial Clonality for Allelic Diversity and Genealogical Structure

M. Vallejo-Marín and M.K. Uyenoyama

Abstract Experimental investigations of homomorphic self-incompatibility (SI) have revealed an unanticipated level of complexity in its expression, permitting fine regulation over the course of a lifetime or a range of environmental conditions. Many flowering plants express some level of clonal reproduction, and phylogenetic analyses suggest that clonality evolves in a correlated fashion with SI in *Solanum* (Solanaceae). Here, we use a diffusion approximation to explore the effects on the evolutionary dynamics of SI of vegetative propagation with SI restricted to reproduction through seed. While clonality reduces the strength of frequency-dependent selection maintaining *S*-allele diversity, much of the great depth typical of *S*-allele genealogies is preserved. Our results suggest that clonality can play an important role in the evolution of SI systems, and may afford insight into unexplained features of allele genealogies in the Solanaceae.

Abbreviations

GSI	Gametophytic self-incompatibility
MRCA	Most recent common ancestor
PSI	Pseudo self-incompatibility
S-RNase	*S*-locus ribonuclease
SC	Self-compatibility

SCR	*S*-locus cysteine rich (also termed SP11) – the pollen determinant in Brassica
SI	Self-incompatibility
SRK	S-Receptor Kinase (the pistil determinant in Brassica)
SSI	Sporophytic self-incompatibility
ΨSCR	SRK-like sequences in Arabidopsis

3.1 Introduction

A watershed era in the study of mating systems began upon the cloning of the *S*-locus over two decades ago (Nasrallah et al., 1985; Anderson et al., 1986). Within evolutionary biology, the revelation of the genetic basis of self-incompatibility (SI) in two model systems precipitated new explorations of the origins, evolutionary modification, and genomic consequences of its expression. In addition to illuminating physiological and regulational mechanisms of SI, the nucleotide sequence of the *S*-locus region itself afforded insight into the evolutionary process over the extraordinary time depth of *S*-allele genealogies.

Virtually every system subjected to close examination has now been found to depart from canonical SI, defined as a one-factor system that both defines compatibility classes and excludes incompatible pollen from fertilization in every reproductive episode. These empirical discoveries invite a broadened concept of SI. Here, we characterize SI as a versatile component of a diversified strategy of reproduction, the evolutionary modification of which spans the full spectrum from total disablement to plastic regulation of its expression.

Section 3.2 begins with a consideration of recent phylogenetic analyses of the transition to self-compatibility (SC). These studies provide evidence of multiple independent departures from canonical SI, and do not exclude subsequent restoration of SI. Further, some transitions from canonical SI may not in fact reflect a loss of function, but rather modification of its expression: a plastic response to physiological state or external environment, for example (see also Chap. 2). Within the Solanaceae, an exemplar for studies of the genetic basis and evolution of SI, a number of species express SI in combination with partial clonal reproduction. Because partial clonality and SI show correlated evolution (Vallejo-Marín and O'Brien, 2007), we develop a diffusion approximation of partial vegetative propagation with SI restricted to reproduction through seed (Sect. 3.3).

We find that while partial clonality reduces the number of *S*-alleles maintained at stochastic equilibrium, it preserves much of the great depth characteristic of *S*-allele genealogies. Our analysis illustrates that the nature of expression of SI influences the inferential framework extending from gene genealogies based on the *S*-locus region to reconstruction of the evolutionary history of mating systems.

3.2 Mating System Dynamics

A characteristic of SI that prevails across the plant kingdom is its evolutionary lability (Barrett, 1988). Stebbins (1974, p 51) described the transition from SI to self-fertilization as one of the evolutionary paths most often taken by flowering plants. While it seems clear that many SC species exist in families in which SI appears to be ancestral (see Chap. 4), distinguishing the modification of SI from its complete disablement requires direct experimental study. In this section, we review some recent analyses of transitions from canonical SI.

3.2.1 Relative Transition Rates

In the Solanaceae, Igic et al. (2006) found that transitions from complete SI have occurred at rates at least several-fold higher than the reverse. They interpreted the failure to find greater support for a model that permits both gain and loss of SI over to a model that precludes gains as evidence of unidirectional loss of gametophytic SI (GSI) in the Solanaceae (see also Chap. 5). However, even the analysis in which they specified the SI status of ancestral nodes indicated virtually identical likelihoods of the two models, and the analysis without assigned ancestral states gave significantly more support to the model that allowed positive rates of gain of SI.

Some cases of apparent loss of SI may in fact constitute transitions from canonical SI to some form of modified expression, as has been described, for example, by Levin (1996). In their survey of the literature on breeding systems in the Asteraceae, Ferrer and Good-Ávila (2007) recognized not only full SI and full SC, but also partial or pseudo-SI (PSI), broadly defined as variation in the strength of SI among individuals or populations in response to floral age, temperature, or various environmental factors (see also Chap. 2). Their phylogenetically independent maximum-likelihood analysis indicated substantial rates of gain of SI from both SC and PSI.

3.2.2 Multiple Origins of SC in Arabidopsis

SRK/SCR-based sporophytic self-incompatibility (SSI, see Chaps. 6–8) has not been detected outside the Brassicaceae. From an analysis of nucleotide variation at cytoplasmic and nuclear loci, Koch et al. (2001) estimated the emergence of this family about 40 million years ago (mya). This figure appears to coincide with an estimate for the origin of this form of SSI (40–50 mya; Uyenoyama, 1995). A key case study is the derivation of SC *Arabidopsis thaliana* from a presumed SI ancestor in the past 5–6 million years, a divergence time supported by analyses of both nucleotide substitutions (Koch et al., 2001) and chromosomal rearrangements (Koch and Kiefer, 2005).

Shimizu et al. (2004) interpreted the low variation at *ΨSCR1*, the pseudogene homologous to the determinant of pollen specificity (*SCR*), as evidence of

a very recent (<0.32 mya) transition to SC through the selective sweep of an *S*-allele with impaired pollen function. Substitution of a loss-of-function *S*-allele has been proposed as a possible route from SI to SC (Charlesworth and Charlesworth, 1979). However, contrary to expectation under the scenario, Tang et al. (2007) found multiple, highly divergent *S*-locus haplotypes in a survey of 96 accessions of *A. thaliana*. Shimizu et al. (2008) confirmed these results in accessions from Europe (haplogroup A) and African island populations (haplogroup B). Haplogroup A, found in 96% of world wide accessions, is associated with *ΨSCR1*, but *SCR* in haplogroup B shows no obvious disabling mutations. In haplotype B, it is *SRK*, the regulator of the stylar response to incompatible pollen, that shows a frameshift mutation (Tang et al., 2007; Shimizu et al., 2008).

Nasrallah et al. (2004) showed that *A. thaliana* accessions differ in their response to the acquisition by transformation of functional *SRK* and *SCR* genes from its SI congener *A. lyrata*. The only accession among the seven tested that showed full expression of SI upon transformation (C24) contains a highly rearranged *S*-locus region, the apparent product of recombination between *S*-haplotypes (Sherman-Broyles et al., 2007). Other accessions show pseudo-SC, with expression of SI breaking down in later stages of flower development. Liu et al. (2007) showed that a locus distinct from both *SCR* and *SRK* modified the age-dependent pattern of SI expression in a PSC transformant.

These studies demonstrate that the transition of *A. thaliana* from SI to SC derives from multiple origins and genetic mechanisms (see also Chap. 6). Of particular relevance to our discussion is the possibility that at least some of these transitions may have involved partial or modified expression of SI rather than the immediate and total loss of function.

3.2.3 Modified Forms of SI

Mounting empirical evidence suggests that in many taxa the expression of SI within individuals or populations may be restricted to certain developmental stages or environmental conditions (Levin 1996; Good-Ávila et al., in Chap. 2).

3.2.3.1 Partial SI

Partial expression of SI occurs in many taxa representing taxonomic groups across the plant kingdom (Levin, 1996). In one of the few phylogenetic analyses of partial SI, Ferrer and Good-Ávila (2007) found that in the Asteraceae, 10% of the 571 species studied expressed some form of PSI. Good-Ávila and Stephenson (2002) found a heritable increase in self-fertility with floral age in *Campanula rapunculoides* (Campanulaceae). In *Solanum carolinense* (Solanaceae), Travers et al. (2004) used direct experimental manipulation to demonstrate an increase in self-fertility with floral age and in plants that were prevented from setting fruit earlier

in the reproductive season. Moreover, by measuring fruit and seed set following self-pollination across clonally replicated genets, they detected a genetic component to the variation in self-fertility. Also in *S. carolinense*, Mena-Alí and Stephenson (2007) found *S*-allele-specific variation in the level of self-fertility, suggesting that this effect reflects a change in the rate of translation or turnover or in the enzymatic activity level of the particular *S-RNase* allele, perhaps due to a linked modifier gene.

3.2.3.2 SI Under Partial Clonal Reproduction

Natural populations of a number of species, including many in the typically colonizing genus *Solanum* (Solanaceae), comprise individuals derived from asexual propagation together with individuals derived from seeds generated through complete SI expression. The capacity for clonal reproduction may reduce the selective advantage of the loss or modification of SI by permitting the extended persistence of isolated SI colonizers. In this case, one might expect correlated evolution of SI status and clonality status to follow from Baker's (1955; 1967) rule.

In the only phylogenetic study of the correlated evolution of SI and life-history to date, Vallejo-Marín and O'Brien (2007) used the Pagel method (Pagel, 1994; Pagel and Meade, 2006) to address the joint evolution of SI and clonality. They detected significant rates of transition from SI to SC under clonality and from clonality to non-clonality under SC (see Fig. 2 of Vallejo-Marín and O'Brien, 2007). The loss of clonality in colonizing SI *Solanum* species may be strongly deleterious, perhaps reflecting that having an alternative means of reproduction can become essential when seed set is limited by an insufficiency of compatible pollen.

3.3 *S*-Locus Evolution Under Partial Clonality

In this section, we use an extension of the one-dimensional diffusion equation approximation of *S*-locus evolution introduced by Wright (1939) to explore the consequences of partial clonal reproduction on the evolutionary dynamics of GSI (see Appendix 1).

3.3.1 Diffusion Approximation

Figure 3.1 depicts a population in which zygotes derived from seeds generated under full expression of SI constitute a fraction γ of reproductives, with the complement derived by clonal propagation. For example, a proportion c of the reproductive resources of an individual, irrespective of *S*-locus genotype, may be devoted to clonal reproduction, with clonally derived offspring (ramets) surviving at rate τ relative to sexually derived offspring (genets).

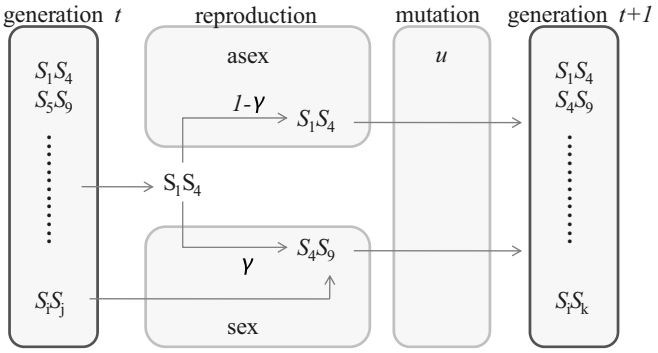

Fig. 3.1 Flow chart of a single generation in the simulations

To develop a one-dimensional diffusion approximation of the high-dimensional dynamics of arbitrary numbers of S-alleles, we assume exchangeability among S-alleles (Wright, 1939; Fisher, 1958, Chap. 4). Over a single generation, the expected change in frequency of a particular S-allele (q) corresponds to

$$\Delta q = \frac{\gamma q(F-q)}{(1-F)(1-2F)}, \tag{3.1}$$

for F denoting homozygosity, approximated by the inverse of the number of S-alleles maintained in the population ($1/n$); and γ the sexual fraction, defined as the proportion of reproductive individuals that are themselves sexually derived:

$$\gamma = \frac{1-c}{c\tau+1-c}. \tag{3.2}$$

Novel (functionally distinct) S-alleles enter the population at rate u per gene per generation, and genetic drift is incorporated through random sampling of N zygotes to form the next generation, reflecting that compatibility reflects interactions between zygotes and gametes rather than between gametes (Fisher, 1958; Wright, 1960).

These expressions give rise to a diffusion approximation with infinitesimal mean and variance coefficients

$$\mu(x) = -\frac{\gamma x(x-F)}{(1-F)(1-2F)} - ux, \tag{3.3a}$$

$$\sigma^2(x) = \frac{x(1-2x)}{2N}. \tag{3.3b}$$

Our diffusion coefficients reduce under full sexual reproduction ($\gamma = 1$) to those of the Wright (1960) model. Partial clonal reproduction decreases the force of SI selection on the infinitesimal mean change (3.3a) by a factor corresponding to the proportion of the population derived by sexual reproduction (3.2).

Drift generates change on the order of the inverse of the effective population size (3.3b). With the exception of $\gamma x(x-F)$ on the order of $1/2N$, the intense frequency-dependent selection induced by SI overwhelms drift (Neuhauser, 1999). Like the Wright (1939) model, our diffusion approximation holds only very close to extinction of the focal allele ($x \approx 1/2N$) or very close to the deterministic equilibrium frequency ($x \approx F = 1/n$). The process jumps between the domains of extinction and common frequencies at rate

$$\lambda = \theta\eta, \tag{3.4}$$

in which

$$\theta = 2Nu \tag{3.5}$$

represents the number of novel S-alleles produced in each generation by mutation and η the probability that an S-allele present in a single copy (frequency $1/2N$) will jump to common frequencies before extinction. Because drift reflects sampling of N zygotes rather than $2N$ genes, θ here is half the scaled mutation rate that customarily arises under random union of gametes. Under our model,

$$\eta = \frac{2n}{(n-1)(n-2)}\left[\gamma - \frac{u(n-1)(n-2)}{n}\right]. \tag{3.6}$$

Under complete clonal reproduction ($\gamma = 0$), mutation and drift of clones alone govern S-allele dynamics. Sexual reproduction ($\gamma > 0$) permits SI selection, but if mutation dominates selection ($u \gg \gamma$), the invasion probability (3.6) converges to zero for $\theta > 1$ (Chap. 15, Sect. 4, Example D of Karlin and Taylor, 1981). Here, we restrict attention to cases where the frequency of sexually derived individuals at reproduction ($M = \gamma N$) is sufficiently large relative to the scaled mutation rate (θ) to ensure a positive jump probability η (3.6) and that sexual reproduction through SI can occur ($n \geq 3$); a necessary condition is

$$\gamma > \frac{u(n-1)(n-2)}{n}.$$

At approximate steady state, the number of common S-alleles (n) segregating in the population reflects a balance between the rate at which S-alleles jump to common frequencies λ (3.4) and the rate at which common alleles jump to extinction (Uyenoyama, 2003). In our model, n is implicitly determined by

$$1 = \theta G e^{\frac{2Mn}{(n-1)(n-2)} - \theta}\sqrt{\frac{2\pi(n-1)}{Mn + \theta(n-1)}}\left\{\left[1 + \frac{\theta(n-1)}{Mn}\right]\frac{n-2}{n}\right\}^{\theta + Mn/(n-1)}, \tag{3.7}$$

in which

$$G = \frac{2n\gamma}{\eta(n-1)(n-2)} = \frac{2Mn}{2Mn - \theta(n-1)(n-2)}$$

and

$$M = \gamma N. \tag{3.8}$$

These expressions indicate that the number of sexually derived individuals at reproduction (M) and the scaled mutation rate (θ) jointly determine allelic diversity at the S-locus. In particular, the effect of partial clonal reproduction on the evolutionary dynamics at the S-locus cannot be explained simply by a reduction in effective population size by the factor γ. The analogous expression derived by Yokoyama and Hetherington (1982) for the Wright model ($\gamma = 1$) also depends on both u and N, and not only their product.

The Wright–Fisher formula for the stationary distribution provides an expression for the density of the frequency of the focal S-allele at approximate steady-state. In this case, the probability that the focal S-allele segregates in the population at a frequency lying in a small interval near x is proportional to

$$\Phi(x) = 2\theta \mathrm{e}^{\frac{2Mn^2x}{(n-1)(n-2)}} (1-2x)^{Mn/(n-1)+\theta-1} x^{-1}. \tag{3.9}$$

3.3.2 S-*Allele Number and Frequency*

Self-incompatibility induces an intense form of frequency-dependent selection (Vekemans and Slatkin, 1994; Neuhauser, 1999) that typically maintains scores of S-alleles. Wright (1939) reduced the description of this high-dimensional evolutionary process to a one-dimensional diffusion equation through a series of inspired approximations. To explore the robustness of our model to similar approximations, we compared theoretical expectations of the number of common S-alleles maintained and their frequency spectrum to the results of numerical simulations (see Appendix 2).

Figure 3.2 presents box plots of the total number of segregating S-alleles over a range of values of γ in a population of $N = 250$ zygotes under a rate of mutation to new S-alleles of $u = 10^{-4}$. Also shown is the prediction from our model, obtained by numerically solving (3.7) for n and determining the largest integer not exceeding the expectation (floor). The median total number of S-alleles observed generally exceeds by 1 or 2 the expected number of common S-alleles, those occurring in approximately equal frequency.

To study the spectrum of allele frequencies, we determined the frequency in the population of each segregating S-allele in 100 independent simulations of a population of 250 zygotes under mutation rate $u = 10^{-4}$. Figure 3.3 shows the total number of alleles segregating in frequencies indicated on the abscissa under two assignments of the proportion of sexually derived zygotes (γ). Also shown is the theoretical stationary distribution (3.9) of the frequency of an S-allele. For a given set of population parameters (N, u, γ), we first determined the expected number of S-alleles as the largest integer not exceeding the root of (3.7) and then assigned this value to n in (3.9) to obtain the solid curves in Fig. 3.3.

In accordance with a prediction (3.9) of our model, a number of S-alleles hover at the edge of extinction, with the frequencies of most alleles clustered around a higher mode. Arrows indicating the expected frequency of a common allele (inverse of the

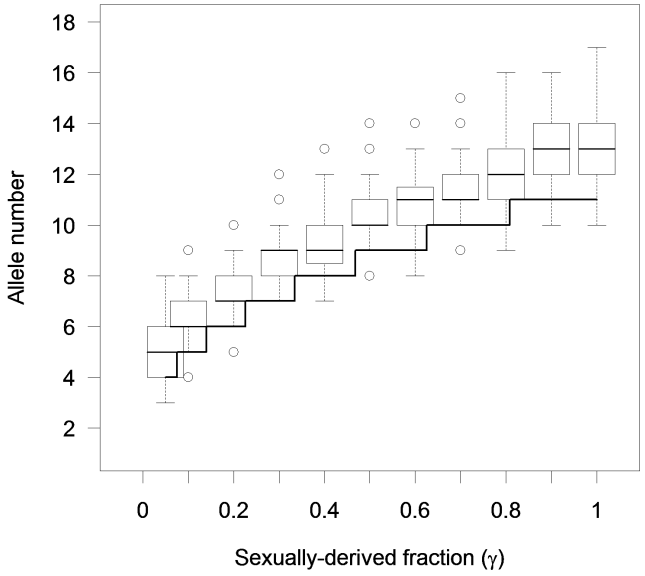

Fig. 3.2 Total number of *S*-alleles segregating in a population of 250 zygotes with a mutation rate of $u = 10^{-4}$. Each box plot summarizes the results of 100 independent simulations, censused in generation 15,000 following the burn-in period (median, first and third quartiles, whiskers at approximate 95% confidence limits, and *circles* for observations beyond the whiskers). Each *horizontal line* represents the expected number of *S*-alleles, corresponding to the largest integer not exceeding the solution for n in (3.7)

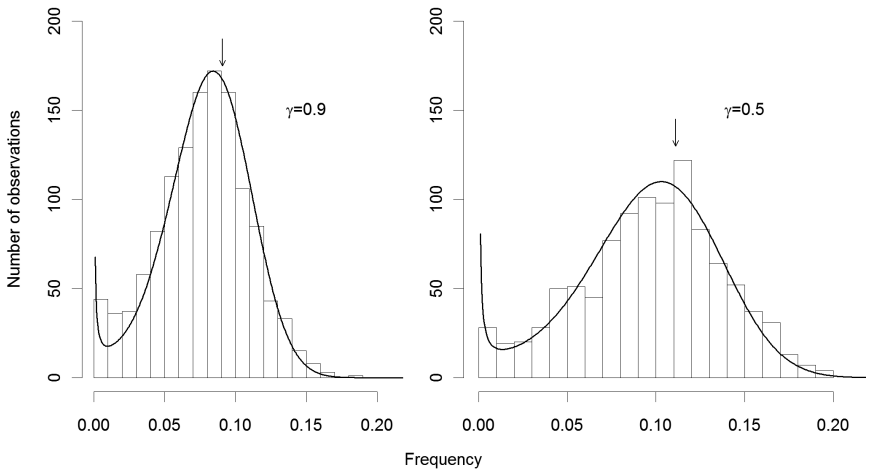

Fig. 3.3 Histograms of *S*-allele frequencies observed in 100 independent simulations compared to expected stationary distributions (3.9) with proportions $\gamma = 0.9$ (*left*) and $\gamma = 0.5$ (*right*) of reproductives sexually derived. Arrows indicate the expected frequency of a common allele $(1/n)$. The *vertical* positions of the *solid curves* are determined by eye. Other parameters as in Fig. 3.2

expected number) lie close to this mode. As γ declines, fewer *S*-alleles segregate; the modal frequency increases and allele frequency vary over a greater range.

3.3.3 Age of the Root

Having established that our diffusion equation approximation provides a good description of a number of classical characteristics of *S*-allele evolution, we now use the model to address the evolutionary process from a genealogical perspective, aspects of which are more difficult to explore using forward-in-time simulations alone.

3.3.3.1 Expansion of Time Scale

Upon the advent of nucleotide sequences sampled from the *S*-locus and adjacent genomic regions, explorations of genealogical relationships among *S*-allele lineages revealed an extensive pattern of deep divergence (Ioerger et al., 1990; Dwyer et al., 1991), on the order of tens of millions of years.

Takahata (1990) characterized genealogies of functionally distinct classes of alleles maintained by exchangeable forms of balancing selection as similar to neutral genealogies, but on a time scale expanded by the scaling factor

$$f = \frac{n^2}{4N\lambda},$$

for n the number of common alleles maintained at stochastic steady state, N the effective population size, and λ (3.4) the rate at which rare allelic lineages become common. This characterization of the process of divergence of *S*-allele lineages suggests that the expected age of the most recent common ancestor (MRCA) of n *S*-alleles corresponds to

$$4Nf(1-1/n) \tag{3.10}$$

(see, for example, Tajima, 1983). Under SI with clonal reproduction (3.4), the scaling factor corresponds to

$$f = \frac{n^2(n-1)(n-2)}{4\theta[2Mn - \theta(n-1)(n-2)]}. \tag{3.11}$$

3.3.3.2 Reduction in the Sexually Derived Fraction

To explore whether the effects of clonality on the evolutionary dynamics are equivalent to a reduction in effective population size alone, we compared the number of common *S*-alleles maintained under partial clonal reproduction (closed symbols in Fig. 3.4) to that maintained under pure non-clonality in a population comprising the

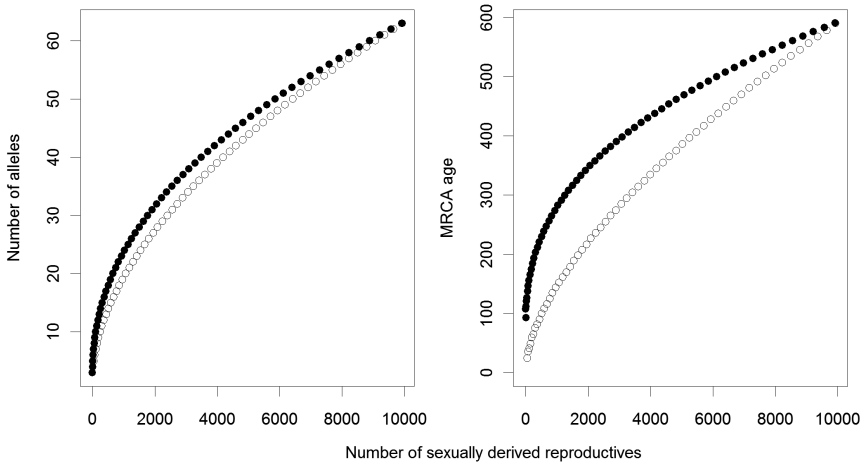

Fig. 3.4 Number of segregating *S*-alleles (n, *left panel*) and expected age of the root of the *S*-allele tree ($4Nf(1-1/n)$, *right panel*) under a rate of mutation to new *S*-alleles of $u = 10^{-5}$. *Closed symbols* correspond to a population of effective size $N = 10^4$ reproducing under partial clonality ($\gamma > 0$). *Open symbols* correspond a population practicing exclusive sexual reproduction with effective population size equal to the sexual fraction ($N\gamma$) in the partially clonal population

same number of sexually derived parents ($N\gamma$; open symbols in Fig. 3.4). The left panel of Fig. 3.4 confirms the positive relationship between the steady-state number of common *S*-alleles (n) and the proportion of sexually derived reproductives (γ) as indicated in Fig. 3.2. Further, it shows that partial clonality maintains somewhat more *S*-alleles than does pure sexuality in a smaller population, possibly reflecting that mutations to new *S*-alleles arise in both the sexual and clonal fractions.

To explore the effect of clonality on the depth of *S*-allele genealogies, we made a similar comparison of the expected age of the MRCA (3.10) using the scaling factor (3.11) induced by SI expression under partial clonality. The right panel of Fig. 3.4 indicates that the age of the root of the genealogy of the segregating *S*-alleles under partial clonality can exceed that under pure sexuality.

3.3.3.3 Discrepancy Between Allele Number and Genealogical Depth

Figure 3.4 suggests that even as the number of *S*-alleles (n) declines with the rate of sexual reproduction (γ), genealogical depth can remain large. To address the implications of an incorrect presumption of the absence of clonality, we conducted a thought experiment. Given a rate of mutation to new *S*-alleles (u) and a number of *S*-alleles (n) segregating in a population undergoing partial clonal reproduction, how great would be the discrepancy between the actual depth of the *S*-allele genealogy and the depth expected under complete sexual reproduction? We simulated the process by comparing, for a given assignment of effective population size (N), mutation rate (u), and number of *S*-alleles (n), the expected age of the MRCA under partial

clonal reproduction (3.10) to the age expected in a purely sexual population with an identical number of common S-alleles with an effective size N^*.

To find the expected age of the MRCA (3.10) in the partially clonal population, we determined, for values of γ over its entire range $(0,1)$, the expected numbers of common S-alleles (n). Because n is restricted to integer values, a range of γ values is consistent with each value of n; we chose the minimum value of γ to represent the range. For each pair of values for n and γ, we obtained the expected age of the MRCA from (3.10). To find the age of the MRCA in a purely sexual population, given the same values of n and u, we first solved the Yokoyama and Hetherington (1982) equation, which corresponds to (3.7) with $\gamma = 1$, for the range of effective population sizes consistent with n, and then obtained the corresponding scaling factor (f^*) from (3.11) that would be appropriate for a completely sexual population of size equal to the number of sexually-derived reproductives in the partially clonal population.

Figure 3.5 shows the ratio the ages of the MRCA ($N^* f^* / Nf$) as a function of the sexual proportion γ. Each S-allele number n determines a range of N^* values, giving rise to the range of $N^* f^* / Nf$ ratios. Except for values of γ very close to 1 (absence of clonality), where the two models are identical, $N^* f^* / Nf$ lies below unity. This implies that wrongly presuming the absence of clonal reproduction would cause the depth of the S-allele genealogy to appear too deep relative to the number of segregating S-alleles (n). Figure 3.5 indicates that under high levels of clonality

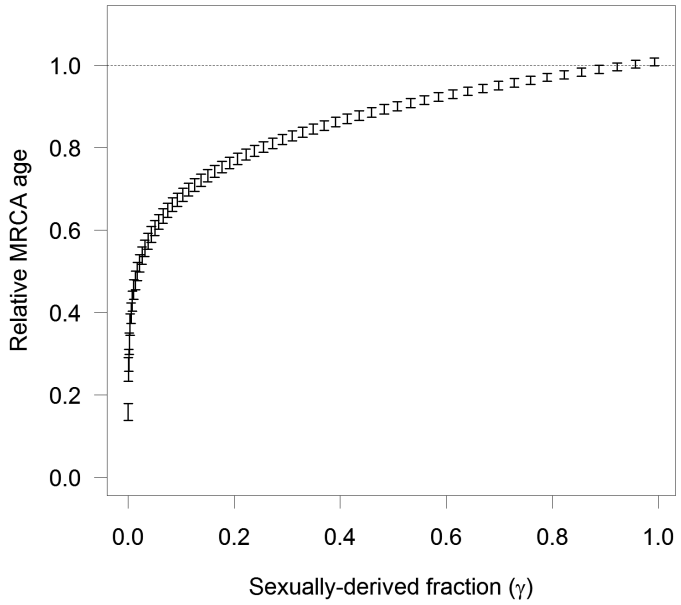

Fig. 3.5 Expected age of the MRCA in a purely sexual population as a proportion of the expected age of the MRCA in a partially clonal population of effective size $N = 10^4$, under a mutation rate $u = 10^{-5}$

(low γ), the actual genealogical depth can exceed the depth expected under non-clonality by several-fold.

3.4 Discussion

3.4.1 Clonality in the Solanaceae

Our diffusion approximation suggests that partial clonal reproduction can have profound effects on *S*-allele diversity and their genealogical span. We expect that clonality would constitute an important influence on the evolutionary dynamics at the *S*-locus only if its history had a genealogical depth comparable to that of *S*-alleles. Clonal reproduction characterized all (33/33) SI *Solanum* species studied by Vallejo-Marín and O'Brien (2007) and 68.5% (37/54) of SC species. While no formal treatment exists of the prevalence or the ancestral clonality status within the Solanaceae, the widespread occurrence of herbaceous perennials in this family (*e.g., Solanum*), together with the observation that the majority of herbaceous perennial can reproduce vegetatively (Richards, 1986), suggests that clonality may be a common feature of this group. Groenendale et al. (1996) estimate that 50–75% of species in this family are clonal, and the closest related family in their analysis (Convolvulaceae) also shows a high incidence of clonality. These observations are consistent with the view that clonality has played a substantial role in the history of groups in the Solanaceae, including the genus *Solanum* in particular.

While extensive trans-specific sharing of *S*-allele lineages appears to be characteristic of all systems of SI, differences in genealogical structure are apparent among groups (Uyenoyama, 1997; Richman and Kohn, 1999). In particular, the analysis of Uyenoyama and Takebayashi (2004) suggested that a difference in effective population size alone between *Solanum carolinense* (with an estimated 13 segregating *S*-alleles) and three species of *Physalis* (with an estimated 72 segregating *S*-alleles) is unlikely to account for the differences in tempo and mode of *S*-allele diversification in these solanaceous taxa. Even though *S. carolinense* shows the lowest number of segregating *S*-alleles among species for which estimates are available (Lawrence, 2000), its *S*-allele lineages span almost the entire genealogical range contained in the Solanaceae (Uyenoyama and Takebayashi, 2004). Our analysis raises the possibility that clonal reproduction might contribute to this pattern of great genealogical depth in spite of low *S*-allele number.

3.4.2 Evolutionary Stability of Partial SI

Although partial expression of SI has traditionally been characterized as a transient state between complete SI and complete SC (Goodwillie et al., 2005), PSI can also represent a persistent condition.

Uyenoyama et al. (2001) explored the generation of new *S*-allele specificities through the successive loss and restoration of recognition between the pollen and stylar components of SI. Despite the shared evolutionary fate of factors under absolute linkage to the *S*-locus, evolutionary conflict between the male and female components can promote both mutations that impair pollen SI function, permitting an escape from rejection, and compensating mutations in the stylar component, restoring rejection (Newbigin and Uyenoyama, 2005). Under certain conditions, haplotypes with impaired pollen SI function can persist in stable polymorphism with full-function haplotypes (Uyenoyama et al., 2001).

Beyond the persistence of a partial breakdown until the restoration of full function, partial expression of SI can in fact constitute an evolutionarily stable state, resistant to the invasion of modifier alleles that increase as well as decrease the level of expression (Vallejo-Marín and Uyenoyama, 2004). In that model, we considered incomplete reproductive compensation of rejected mating opportunities through the expression of SI, perhaps as a consequence of low pollinator service or restriction of the domain of pollen transfer to the local neighborhood (Wilcock and Neiland, 2002). Selection on a modifier locus that controls SI expression level comprises two components: a direct effect (offspring number weighted by relatedness at the modifier locus) and a disequilibrium effect (greater rates of rejection and lower rates of reproductive compensation in *S*-locus heterozygotes than homozygotes). In a restricted parameter range, exact cancellation of these two components gives rise to an evolutionary stable intermediate level of SI expression.

3.4.3 Paradoxical Effects on Mating Systems

By reducing the intensity of frequency-dependent *S*-locus selection (3.2), clonal reproduction permits the maintenance of fewer segregating *S*-alleles (Fig. 3.2). Further, by promoting the spatial clumping of individuals with identical genotypes, clonality may exacerbate pollen limitation by reducing the fraction of compatible pollen received (Charpentier, 2002). Both effects serve to increase the evolutionary cost of SI under clonality.

Paradoxically, clonality may also preserve SI under conditions of pollen limitation by promoting reproductive assurance through vegetative reproduction (Baker, 1955, 1967; Vallejo-Marín and O'Brien, 2007). For example, Pannell and Barrett (1998) demonstrated that perennial SI taxa or those that maintain a persistent seed bank are less likely to become locally extinct. Similarly, asexual reproduction may permit isolated colonists to propagate themselves until conditions favorable for sexual reproduction arise (Baker, 1955; Baker and Cox, 1984).

These considerations suggest that under conditions of pollen limitation, clonality can promote the maintenance or the breakdown of SI: a contrast Vallejo-Marín (2007) has dubbed the SI-clonality paradox (see also Honnay and Jacquemyn, 2008). Similarly, SI itself has paradoxical effects on genetic load: by reducing the

expression and purging of recessive deleterious mutations, especially in regions tightly linked to the *S*-locus, SI can shelter and increase genetic load (Uyenoyama, 2003; Stone, 2004). Because clonality and SI affect multiple components of reproduction, the resolution of the various conflicting evolutionary pressures may depend on many ecological factors (including plant density, spatial distribution of genotypes, and pollinator type and availability) as well as genetic factors (including epistasis among deleterious mutations and genomic structure in the vicinity of the *S*-locus).

3.5 Conclusions

Adoption of an unambiguous definition of the SI phenotype has been essential throughout the long history of the study of SI, and especially in the identification and cloning of the cardinal components of the *S*-locus. At the present juncture, following two remarkable decades of molecular-level analysis in several systems of homomorphic SI, a picture has emerged of the *S*-locus as a dynamic genomic region replete with evolutionary conflict, death, and rebirth.

Homology can now be defined at the genetic as well as phenotypic level (Kusaba et al., 2001), permitting the reconstruction of the history of SI even in taxa that do not express the SI phenotype (Tang et al., 2007; Shimizu et al., 2008). These developments suggest the possibility of now belaying the study of SI on the genetics to explore a wider diversity of SI phenotypes.

Recent experimental studies have revealed an array of departures from canonical SI, consistent with a characterization of SI systems as highly adaptable and responsive to the exigencies of survival and reproduction. As yet, rather few studies have attempted to explore the evolutionary processes through which such diversification arise. Here, we have addressed a simple departure from canonical SI, incorporation of clonal propagation together with full SI expression in reproduction through seeds. We found that the addition of a single parameter to the classical Wright (1939) model of GSI gives rise to new evolutionary dynamics. In particular, while clonality reduces the number of segregating *S*-alleles, much of the extraordinary depth of *S*-allele genealogies persists.

Recognition of SI as a dynamic component of the reproductive phenotype invites experimental investigation of its diverse manifestations. A deepening understanding of the genetic mechanisms of modification in turn demands the theoretical explorations of their implications for the evolutionary process and empirical reconstructions of genealogical history.

Acknowledgements MVM was supported during part of this study by a post-doctoral fellowship from the Canada Research Chair program to SCH Barrett. US Public Health Service grant GM 37841 (MKU) provided partial support for this research.

Appendix 1: Diffusion Equation Approximation

At reproduction, the frequency in the population of zygotes bearing S-locus genotype S_iS_j is

$$TP'_{ij} = c\tau P_{ij} + (1-c)\sum_{k\neq i,j}\left[\frac{q_iP_{jk}}{1-q_j-q_k} + \frac{q_jP_{ik}}{1-q_i-q_k}\right]/2, \tag{3.12}$$

for c is the proportion of reproductive effort invested in clonal reproduction, τ the relative viability of clonally derived offspring, and T a normalizer to be determined. We address the change in frequency of a specific S-allele, arbitrarily designated S_i, imposing the assumption of equality among the frequencies of any of the other $n-1$ S-alleles presently segregating in the population.

For q the frequency of the focal allele S_i, the frequency of any other allele ($\hat{q}$) satisfies

$$(n-1)\hat{q} + q = 1.$$

Similarly, let P denote the frequency of the genotype comprising S_i together with any other allele, and $\hat{P}$ the frequency of any given genotype comprising alleles other than S_i:

$$\binom{n-1}{2}\hat{P} + (n-1)P = 1. \tag{3.13}$$

Because the allelic and genotypic frequencies are related by

$$q = (n-1)P/2, \tag{3.14}$$

a description of the population requires only a single free variable (q).

From (3.12) and an analogous recursion in the frequency of any given genotype that does not carry S_i, we obtain

$$TP' = c\tau P + (1-c)(n-2)\left[\frac{q\hat{P}}{1-2\hat{q}} + \frac{\hat{q}P}{1-q-\hat{q}}\right]/2$$
$$T\hat{P}' = c\tau\hat{P} + (1-c)\hat{q}\left[\frac{(n-3)\hat{P}}{1-2\hat{q}} + \frac{P}{1-q-\hat{q}}\right].$$

Together with (3.13), these expressions determine the normalizer T:

$$T = c\tau + 1 - c.$$

Using (3.14), the expected change in allele frequency (3.1) derives from the expression for the change in P. From (3.1), the derivation of the diffusion approximation follows as described by Uyenoyama (2003). Analysis of the diffusion equation derived here follows that of Uyenoyama (2003), with the replacement of a and b in that article by

$$a^* = \frac{\gamma}{(1-F)(1-2F)}$$
$$b^* = \frac{\gamma}{2(1-F)} + u.$$

Appendix 2: Simulations

Each simulation was started with a population of N individuals, heterozygous at the S-locus. This initial population had k S-alleles in equal frequencies. During each reproductive cycle, a new set of N individuals was created from the parental generation (Fig. 3.1). Each new offspring was formed by randomly selecting a parental individual as the maternal plant, and then allowing it to reproduce either clonally or sexually with probabilities c and $1-c$, respectively. We assumed no differences in any components of fitness among individuals or among sexual and clonal offspring (i.e. $\tau = 1$), which in turn implies $c = 1-\gamma$. For sexual reproduction, a new zygote was created by mating the maternal plant with a randomly selected pollen grain. All individuals, regardless of whether they reproduced by seed or vegetatively, contributed to the pollen pool. We assumed the complete expression of SI during the sexual component of the reproductive cycle. If the selected pollen was incompatible with the maternal plant, a new pollen grain was drawn. This process was continued until a compatible pollen grain was found, or until 900 pollination attempts were tried, after which we selected a new maternal parent. For the clonal component of reproduction, the maternal parent was used to generate a genetically identical offspring. New alleles at the S-locus were introduced by mutation at rate u. Mutations were introduced after the zygote was formed, i.e., we assumed that new mutations are not expressed in the gametophyte. Mutations occurred either in the maternal or paternal S-allele copy but not in both, and were equally likely to occur in sexually or clonally produced offspring.

To allow the system to stabilize before measurements were taken, we introduced a burn-in period that lasted until the initial k S-alleles coalesced to a single ancestor. After the burn-in period, we iterated the system for 15,000 more reproductive cycles before censusing the population. We repeated this process 100 times for each parameter combination to generate the simulation results.

References

Anderson MA, Cornish EC, Mau SL, Williams EG, Hoggart R, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penschow JD, Niall HD, Tregear GW, Coghlan JP, Crawford RJ, Clarke AE (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature 321:38–44

Baker HG (1955) Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347–349

Baker HG (1967) Support for Baker's Law – as a rule. Evolution 21:853–856

Baker HG, Cox PA (1984) Further thoughts on dioecism and islands. Ann Mo Bot Gard 71:244–253

Barrett S (1988). The evolution, maintenance, and loss of self-incompatibility systems. In: Lovett Doust J, Lovett Doust L (eds) Plant reproductive ecology: Patterns and strategies. Oxford University Press, Oxford, pp 98–124

Charlesworth D, Charlesworth B (1979) The evolution and breakdown of *S*-allele systems. Heredity 43:41–55

Charpentier A (2002) Consequences of clonal growth for plant mating. Evol Ecol 15:521–530

Dwyer KG, Balent MA, Nasrallah JB, Nasrallah ME (1991) DNA sequences of self-incompatibility genes from *Brassica campestris* and *B. oleracea*: Polymorphism predating speciation. Plant Mol Biol 16:481–486

Ferrer MM, Good-Ávila SV (2007) Macrophylogenetic analyses of the gain and loss of self-incompatibility in the Asteraceae. New Phytol 173:401–414

Fisher RA (1958) The genetical theory of natural selection, 2nd edn. Oxford University Press, Dover, New York

Good-Ávila SV, Stephenson AG (2002) The inheritance of modifiers conferring self-fertility in the partially self-incompatible perennial, *Campanula rapunculoides* L. (Campanulaceae). Evolution 56:263–272

Goodwillie C, Kalisz S, Eckert CG (2005) The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. Annu Rev Ecol Evol Syst 36:47–79

Groenendale JMv, Klimes L, Klimesova J, Hendriks RJJ (1996) Comparative ecology of clonal plants. Philos Trans R Soc Lond B 351:1331–1339

Honnay O, Jacquemyn H (2008) Mating system evolution under strong clonality: towards self-compatibility or self-incompatibility? Evol Ecol DOI 10.1007/s10682-007-9207-3

Igic B, Bohs L, Kohn JR (2006) Ancient polymorphism reveals unidirectional breeding system shifts. Proc Natl Acad Sci USA 103:1359–1363

Ioerger TR, Clark AG, Kao Th (1990) Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. Proc Natl Acad Sci USA 87:9732–9735

Karlin S, Taylor HM (1981) A second course in stochastic processes. Academic, New York

Koch MA, Kiefer M (2005) Genome evolution among cruciferous plants: A lecture from the comparison of the genetic maps of three diploid species – *Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. Am J Bot 92:761–767

Koch MA, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: Evidence from coding plastidic *matK* and nuclear *Chs* sequences. Am J Bot 88:534–544

Kusaba M, Dwyer K, Hendershot J, Vrebalov J, Nasrallah JB, Nasrallah ME (2001) Self-incompatibility in the genus *Arabidopsis*: Characterization of the *S*-locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. Plant Cell 13:627–643

Lawrence MJ (2000) Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. Ann Bot (Lond) 85(Suppl. A):221–226

Levin DA (1996) The evolutionary significance of pseudo-self-fertility. Am Nat 148:321–332

Liu P, Sherman-Broyles S, Nasrallah ME, Nasrallah JB (2007) A cryptic modifier causing transient self-incompatibility in *Arabidopsis thaliana*. Curr Biol 17:734–740

Mena-Alí JI, Stephenson AG (2007) Segregation analyses of partial self-incompatibility in self and cross progeny of *Solanum carolinense* reveal a leaky s-allele. Genetics 177:501–510

Nasrallah JB, Kao Th, Goldberg ML, Nasrallah ME (1985) A cDNA clone encoding an *S*-locus-specific glycoprotein from *Brassica oleracea*. Nature 318:263–267

Nasrallah ME, Liu P, Sherman-Broyles S, Nasrallah JB (2004) Natural variation in expression of self-incompatibility in *Arabidopsis thaliana*: Implications for the evolution of selfing. Proc Natl Acad Sci USA 101:16070–16074

Neuhauser C (1999) The ancestral graph and gene genealogy under frequency-dependent selection. Theor Popul Biol 56:203–214

Newbigin E, Uyenoyama MK (2005) The evolutionary dynamics of self-incompatibility systems. Trends Genet 21:500–505

Pagel M, Meade A (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. Am Nat 167:808–825
Pagel MD (1994) Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. Philos Trans R Soc Lond B 255:37–45
Pannell JR, Barrett SCH (1998) Baker's law revisited: Reproductive assurance in a metapopulation. Evolution 52:657–668
Richards AJ (1986) Plant breeding systems. George Allen & Unwin, London
Richman AD, Kohn JR (1999) Self-incompatibility alleles from *Physalis*: Implications for historical inference from balanced genetic polymorphisms. Proc Natl Acad Sci USA 96:168–172
Sherman-Broyles S, Boggs N, Farkas A, Liu P, Vrebalov J, Nasrallah ME , Nasrallah JB (2007) *S*-locus genes and the evolution of self-fertility in *Arabidopsis thaliana*. Plant Cell 19:94–106
Shimizu KK, Cork JM, Caicedo AL, Mays CA, Moore RC, Olsen KM, Ruzsa S, Coop G, Bustamante CD, Awadalla P, Purugganan MD (2004) Darwinian selection on a selfing locus. Science 306:2081–2084
Shimizu KK, Shimizu-Inatsugi R, Tsuchimatsu T, Purugganan MD (2008) Independent origins of self-compatibility in *Arabidopsis thaliana*. Mol Ecol 17:704–714
Stebbins GL (1974) Flowering plants: Evolution above the species level. Harvard University Press, Cambridge, MA
Stone JL (2004) Sheltered load associated with S-alleles in *Solanum carolinense*. Am J Bot 92:335–342
Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460
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
Tang C, Toomajian C, Sherman-Broyles S, Plagnol V, Guo YL, Ho TT, Clark RM, Nasrallah JB, Weigel D, Nordborg M (2007) The evolution of selfing in *Arabidopsis thaliana*. Science 317:1070–1072
Travers SE, Mena-Alí JI, Stephenson AG (2004) Plasticity in the self-incompatibility of *Solanum carolinense*. Plant Species Biol 19:127–135
Uyenoyama MK (1995) A generalized least-squares estimate for the origin of sporophytic self-incompatibility. Genetics 139:975–992
Uyenoyama MK (1997) 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
Uyenoyama MK, Takebayashi N (2004) Genus-specific diversification of mating types. In: Singh R, Uyenoyama MK (eds) The evolution of population biology. Cambridge University Press, New York, pp 254–271
Uyenoyama MK, Zhang Y, Newbigin E (2001) On the origin of self-incompatibility haplotypes: Transition through self-compatible intermediates. Genetics 157:1805–1817
Vallejo-Marín M (2007) The paradox of clonality and the evolution of self-incompatibility. Plant Signal Behav 2:265–266
Vallejo-Marín M, O'Brien HE (2007) Correlated evolution of self-incompatibility and clonal reproduction in *Solanum* (Solanaceae). New Phytol 173:415–421
Vallejo-Marín M, Uyenoyama MK (2004) On the evolutionary costs of self-incompatibility: Incomplete reproductive compensation due to pollen limitation. Evolution 58:1924–1935
Vekemans X, Slatkin M (1994) Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics 137:1157–1165
Wilcock C, Neiland R (2002) Pollination failure in plants: Why it happens and when it matters. Trends Plant Sci 7:270–277
Wright S (1939) The distribution of self-sterility alleles in populations. Genetics 24:538–552
Wright S (1960) On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of a given size: A reexamination. Biometrics 16:61–85
Yokoyama S, Hetherington LE (1982) The expected number of self-incompatibility alleles in finite plant populations. Heredity 48:299–303