

Research review

Historical inferences from the self-incompatibility locus

Author for correspondence:

Joshua R. Kohn

Tel: +1 858 534 8233

Fax: +1 858 534 7108

Email: jkohn@ucsd.edu

Received: 14 June 2003

Accepted: 15 August 2003

doi: 10.1046/j.1469-8137.2003.00952.x

Boris Igic¹, Lynn Bohs² and Joshua R. Kohn¹

Division of Biology, University of California, San Diego, USA; ¹Division of Biology, Box 0116, La Jolla, CA 92093, USA; ²Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

Summary

Key words: self-incompatibility, bottleneck, diversification, speciation, phylogeny, ancestral state.

Ancient polymorphism preserved at the self-incompatibility locus facilitates investigation of historical occurrences far older than extant species. We outline two ways in which studies of the S-locus can provide insights into patterns of speciation. First, we review evidence concerning the prevalence of founder events in speciation. A dramatic population size reduction is expected to reduce sequence diversity at the S-locus for millions of years. Only one potential bottleneck is preserved at the S-locus of the Solanaceae, suggesting that severe population size restrictions rarely occur in successful lineages. This must be interpreted with caution because of the restrictive conditions under which bottlenecks at the S-locus will be preserved. Second, S-locus polymorphism provides a novel opportunity to reconstruct, with considerable certainty, the presence or absence of self-incompatibility among ancestral taxa. We demonstrate this approach using a phylogenetic analysis and find that transitions from self-incompatibility to self-compatibility are common and essentially irreversible in the Solanaceae. Self-incompatibility is ancestral, but self-compatible taxa currently outnumber self-incompatible ones. Either self-incompatibility is going extinct, or the presence of incompatibility increases the diversification rate, maintaining a mixture of self-incompatible and self-compatible species in equilibrium. We outline how phylogenetic approaches can be used to determine the effect of incompatibility on diversification.

© *New Phytologist* (2003) **161**: 97–105

Introduction

The ability of plants to recognize and reject their own pollen (self-incompatibility) has been recognized as an important trait at least since Darwin's (1876) studies of the effects of inbreeding on fitness. Because they are subject to diversifying selection, self-recognition loci, such as the self-incompatibility locus in flowering plants, harbor extreme polymorphism in terms of the numbers of alleles maintained, the molecular

divergence among alleles, and the time to coalescence of the polymorphism. Therefore, the study of the self-incompatibility (S-) locus can provide information concerning mating system transitions with potentially important population genetic consequences, while also providing a source of historical information that stretches further back in time than neutral polymorphism. In this paper we review two uses of the self-incompatibility locus both as a source of historical inference about population sizes through time and as a mating system

trait that may affect the rates at which speciation and extinction occur.

Homomorphic self-incompatibility systems are generally split into two types: gametophytic, in which the genotype of the haploid male gametophyte determines specificity; and sporophytic, in which the genotype of the pollen parent determines specificity. Many of the applications of self-incompatibility discussed in this paper apply to both types of systems but we concentrate here on the RNase-based gametophytic self-incompatibility (GSI). This form of incompatibility is the most phylogenetically widespread, occurring in at least three families: the Solanaceae (Anderson *et al.*, 1986; McClure *et al.*, 1989), Rosaceae (Sassa *et al.*, 1993) and Scrophulariaceae (Xue *et al.*, 1996). It is well established that the female component of the SI reaction is an RNase, which both recognizes and rejects self-pollen (Anderson *et al.*, 1986; McClure *et al.*, 1989). Other mechanisms of incompatibility are known (Franklin-Tong & Franklin, 2003; Hiscock & Tabah, 2003) but molecular data from natural populations are much more abundant for RNase-based GSI systems than for any others.

The gametophytic S-locus of flowering plants has attracted the attention of population geneticists ever since Emerson (1939) reported extraordinary S-polymorphism in the narrow endemic *Oenothera organensis*. Wright (1939) recognized the power of negative frequency dependent selection to promote extreme polymorphism and produced formulae that relate the number of alleles maintained in populations to population effective size. Strong negative frequency dependence arises because rare alleles have access to more potential mates (Clark & Kao, 1994). The allele in the haploid pollen assays two alleles in each diploid style for a match and it can be shown that, at equilibrium with n alleles in a population, the advantage of the $n + 1$ st allele is $n/(n - 2)$. Therefore, even with 20 alleles in a large population, a newly arriving 21st allele has an average selective advantage of more than 11%!

Crossing studies using plants from natural populations have repeatedly found abundant functional polymorphism with dozens of S-alleles maintained in single populations at this locus (reviewed in Lawrence, 2000). More recently, progress in understanding of the molecular basis of incompatibility in a number of plant families (reviewed in Franklin-Tong & Franklin, 2003; Hiscock & Tabah, 2003) led to the development of molecular techniques for assaying S-allele diversity in natural populations. In parallel, application of coalescence approaches to loci under diversifying selection led to the prediction (Takahata, 1990; Vekemans & Slatkin, 1994), confirmed by molecular data (Ioerger *et al.*, 1990; Richman *et al.*, 1995, 1996a,b) that negative frequency dependent selection greatly increases the time-depth of polymorphism at the S-locus relative to neutral polymorphism. Such extreme and ancient polymorphism can be used as a tool to answer questions related to speciation and the effect that mating system traits may have upon it.

In the Solanaceae, the major outcomes of diversifying selection on S-RNase evolution were immediately apparent from

molecular population surveys and interspecific studies (Ioerger *et al.*, 1990; Richman *et al.*, 1995, 1996b) which found large numbers of functionally distinct S-alleles in natural populations. In addition, alleles from the same species often differed at more than 50% of amino acid residues, an outcome of both strong selection to diversify and great age of the polymorphism. Finally, phylogenetic analyses found that S-alleles from different species and genera were often more closely related to one another than were alleles within the same species. This phenomenon, known as trans-specific or trans-generic polymorphism arises when polymorphism in a common ancestor is passed down, with modification, to multiple descendent taxa (Fig. 1). The alternative hypothesis, that phylogenetic reconstruction inappropriately unites alleles from different taxa because of rampant homoplasy, is both unlikely (trans-specific polymorphism is the expectation under diversifying selection) and has been rejected by a shared polymorphism test (Ioerger *et al.*, 1990).

Population genetic inferences

A phylogenetic reconstruction of S-alleles from Solanaceae is shown in Fig. 2. The key feature to note is that all sampled species possess alleles representing multiple lineages that predate the diversification of the six genera represented. This is diagnosed by the existence of multiple well-supported

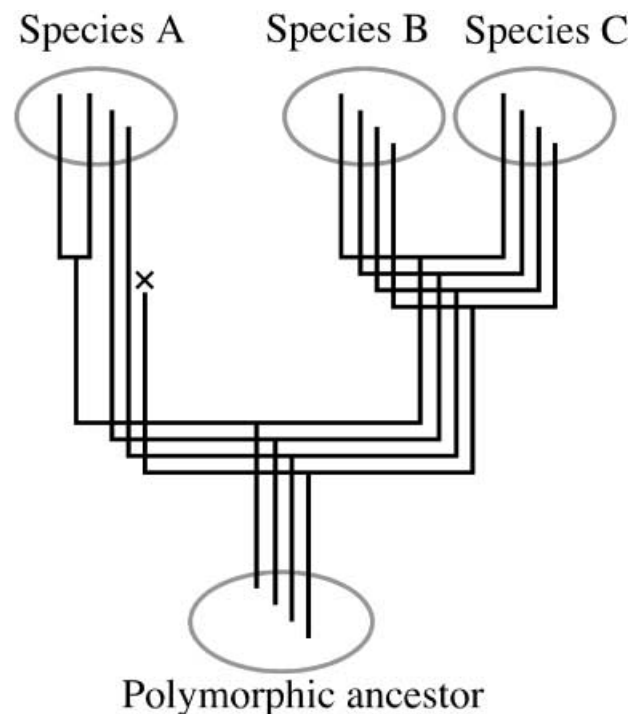


Fig. 1 Trans-specific polymorphism. Polymorphism present in the ancestral species is passed on, with modification, to descendent taxa. Alleles from the same lineage, rather than of the same species, will group together in phylogenetic reconstructions.

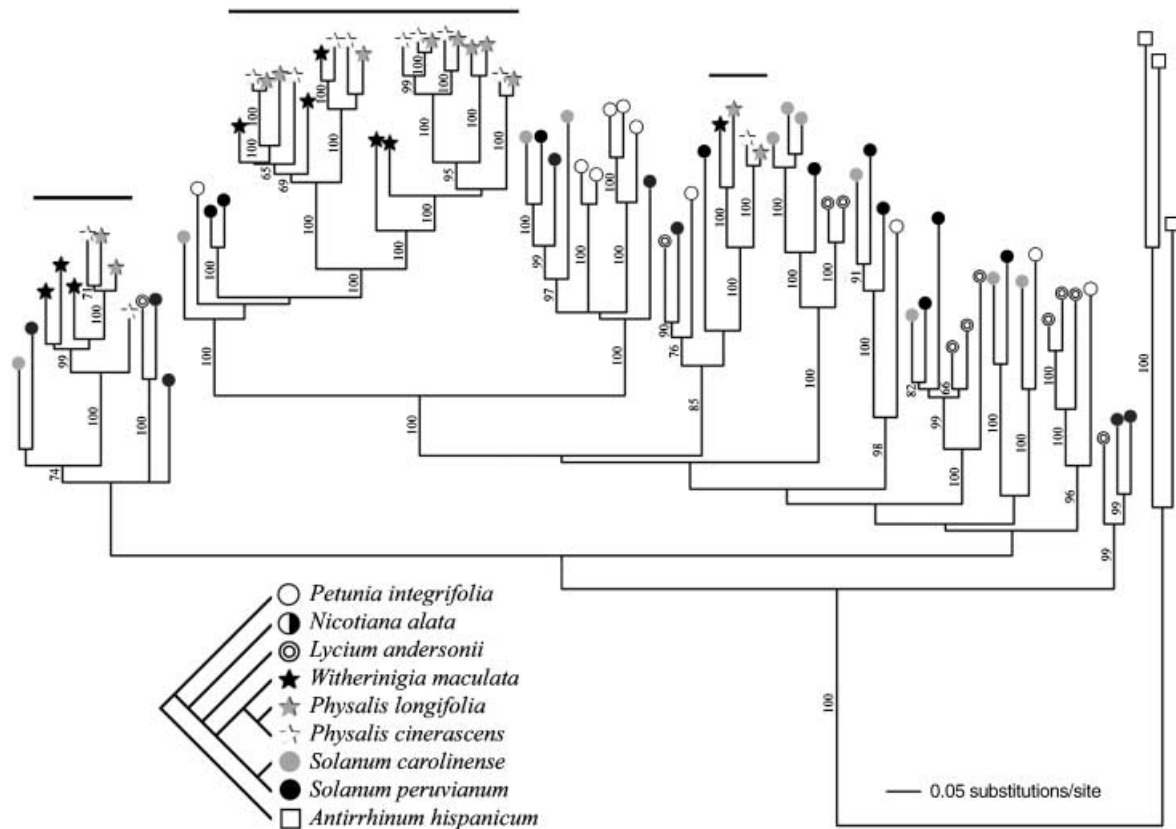


Fig. 2 A phylogenetic tree of S-alleles from 8 SI Solanaceae species shows extensive trans-generic polymorphism. The most likely tree obtained during an Markov Chain Monte Carlo search in Mr Bayes (Huelsenbeck & Ronquist, 2001) is shown. Posterior probabilities receiving > 60% support from 1600 trees kept after 1 million generations are shown above branches. Historically bottlenecked species (*Witheringia maculata*, *Physalis cinerascens*, and *P. longifolia*) possess alleles from three allelic lineages. The sequence data used for this analysis are from Richman *et al.* (1995), Golz *et al.* (1998), Richman & Kohn (1999), Wang *et al.* (2001), Lu (2001), and Kondo *et al.* (2002a). A representative, rather than exhaustive, sample of alleles and taxa were used to simplify the presentation. More exhaustive analyses (Richman & Kohn, 2000; Lu, 2001) confirm the results shown here.

clades comprised of alleles drawn from different genera. The last common ancestor of these genera is thought to have lived approx. 30–40 million yr ago (Ioerger *et al.*, 1990). As predicted by theory, diversifying selection succeeds in preventing complete lineage sorting over this entire time period and reciprocal monophyly of S-alleles is not observed among any pair of taxa.

Takahata (1993) pointed out the possibility of using loci under diversifying selection for 'paleo population genetic' purposes. The expected coalescent structure of polymorphism under several forms of diversifying selection, including that found at the gametophytic S-locus, closely resembles that of neutral polymorphism with the exception of a greatly increased time scale. This makes possible, at least theoretically, the inference of population sizes through time extending back tens of millions of years. In practice, such inference is made difficult by the fact that S-allele genealogies from individual taxa often fail to meet shape expectations of a constant birth-death process (Uyenoyama, 1997; Richman & Kohn, 2000). Nevertheless, severe and/or repeated population restrictions in the histories

of lineages are expected to leave a lasting impression on the S-locus. Such historical restrictions should be diagnosable as a reduction in the number of S-lineages followed by diversification of alleles from within this reduced set and should be observable for tens of millions of years following the event.

A putative historical bottleneck occurred in the common ancestor of the genera *Physalis* and *Witheringia*. All sampled species in these two genera possess S-alleles that are restricted to only three lineages, whereas all other sampled Solanaceae contain many more ancient S-lineages (Richman *et al.*, 1996b; Richman & Kohn, 2000; Lu, 2001). This implies that a reduction in S-allele diversity occurred before the divergence of *Physalis* and *Witheringia*. Subsequent diversification of the remaining three S-lineages resulted in current species in these two genera having numbers of alleles similar to those of other Solanaceae, but these alleles show much lower levels of molecular divergence. Thus estimates of current population sizes based on allele numbers for species in the genera *Physalis* and *Witheringia* are unremarkable among Solanaceae but long-term population sizes estimated from the number of ancient

S-lineages are at least an order of magnitude smaller (Richman *et al.*, 1996b; Richman & Kohn, 2000).

It is possible that forces other than bottlenecks could have caused the observed pattern. For instance, Uyenoyama (1997, 2003) suggests that genetic load linked to the obligately heterozygous S-locus could reduce the rate at which new alleles enter populations slowing the rate of lineage sorting even further than expected under negative frequency-dependent selection. Under this view, the ancestor of *Physalis* and *Witheringia* lost the load associated with at least some S-alleles, which then proliferated more rapidly causing higher rates of lineage sorting. Uyenoyama (1997) suggests, however, that the loss of load itself may have been caused by fixation of deleterious recessives held in common by the few alleles remaining after a population bottleneck. It is also conceivable that certain S-lineages evolved to a molecular space that made mutations to new specificities more likely and these proliferated. However, because separate lineages of S-alleles appear to have initiated rapid diversification at approximately the same time (Fig. 2; Richman, 2000), a bottleneck remains the best explanation of the observed pattern.

It may always be difficult to tie a genetic pattern created millions of years ago, such as the reduction of molecular diversity at the S-locus in *Physalis* and *Witheringia*, to a specific cause such as a bottleneck. Nevertheless, one should not lose sight of the more general observation that only one such putative event is observed in the current sample. There is no evidence for severe restrictions at the S-locus throughout the history of the vast majority of SI Solanaceae. If severe bottlenecks were commonly associated with speciation, we would expect to see increased lineage sorting at the S-locus. Thus it would appear that most speciation events do not entail a severe restriction in population size. A major caveat, however, is that if bottlenecks are accompanied by a complete loss of incompatibility, such events would be undetectable by analyses of S-allele diversity. All conversions to self-compatibility are expected to collapse polymorphism at the S-locus, whether or not such conversions are accompanied by restrictions in population size. Even if SI is not lost, little evidence of an historical bottleneck at the S-locus will be detectable unless bottlenecks are severe, prolonged, and subsequent gene flow is prevented. This is because even rather small populations can harbor substantial S-diversity (a population of size 100 harbors 6 alleles at equilibrium) and it may take dozens of generations to reach equilibrium allele numbers following a reduction in population size (Crosby, 1966). Finally, because gene flow at the S-locus is enhanced by frequency-dependent selection, subsequent gene flow will tend to preferentially draw in alleles from other populations if S-allele number is below equilibrium. The pattern of trans-specific evolution at the S-locus observed in the Solanaceae shows that such severe population restrictions have not occurred during the historical diversification leading to the vast majority of SI taxa.

The history of self-incompatibility in the Solanaceae

Phylogenetic analyses of S-RNases from the Solanaceae, Scrophulariaceae and Rosaceae, together with all available plant RNases of the same gene family, find that RNase-based GSI is ancestral in approximately three quarters of eudicot families (Igic & Kohn, 2001; Steinbachs & Holsinger, 2002). This implies that RNase-based SI is ancestral even in lineages that have lost this form of SI and subsequently acquired a new mechanism for self-recognition. Consequently, several other forms of SI have independent origins, including the well-characterized, nonhomologous mechanism that operates in the Brassicaceae (Igic & Kohn, 2001; Nasrallah, 2002).

The loss of self-incompatibility is certainly reversible on time scales equal to the divergence of families with nonhomologous SI. However, transitions from SC to SI are expected to be rare within lineages roughly at the rank of families (Igic & Kohn, 2001; Steinbachs & Holsinger, 2002). The wide distribution, persistence, and occasional origination of nonhomologous forms of SI suggest that incompatibility plays an adaptive role in the evolution of plant lineages. Nevertheless, many species are self-compatible, and a close examination of extant taxa reveals a pattern of commingled SC and SI populations and species. This pattern is well documented in the three families with RNase-based SI (Whalen & Anderson, 1981; Ueda & Akimoto, 2001; Vieira & Charlesworth, 2002, B. Igic, unpublished), as well as at least six other families with uncharacterized or nonhomologous SI systems (Arroyo, 1981; Gibbs, 1990; Weller *et al.*, 1995; Goodwillie, 1999). This pattern suggests that SI systems are frequently lost but rarely regained within families. Independent evidence for this is found in molecular genetic studies that suggest involvement of many genes and existence of many routes leading to the loss of SI (Kondo *et al.*, 2002b; Stone, 2002; Tsukamoto *et al.*, 2003a,b).

Frequent loss relative to gain of SI begs the question of how SI persists in the face of extreme inequality of character state transition rates. Here we use data from the Solanaceae to demonstrate a general method to reconstruct the evolutionary history of self-incompatibility in groups where the molecular basis of SI is at least partially known. We calculate the transition rates between SI and SC character states, and conduct a test that fails to reject the hypothesis of irreversibility of loss of SI. Finally we discuss the implications of irreversibility with respect to the continued persistence of SI and outline a method for testing the effect of a binary character state (SI vs SC) on diversification rates.

Unique properties in ancestral state reconstruction

Reconstructions of ancestral states often attempt to infer the evolutionary path of a complex character that is frequently lost and difficult to gain (Kohn *et al.*, 1996; Omland, 1997;

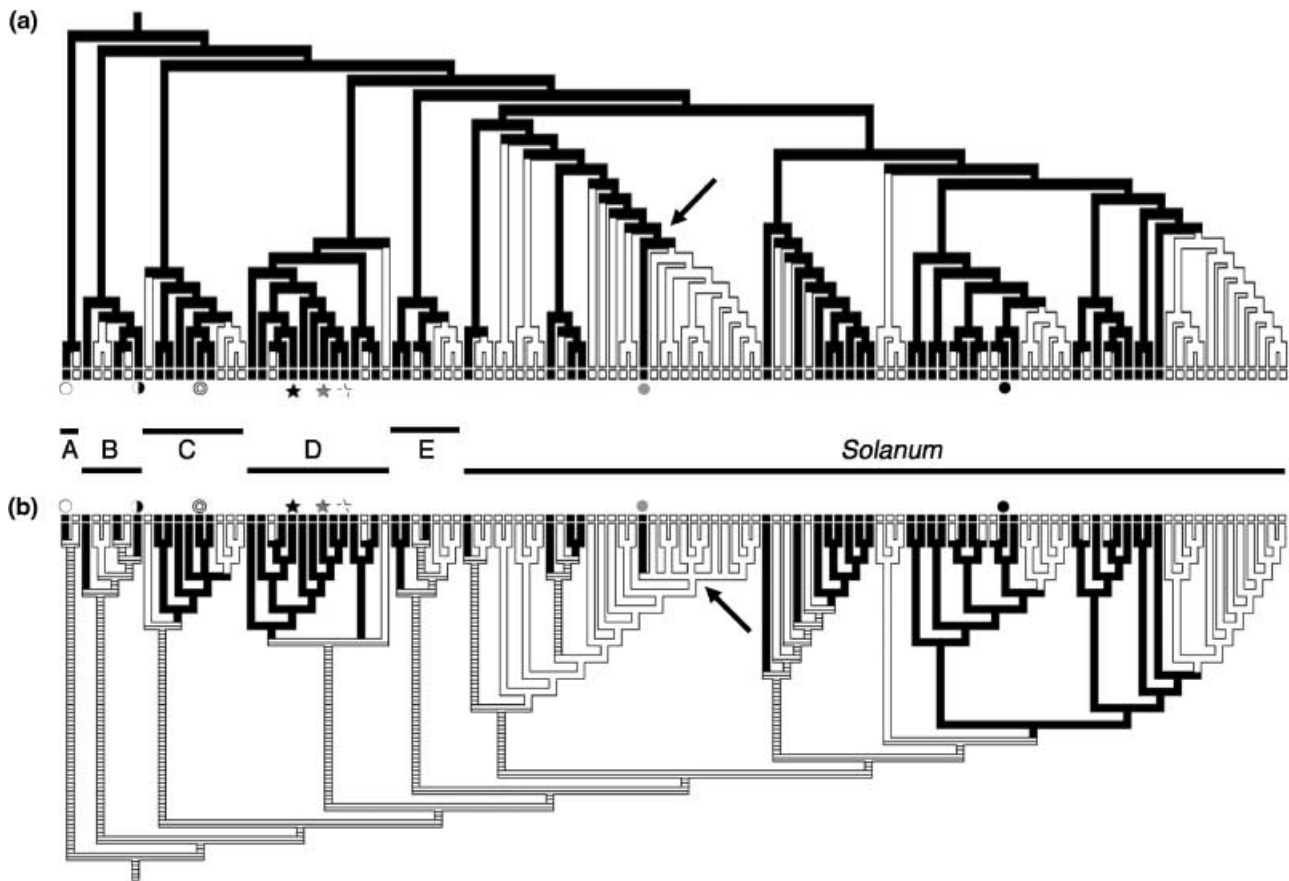


Fig. 3 Solanaceae species cladogram, SI lineages in black, SC in white, uncertain character state hatched. Bottom: Unweighted analysis (1 : 1; gain and loss of SI equally weighted) produces a minimum of 2 reversals. Species labelled with circles and stars appear in Fig. 2 and show *trans*-generic S-allele sharing. Top: Parsimony analysis with 5 : 1 weight applied to favor loss of SI. This is the minimum weighting needed to connect, with uninterrupted ancestry of SI, all SI taxa showing *trans*-specific S-polymorphism. This weighting produces no inferred reversals from SC to SI. Parsimony analyses were conducted in MacClade 4.0 (Maddison & Maddison, 2000). See text for results of maximum likelihood (ML) analyses. Black arrows show an important polytomy and its conservative resolution used in the ML analyses presented. Labeled groups follow Olmstead *et al.* (1999) and are encoded as follows: A, subfam. Petunioidae; B, subfam. Nicotianoideae; C–E, subfam. Solanoideae, with C, (tribes Lycieae + Hyoscyameae); D, (tribe Datureae + genus *Jaltomata*); E, tribe Capsiceae; *Solanum* (genus).

Cunningham, 1999). Such studies generally rely on arbitrary weighting schemes to find the transition parameters and ancestral states of characters. Because the ancestral states are usually unknown, studies of character history cannot objectively or accurately evaluate the application of various transition weights (in a parsimony framework) or accurately estimate transition parameters (maximum likelihood (ML)). *Trans*-specific polymorphism can be used to objectively assign ancestral states to many internal nodes of a species-level phylogeny and assess the transition rates between SI and SC. The knowledge of states at internal nodes, lacking from most current studies, is akin to possession of fossil character states, considered a critical component for accurate assessment of character evolution (Cunningham, 1999; Takebayashi & Morrell, 2001).

For reconstruction of the history of incompatibility, we rely heavily on the system's unorthodox preservation of polymor-

phism. An important inference from the observed *trans*-generic sharing of S-alleles is that SI has been present continuously since the common ancestor of all sampled SI species (Fig. 3, marked species). When SI is broken down, whether by mutation in the S-locus or an unlinked modifier locus (Stone, 2002), the previously numerous S-alleles become effectively neutral, and are expected to be lost by drift. Because the S-locus requires three alleles to function, or else all individuals are mutually incompatible, the loss of polymorphism at the S-locus is expected to make the regain of RNase-based GSI difficult. Fig. 2 records only a single case of recovery from restricted numbers of S-lineages (the common ancestor of *Physalis*/*Witheringia* is represented by only three lineages) and no cases of rediversification from fewer than three lineages. Re-invention of the same system after fixation on a single allele is an unlikely proposition, but should it ever have occurred in the Solanaceae, it would stand in sharp contrast

to the observed pattern of trans-specific polymorphism (Fig. 2). We use trans-specific evolution at the S-locus, and the implied continuous presence of SI connecting sampled species, to infer with certainty the ancestral states of deep nodes. SI therefore offers a unique opportunity for reconstructing the evolution of a binary mating system character, with SI or SC states, and assessment of the effect of each character state on patterns of diversification.

Future studies of the S-locus may provide for a finer scale of analysis. Loss of SI reduces the mean coalescence time of S-locus polymorphism to $4N$ generations. Because the SC allele is assumed to be favored during SI loss, we expect this time is often an overestimate. Thus loss of SI is expected to result in fixation on a single S-allele in a comparatively short time. If closely related self-compatible taxa are fixed on the same S-allele, they likely share a single ancestral transition to the SC condition. However, if they are fixed on different S-alleles, then multiple transitions from one or more SI ancestors may have occurred. Therefore, studies of the S-locus in SC taxa could provide information on both phylogenetic relationships and the numbers of transitions from SI to SC among a group of related species.

Available studies corroborate the expectation of rapid loss of S-allele diversity following transition to SC (Golz *et al.*, 1998; Kondo *et al.*, 2002a), but fixation is not always observed (Kondo *et al.*, 2002a; Tsukamoto *et al.*, 2003b) perhaps as a result of population structure and repeated recent transitions to SC. These observations complicate the pattern of S-locus evolution during and after the transition to SC. For instance, the finding of more than one SC allele within and among SC *Solanum* section *Lycopersicum* species (Kondo *et al.*, 2002a) could result from multiple independent transitions to SC or incomplete fixation before diversification of SC taxa. Recent diversification of this group makes it difficult to establish whether or not the SC species are monophyletic, further clouding inference. At present, it seems reasonable to assume that if SC species share the same S-allele, they share a single historical transition to SC but the converse may not always be true and further work on the fate of S-alleles in SC species is needed.

Data from the Solanaceae: irreversible loss of SI

We collected a database of incompatibility character states from 410 species, or roughly 20% of Solanaceae (primarily from Whalen & Anderson (1981)). We also constructed a composite tree (Fig. 3) made from a number of recent molecular phylogenies of Solanaceae. The composite tree consists of a skeletal family phylogeny (Olmstead *et al.*, 1999; Bohs & Olmstead, 2001), augmented by insertion of clades obtained from detailed phylogenetic studies of established monophyletic groups (Spooner & Systma, 1992; Anderson *et al.*, 1996; Kardolus *et al.*, 1998; Mace *et al.*, 1999; Miller & Venable, 2000; Bohs & Olmstead, 2001; Fukuda *et al.*, 2001;

Marshall *et al.*, 2001; Peralta & Spooner, 2001; Walsh & Hoot, 2001; L. Bohs, unpublished). The 120 taxa whose compatibility status is known and whose relationships can be inferred from existing molecular phylogenies are drawn in Fig. 3. The proportion of SI and SC species on the tree (39% SI, 57% SC, 4% polymorphic) is not significantly different from those in the 410 Solanaceae whose compatibility status is known ($P = 0.44$). Subsequent analyses required fully dichotomous trees. We resolved all polytomies in a way that would result in the fewest possible transitions from SI to SC – resolution that is conservative with the respect to the hypothesis of irreversibility. ML procedures require branch length information, which is unavailable in this composite tree. Therefore, we set all branches equal to unit length. Our composite tree is based on a variety of data sources which are not strictly comparable. If data for all taxa on the tree were analyzed simultaneously, a different topology might result. With this uncertainty in mind, we performed sensitivity analyses (Donoghue & Ackerly, 1996), which revealed that the obtained data are robust to modest changes in branching order (B. Igic, unpublished). Details of the methods used in this paper can be obtained from the authors.

Results of reconstructions of transition parameters and ancestral states are presented in Table 1. If molecular data are ignored, and the ancestral states are reconstructed based solely on the character states of extant taxa, maximum likelihood analyses reject the one-rate model ($-2 \Delta L = 5.69$, $P < 0.02$) and the irreversibility of loss of SI ($-2 \Delta L = 22.07$, $P < 0.001$). In other words, the use of unrestricted separate forward (SI to SC, β) and reverse (SC to SI, α) transition rates fits significantly better than use of a single parameter ($\alpha = \beta$). More importantly, however, use of the unrestricted two-parameter

Table 1 Log-likelihood tests for parameter estimates obtained from enforcement of restricted models compared with an unrestricted two-parameter model

	α	β	Log-likelihood	<i>P</i> -value
(a) Parameters and likelihood values were calculated from a tree whose internal states connecting molecularly characterized SI species were not fixed				
Unrestricted	0.093	0.218	-143.450	-
$\alpha = \beta$	0.188	0.188	-146.295	< 0.02
$\alpha = 0$	0	0.222	-154.484	< 0.001
(b) Parameters and likelihood values calculated from a tree whose internal states connecting molecularly characterized SI species were fixed				
Unrestricted	0.0005	0.2213	-154.51	-
$\alpha = \beta$	0.1897	0.1897	-170.35	<< 0.001
$\alpha = 0$	0	0.2216	-154.52	0.888 NS

Analyses were performed in *Discrete* (Pagel, 1999). Model restrictions are listed in the first column. $\alpha =$ SC to SI transition rate; $\beta =$ SI to SC transition rate

model is also significantly better than irreversibility ($\alpha = 0$, unrestricted β). Thus, irreversibility is not supported using this model. Consideration of the molecular genetic data, in particular the knowledge that SI was present in all nodes connecting the extant SI taxa with trans-specific polymorphism at the S-locus, is expected to provide improved accuracy of reconstruction of ancestral states and transition rates. This analysis also strongly rejects the one-parameter model ($-2 \Delta L = 31.68$, $P < 0.001$), but cannot reject irreversibility ($-2 \Delta L = 0.02$, $P = 0.888$). Therefore, failure to consider the molecular genetic data results in a false rejection of irreversibility.

Parsimony reconstruction also suggests that equal weight of gain and loss of SI is untenable. Equal weighting results in reconstruction of repeated gains of SI. Fig. 3 shows the results of the application of equal weights (1 : 1) and minimum unequal weight required for recovering an unbroken line of SI leading to taxa with known trans-specific lineages.

Although transition to SC was long thought to be common (Stebbins, 1974), quantitative empirical evidence has lagged. While there is support for occasional origination of non-homologous SI in various angiosperms, transitions to SC appear to be remarkably common, by comparison. Therefore, while loss of SI has been essentially irreversible in the Solanaceae, it has obviously not been strictly irreversible at deeper taxonomic levels. While the relatively closely related Scrophulariaceae and distantly related Rosaceae share a homologous SI system with the Solanaceae, the Convolvulaceae, a family sister to Solanaceae, contains many SI species with nonhomologous sporophytic incompatibility (Kowiyama *et al.*, 1980). If the common ancestor of Convolvulaceae was SC, its SI system represents a gain. However, measured as a rate, the loss of SI remains much more common than its reversal, so as to be considered *de facto* irreversible in higher lineages.

The consequences of SI: a deterministic model

The above analysis implies that SI was ancestral in Solanaceae, with transitions to SC common and irreversible, and that SC species currently outnumber SI species in the Solanaceae. Two explanations for the observed pattern (Fig. 3) are possible. First, the pattern might imply that SI and SC are not in equilibrium, such that SI, while ancestral, is headed for extinction. This hypothesis is difficult to reconcile with the reconstructed history of RNase-based GSI. Second, long-term persistence of SI, in the face of constant attrition as a result of largely irreversible transitions to SC, may imply that SI lineages have an advantage in evolutionary time, measured by an increased net diversification (speciation minus extinction) rate relative to SC lineages.

Given that SC lineages may contain primarily outcrossing taxa, the detection of a comparative advantage for SI lineages would be remarkable. Such a conservative comparison of the macroevolutionary effects of outcrossing is expected to suffer from a lack of statistical power as a result of large variance in

outcrossing rates among SC taxa. However, a large sample of taxa may offset this potential weakness.

A simple deterministic model (A. Lande *et al.*, unpublished) suggests conditions under which a constant proportion of SI and SC species will be maintained at equilibrium and those under which SI decreases towards extinction. Assuming that SI loss is irreversible ($\alpha = 0$; see above), and allowing exponential growth of species numbers, the growth rates for species numbers of each character state are as follows:

$$dN_{SC}/dt = r_{SC}N_{SC} + \beta N_{SI}$$

$$dN_{SI}/dt = r_{SI}N_{SI} - \beta N_{SI}$$

where N_{SC} and N_{SI} are numbers of SC and SI species, r_{SI} and r_{SC} are net diversification rates of SI and SC species, with a transition rate β (SI to SC). There exists an equilibrium proportion of SI species, $p_{SI} = 1 - \beta / (r_{SI} - r_{SC})$, if $r_{SI} > r_{SC} + \beta$. Equilibrium conditions require that the net diversification rate of SI species is larger than the sum of the net diversification rate of SC and the transition rate β . Otherwise, if $r_{SI} < r_{SC} + \beta$, then the proportion of SI species declines to 0. This simple model suggests that future studies should aim to determine transition and diversification rate parameters to test whether SI is destined for extinction or equilibrium frequency. In particular, lineages-through time approaches (Harvey *et al.*, 1994; Nee *et al.*, 1994; Pybus & Harvey, 2000) and Pagel's method of detecting lineage-specific speciation rates (Lutzoni & Pagel, 1997; Pagel, 1997) may be used to test whether or not the equilibrium hypothesis is tenable and to find if rates of speciation and extinction differ for the two character states.

We hypothesize that the presence of SI is unlikely to increase the speciation rate because populations of selfing taxa experience higher levels of genetic isolation, an aid to differentiation and speciation. Rather, we suspect that incompatibility increases levels of genetic variation in taxa that express it and also increases the cohesiveness of species, helping to maintain effectively high gene flow in the face of geographical isolation. This may serve to substantially decrease the extinction rate of SI relative to SC taxa. Such a hypothesis would be consistent with the general finding that SI taxa do not outnumber SC taxa in groups in which SI occurs (e.g. the Solanaceae; Heilbut, 2000) and could also explain the long term persistence of incompatibility in the face of constant and near-irreversible transitions to SC.

Acknowledgements

We thank A. Angert, T. Case, R. Lande, T. Near, T. Price, and K. Roy for valuable discussions. P. Bernhardt, R. Hanneman, T. Holtsford, J. Miller, T. Mione, M. Nee, R. Olmstead, S. Smith, M. Whitson provided expertise on various groups of Solanaceae. This work was funded in part by NSF DEB 0108173 to J.R.K. and DEB-0309184 to B.I. and J.R.K.

References

- Anderson MA, Cornish EC, Mau S-L, Williams EG, Hoggart R, Atkinson A, Bonig I, Grego B, Simpson RJ, Roche PJ, Haley J, Penschow J, Niall HD, Tregear G, Coghlan J, Crawford R, Clarke AE. 1986. Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana glauca*. *Nature* 321: 38–44.
- Anderson GJ, Jansen RK, Kim Y. 1996. The Origin and Relationships of the Pepino, *Solanum muricatum* (Solanaceae): DNA Restriction Fragment Evidence. *Economic Botany* 50: 369–380.
- Arroyo MTK. 1981. Breeding systems and pollination Biology in Leguminosae. In: Polhill RM and Raven PH, eds. *Advances in Legume Systematics*. Royal Botanic Gardens, Kew, 723–769.
- Bohs L, Olmstead RG. 2001. A reassessment of *Normania* and *Triguera* (Solanaceae). *Plant Systematics and Evolution* 228: 33–48.
- Clark AG, Kao T-H. 1994. Self-incompatibility: theoretical concepts and evolution. In: Williams EG, Clarke AE, Knox RB, eds. *Genetic control of self-incompatibility and reproductive development in flowering plants*. Boston, MA, USA: Kluwer Academic Publishers, 220–242.
- Crosby JL. 1966. Self-incompatibility alleles in the population of *Oenothera organensis*. *Evolution* 20: 567–579.
- Cunningham CW. 1999. Some limitations of ancestral character-state reconstruction when testing evolutionary hypotheses. *Systematic Biology* 48: 665–674.
- Darwin C. 1876. *The effects of cross and self fertilisation in the vegetable kingdom*. London, UK: John Murray.
- Donoghue MJ, Ackerly DD. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 351: 1241–1249.
- Emerson S. 1939. A preliminary survey of the *Oenothera organensis* population. *Genetics* 24: 524–537.
- Franklin-Tong VE, Franklin FCH. 2003. The different mechanisms of gametophytic self-incompatibility. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 358: 1025–1032.
- Fukuda T, Yokoyama J, Ohashi H. 2001. Phylogeny and biogeography of the genus *Lycium* (Solanaceae): Inferences from chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* 19: 246–258.
- Gibbs PE. 1990. Self-incompatibility in flowering plants: a neotropical perspective. *Revista Brasileira de Botânica* 13: 125–136.
- Golz JF, Clarke AE, Newbigin E, Anderson E. 1998. A relic S-RNase is expressed in the styles of self-compatible *Nicotiana glauca*. *Plant Journal* 16: 591–599.
- Goodwillie C. 1999. Multiple origins of self-compatibility in *Limnanthus* section *Leptasiphon*. Phylogenetic evidence from ITS sequence data. *Evolution* 53: 1387–1395.
- Harvey PH, May RM, Nee S. 1994. Phylogenies without Fossils. *Evolution* 48: 523–529.
- Heilbut JC. 2000. Lower species richness in dioecious clades. *American Naturalist* 156: 221–241.
- Hiscock SJ, Tabah DA. 2003. The different mechanisms of sporophytic self-incompatibility. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 358: 1037–1045.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.
- Igic B, Kohn JR. 2001. Evolutionary relationships among self-incompatibility RNases. *Proceedings of the National Academy of Sciences, USA* 98: 13167–13171.
- Ioerger TR, Clark AG, Kao T-H. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proceedings of the National Academy of Sciences, USA* 87: 9732–9735.
- Kardolus JP, van Eck HJ, van den Berg RG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy. (*Solanaceae*) *Plant Systematics and Evolution* 210: 87–103.
- Kohn JR, Graham SW, Morton B, Doyle JJ, Barrett SCH. 1996. Reconstruction of the evolution of reproductive characters in Pontederiaceae using phylogenetic evidence from chloroplast DNA restriction variation. *Evolution* 50: 1454–1469.
- Kondo K, Yamamoto M, Itahashi R, Sato T, Egashira H, Hattori T, Koyama Y. 2002a. Insights into the evolution of self-compatibility in *Lycopersicon* from a study of stylar factors. *Plant Journal* 30: 143–153.
- Kondo K, Yamamoto M, Matton DP, Sato T, Hirai M, Norioka S, Hattori T, Koyama Y. 2002b. Cultivated tomato has defects in both S-RNase and HT genes required for stylar function of self-incompatibility. *Plant Journal* 29: 627–636.
- Koyama Y, Shimano N, Kawase T. 1980. Genetic analysis of incompatibility in the diploid *Ipomoea* species closely related to the sweet potato. *Theoretical and Applied Genetics* 58: 149–155.
- Lawrence MJ. 2000. Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Annals of Botany* 85: 221–226.
- Lu Y. 2001. Roles of lineage sorting and phylogenetic relationships in the genetic diversity at the self-incompatibility locus of Solanaceae. *Heredity* 86: 195–205.
- Lutzoni F, Pagel M. 1997. Accelerated evolution as a consequence of transitions to mutualism. *Proceedings of the National Academy of Sciences, USA* 94: 11422–11427.
- Mace ES, Gebhardt CG, Lester RN. 1999. AFLP analysis of genetic relationships in the tribe Daturae (Solanaceae). *Theoretical and Applied Genetics* 99: 634–641.
- Marshall JA, Knapp S, Davey MR, Power JB, Cocking EC, Bennett MD, Cox AV. 2001. Molecular systematics of *Solanum* section *Lycopersicon* (*Lycopersicon*) using the nuclear ITS region. *Theoretical and Applied Genetics* 103: 1216–1222.
- McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE. 1989. Style self-incompatibility gene products of *Nicotiana glauca* are ribonucleases. *Nature* 342: 955–957.
- Miller JS, Venable DL. 2000. Polyploidy and the evolution of gender dimorphism in plants. *Science* 289: 2335–2338.
- Nasrallah JB. 2002. Recognition and rejection of self in plant reproduction. *Science* 296: 305–308.
- Nee S, May RM, Harvey PH. 1994. The Reconstructed Evolutionary Process. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 344: 305–311.
- Olmstead RG, Sweere JA, Spangler RE, Bohs L, Palmer JD. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: Nee M, Symon DE, Lester RN, Jessop JP, eds. *Solanaceae IV*. Richmond, UK: Royal Botanic Gardens, Kew, 111–137.
- Omland KE. 1997. Examining Two Standard Assumptions of Ancestral State Reconstruction: Repeated Loss of Dichromatism in Dabbling Ducks (Anatini). *Evolution* 51: 1636–1646.
- Pagel M. 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta* 26: 331–348.
- Pagel M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* 48: 612–622.
- Peralta IE, Spooner DM. 2001. Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* Mill. Wettst. subsection *Lycopersicon*). *American Journal of Botany* 88: 1888–1902.
- Pybus OG, Harvey PH. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London Series B-Biological Sciences* 267: 2267–2272.
- Richman A. 2000. Evolution of balanced genetic polymorphism. *Molecular Ecology* 9: 1953–1963.
- Richman AD, Kao TH, Schaeffer SW, Uyenoyama MK. 1995. S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* 75: 405–415.
- Richman AD, Kohn JR. 2000. Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Molecular Biology* 42: 169–179.

- Richman AD, Kohn JR. 2000. Self-incompatibility alleles from *Physalis*: Implications for historical inference from balanced genetic polymorphisms. *Proceedings of the National Academy of Sciences, USA* 96: 168–172.
- Richman AD, Uyenoyama MK, Kohn JR. 1996a. S-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. *Heredity* 76: 497–505.
- Richman AD, Uyenoyama MK, Kohn JR. 1996b. Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. *Science* 273: 1212–1216.
- Sassa H, Hirano H, Ikehashi H. 1993. Identification and characterization of stylar glycoproteins associated with self-incompatibility genes in Japanese pear, *Pyrus serotina* Rehd. *Molecular and General Genetics* 242: 17–25.
- Spooner DM, Systma KJ. 1992. Reexamination of Series Relationships of Mexican and Central American Wild Potatoes (*Solanum* sect. *Petota*): Evidence from Chloroplast DNA Restriction Site Variation. *Systematic Botany* 17: 432–448.
- Stebbins GL. 1974. *Flowering plants: evolution above the species level*. Cambridge, MA, USA: Harvard University Press.
- Steinbachs JE, Holsinger KE. 2002. S-RNase-mediated Gametophytic Self-incompatibility is Ancestral in Eudicots. *Molecular Biology and Evolution* 19: 825–829.
- Stone JL. 2002. Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. *Quarterly Review of Biology* 77: 17–32.
- Takahata N. 1990. A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proceedings of the National Academy of Sciences, USA* 87: 2419–2423.
- Takahata N. 1993. Evolutionary genetics of human paleopopulations. In: Takahata N, Clarke AG, eds. *Mechanisms of molecular evolution*. Sunderland, MA, USA: Sinauer Associates, 1–21.
- Takebayashi N, Morrell PL. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany* 88: 1143–1150.
- Tsukamoto T, Ando T, Kokubun H, Watanabe H, Sato T, Masada M, Marchesi E, Kao T-H. 2003a. Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by a modifier locus that suppresses the expression of an S-RNase gene. *Sexual Plant Reproduction* 15: 255–263.
- Tsukamoto T, Ando T, Takahashi K, Omori T, Watanabe H, Kokubun H, Marchesi E, Kao T-H. 2003b. Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. *Plant Physiology* 131: 1903–1912.
- Ueda Y, Akimoto S. 2001. Cross- and self-compatibility in various species of the genus *Rosa*. *Journal of Horticultural Science and Biotechnology* 76: 392–395.
- Uyenoyama MK. 1997. Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. *Genetics* 137: 1389–1400.
- Uyenoyama MK. 2003. Genealogy dependent viability among self-incompatibility genotypes. *Theoretical Population Biology* 63: 281–293.
- Vekemans X, Slatkin M. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics* 137: 1–9.
- Vieira CP, Charlesworth D. 2002. Molecular variation at the self-incompatibility locus in natural populations of the genera *Antirrhinum* and *Misopates*. *Heredity* 88: 172–181.
- Walsh BM, Hoot SB. 2001. Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: The chloroplast atpB-rbcL spacer region and nuclear *waxy* introns. *International Journal of Plant Sciences* 162: 1409–1418.
- Wang X, Hughes AK, Tsukamoto T, Ando T, Kao T-H. 2001. Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self-incompatibility (S) locus in *Petunia inflata*. *Plant Physiology* 125: 1012–1022.
- Weller SG, Donoghue MJ, Charlesworth D. 1995. The evolution of self-incompatibility in flowering plants: A phylogenetic approach. In: Hoch PC, Stephenson AG, eds. *Experimental and molecular approaches to plant biosystematics*. St Louis, MO, USA: Missouri Botanical Garden, 355–382.
- Whalen MD, Anderson GJ. 1981. Distribution of Gametophytic Self-Incompatibility and Infrageneric Classification in Solanum. *Taxon* 30: 761–767.
- Wright S. 1939. The distribution of self-sterility alleles in populations. *Genetics* 24: 538–552.
- Xue Y, Carpenter R, Dickinson H, Cohen ES. 1996. Origin of allelic diversity in *Antirrhinum* S locus RNases. *Plant Cell* 8: 805–814.



About New Phytologist

- *New Phytologist* is owned by a non-profit-making **charitable trust** dedicated to the promotion of plant science, facilitating projects from symposia to open access for our Tansley reviews. Complete information is available at www.newphytologist.org
- Regular papers, Letters, Research reviews, Rapid reports and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as-ready' via *OnlineEarly* – average first decisions are just 5–6 weeks. Essential colour costs are **free**, and we provide 25 offprints as well as a PDF (i.e. an electronic version) for each article.
- For online summaries and ToC alerts, go to the website and click on 'Journal online'. You can take out a **personal subscription** to the journal for a fraction of the institutional price. Rates start at £108 in Europe/\$193 in the USA & Canada for the online edition (click on 'Subscribe' at the website)
- If you have any questions, do get in touch with Central Office (newphytol@lancaster.ac.uk; tel +44 1524 592918) or, for a local contact in North America, the USA Office (newphytol@ornl.gov; tel 865 576 5261)