Expression and Trans-Specific Polymorphism of Self-Incompatibility RNases in *Coffea* (Rubiaceae)

Michael D. Nowak[1][*][¤], **Aaron P. Davis**[2], **François Anthony**[3], **Anne D. Yoder**[1]

1 Department of Biology, Duke University, Durham, North Carolina, United States of America, 2 Royal Botanic Gardens, Kew, Richmond, Surrey, United Kingdom, 3 Institut de Recherche pour le Développement, centre de Montpellier, Unité Mixte de Recherche "Résistance des Plantes aux Bioagresseurs", Montpellier, France

Abstract

Self-incompatibility (SI) is widespread in the angiosperms, but identifying the biochemical components of SI mechanisms has proven to be difficult in most lineages. *Coffea* (coffee; Rubiaceae) is a genus of old-world tropical understory trees in which the vast majority of diploid species utilize a mechanism of gametophytic self-incompatibility (GSI). The S-RNase GSI system was one of the first SI mechanisms to be biochemically characterized, and likely represents the ancestral Eudicot condition as evidenced by its functional characterization in both asterid (Solanaceae, Plantaginaceae) and rosid (Rosaceae) lineages. The S-RNase GSI mechanism employs the activity of class III RNase T2 proteins to terminate the growth of "self" pollen tubes. Here, we investigate the mechanism of *Coffea* GSI and specifically examine the potential for homology to S-RNase GSI by sequencing class III RNase T2 genes in populations of 14 African and Madagascan *Coffea* species and the closely related self-compatible species *Psilanthus ebracteolatus*. Phylogenetic analyses of these sequences aligned to a diverse sample of plant RNase T2 genes show that the *Coffea* genome contains at least three class III RNase T2 genes. Patterns of tissue-specific gene expression identify one of these RNase T2 genes as the putative *Coffea* S-RNase gene. We show that populations of SI *Coffea* are remarkably polymorphic for putative S-RNase alleles, and exhibit a persistent pattern of trans-specific polymorphism characteristic of all S-RNase genes previously isolated from GSI Eudicot lineages. We thus conclude that *Coffea* GSI is most likely homologous to the classic Eudicot S-RNase system, which was retained since the divergence of the Rubiaceae lineage from an ancient SI Eudicot ancestor, nearly 90 million years ago.

Introduction

Many bisexual flowering plants avoid the deleterious effects of inbreeding by employing genetically controlled self-incompatibility (SI) mechanisms to ensure outcrossing [1]. SI mechanisms provide the biochemical machinery necessary for plants to recognize and reject their own pollen as well as non-self pollen with a genotype sufficiently similar to elicit activation of the SI mechanism. SI plants thus require a pollen donor with a divergent genotype for successful fertilization, a mating system known as "obligate outcrossing" [2]. SI plays an important role in shaping the spatial and temporal distribution of genetic diversity in plant populations and is thought to influence patterns of lineage diversification in clades within which these mechanisms are utilized [3,4].

It is likely that more than half of all flowering plant species employ biochemical SI mechanisms [5,6], but identifying the loci involved in SI (S-loci) has proven to be difficult [7]. In flowering plants, the biochemical components of only three SI mechanisms have been well-characterized (see [8] for review): 1) the Papaveraceae calcium signaling system [9–11]; 2) the Brassicaceae sporophytic SRK and SP11/SCR SI system [12,13]; and 3) the Eudicot gametophytic S-RNase system [14]. The unique biochemical mechanisms characterized in both the Papaveraceae and Brassicaceae point to

independent evolution of these SI systems in the respective ancestors of these families. In contrast, the S-RNase system is known to function in both of the two major Eudicot clades: the asterids (i.e. Solanaceae and Plantaginaceae), and the rosids (i.e. Rosaceae). Furthermore, there is now reliable phylogenetic evidence for homology of the S-RNase gene between these three families that strongly supports a single evolutionary origin for the S-RNase SI mechanism in an ancient Eudicot lineage ancestral to both asterids and rosids [15–17]. The S-RNase SI mechanism is a type of gametophytic self-incompatibility (GSI) that functions through the selective deactivation of RNase cytotoxicity in non-self pollen tubes as they grow through the style. The S-locus in Eudicot species utilizing the S-RNase GSI system carries genes with reproductive tissue-specific expression patterns (i.e. pistil or pollen tube; [18–22]). The S-locus gene expressed in the pistil (i.e. the female floral organ) was the first component of the S-locus to be identified [23], and was found to encode a ribonuclease in the RNase T2 protein family [24]. While species in at least 23 other Eudicot families are known to employ some mechanism of GSI [3], there has been a conspicuous lack of published data from any of these families that would demonstrate utilization of the ancestral Eudicot S-RNase SI mechanism.

We present here the first molecular evidence for the utilization of the ancestral Eudicot S-RNase system in the asterid genus *Coffea*

L. (Rubiaceae). *Coffea* is a genus of substantial agricultural and commercial importance [25], with two species, *C. arabica* L. (Arabica) and *C. canephora* A. Froehner (Robusta), used in the production of the beverage coffee. There are approximately 103 species in the genus, most of which are understory trees, distributed in the tropical belt and found in Africa, Madagascar, the Mascarene Islands, and the Comoros Islands [26,27]. The vast majority of *Coffea* species are known to exhibit a strong GSI response, but three African species (*C. arabica*, *C. anthonyi* Stoff. & F.Anthony, and *C. heterocalyx* Stoff.) are exceptional for their ability to self-fertilize (i.e. self-compatibility, SC) [26,28,29]. Parenthetically, the genus *Psilanthus*, a genus of ca. 22 species that is very closely allied to *Coffea* [30], is known to contain at least one SC species, viz. *P. ebracteolatus* Hiern [31]. *C. arabica* is the allotetraploid product of recent hybridization between the two African species *C. canephora* (west and central Africa) and *C. eugenoides* (central Africa) [30,32,33]. Self-compatibility in *C. arabica* is thus not surprising given the strong correlation between polyploidy and the breakdown of SI previously documented in other angiosperm lineages (e.g. see [34–37]). In contrast to the unique polyploid origin of *C. arabica*, the cause of SC in diploid *C. heterocalyx* and *C. anthonyi* is unknown. Coulibaly et al. [38] exploited *C. heterocalyx* SC to genetically map loci implicated in the *C. canephora* GSI mechanism. A single 5.4 cM *Coffea* self-incompatibility locus (S-locus) was identified, but they were unable to fine map or sequence this locus to identify individual genes.

We present evidence from several sources of data that the *Coffea* GSI mechanism is homologous to the S-RNase GSI system employed by many Eudicot lineages. To assess the homology of the *Coffea* GSI mechanism to S-RNase GSI, we collected sequence polymorphism data from RNase T2 genes expressed in floral tissues from SI and SC *Coffea* species. We compare three specific characteristics of these data to that of previously isolated S-RNase genes. 1) Phylogenetic analyses consistently place all S-RNase genes within the Class III RNase T2 gene family. 2) Populations of SI plants harbor exceptional allelic diversity and polymorphism at the S-RNase locus. Phylogenetic analyses of S-RNase alleles isolated from related species show that allelic lineages are often shared between species, a pattern known as trans-specific polymorphism. Such phylogenetic patterns are thought to represent the maintenance of ancestral S-allelic diversity through strong negative frequency-dependant selection [39]. 3) Distinct from other Class III T2 RNases, expression of S-RNase genes is generally limited to floral tissues. Below, we describe in detail these characteristics of the S-RNase GSI system and the means by which they are employed to test for homology to the *Coffea* GSI mechanism.

Evidence for S-RNase GSI homology

S-RNase genes are members of the class III RNase T2 gene family. Plant RNase T2 proteins are endoribonucleases with selective activity expressed in a diverse array of tissues in response to physical or pathogen-induced stress [40,41]. RNase T2 proteins have been the subject of many detailed biochemical and functional studies [40,42], and contain several conserved active sites (CAS), which are amino acid motifs involved in the biochemistry of ligand-specific RNase activation [43]. Plant RNase T2 genes are grouped into three classes (i.e. class I, II, and III) based on conserved structural motifs and evolutionary history as defined by phylogenetic placement. Class I and II RNase T2 genes are either expressed in response to cellular stress (Class I) or constitutively expressed (Class II). Class III RNase T2 genes contain either one or two introns, and carry conserved amino acid motifs that have been shown to involved in activating RNase function [17,40,41]. All S-RNase genes known to function in the S-RNase

GSI system encode proteins in the Class III RNase T2 subfamily. S-RNase expression is generally limited to floral tissues [44,45], but Class III RNases that are unrelated to the GSI mechanism (i.e. those isolated from sequenced genomes and EST libraries of SC Eudicot species) tend to exhibit little tissue-specificity in expression and their cellular and biochemical function is generally unknown [41]. These genes have been termed "S-like" [17] or "proto-S-RNases" [41] due to domain structure similarity and phylogenetic placement within the Class III RNase T2 family. We employ phylogenetic analysis to test for the placement of putative *Coffea* S-RNase genes within the Class III RNase T2 family.

S-RNase genes exhibit deep allelic divergence and trans-specific polymorphism. S-RNase genes contain five conserved regions distinguished by specific amino acid motifs (denoted c1 to c5), and between c2 and c3 carry either one (Solanaceae and Plantaginaceae) or two (Rosaceae) regions of elevated sequence variation, or "hypervariable" regions (see Figure 1). Sites within the hypevariable regions appear to be experiencing positive selection (i.e. sites with dN/dS ratios >1.0), suggesting a possible implication in the biochemical trigger of self-recognition [39,46,47,48,49]. Beyond the extremely divergent hypervariable regions, S-RNase alleles display characteristically deep sequence divergence and exceptional levels of polymorphism in SI populations [50]. If *Coffea* GSI is homologous to the S-RNase GSI system, putative S-RNase alleles isolated from *Coffea* populations should show a level of sequence divergence and polymorphism consistent with known S-RNase alleles. Additionally, theoretical expectations and empirical evidence suggest that diploid SC *Coffea* species should show dramatically reduced S-allele polymorphism as the loss of SI will be accompanied by an extreme genetic bottleneck at the S-locus [51,52].

The S-RNase GSI mechanism places the S-locus under negative frequency-dependent selection [39,53,54], which means that the frequency of an S-allele is inversely related to that allele's effect on fitness. Individuals in an SI population carrying rare S-allele genotypes benefit from an abundant supply of potential mates (i.e. pollen donors or recipients) and thus experience a significant fitness advantage. The result of this negative frequency-dependant selection is that the fecundity of an individual within an SI population is inversely proportional to the frequency of that individual's S-alleles in the population. These conditions have important ramifications for the frequency distribution of S-alleles in finite SI populations: the probability that an allele will be fixed is zero, and while S-alleles will eventually be replaced by new specificities, the time-scale of this replacement is far greater than at neutral loci [39,55,56]. Thus, ancestral polymorphisms at the S-locus will be maintained in the population much longer than at neutral loci. As a result, phylogenetic analyses of S-alleles show a pattern of trans-specific polymorphism [39,50,57,58], similar to empirical studies of genes thought to be evolving under negative frequency-dependant selection in animals (e.g. primate MHC alleles; [59]), and fungi (e.g. mating pheromone receptors; [60]). If *Coffea* GSI is homologous to the Eudicot S-RNase GSI mechanism, putative S-RNase alleles from *Coffea* species should show a clear pattern of trans-specific polymorphism reflecting ancestral allelic diversity.

S-RNase expression is limited to pistil tissues. The existence of S-like class III RNase T2 genes in SC Eudicot lineages suggests that a more rigorous assay of homology is required beyond phylogenetic placement alone. While S-RNase genes are expressed primarily in female floral organs (i.e. the pistil; [44,45]), S-like RNase genes generally show little tissue-specific expression (i.e. they are expressed constitutively; [41]). Therefore, if we have correctly identified the *Coffea* S-RNase, we expect to find
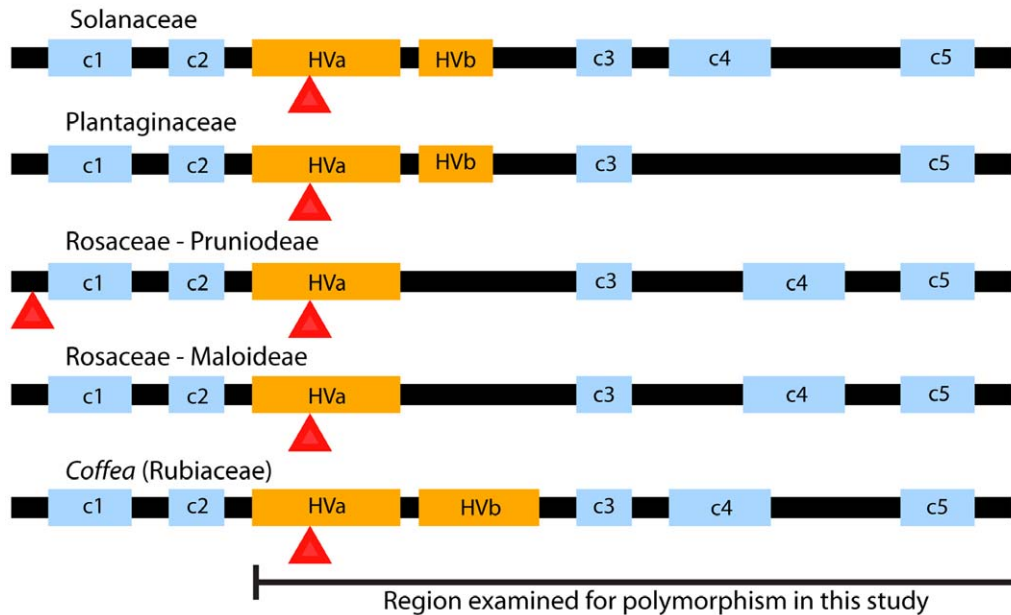
Figure 1. Eudicot S-RNase protein conserved domain structure. The conserved domain structures of S-RNases employed in GSI Eudicot lineages and the conserved domains revealed in this study of *Coffea* are labeled accordingly. Canonical S-RNase conserved regions are labeled "c1"– "c5", and hyper-variable regions are labeled "HVa" and "HVb". The red arrowheads below the protein constructs indicate intron positions. Intron position data and presence of conserved region c1 was validated in *Coffea* based on genomic sequence and 5′RACE data from *C. arabica* and *C. canephora* alone (data not shown). Results relating to RNase polymorphism reported in this study are based on 3′RACE data from the hyper-variable regions (i.e. HVa) to the end of transcripts (i.e. through c5). This figure was adapted from Vieira and Charlesworth [91].
doi:10.1371/journal.pone.0021019.g001

expression limited to pistil tissues, whereas putative S-like RNase T2 genes are unlikely to display the same tissue-specificity.

Results

African and Madagascan *Coffea* species exhibit strong gametophytic self-incompatibility

Microscopic observations of self pollen tube growth in *C. canephora* were found to be consistent with a gametophytic SI mechanism [38,61], but SI in all other *Coffea* species has been assumed by the observation of reduced seed set in self versus outcross pollinations [62,63]. While this is a valid metric for evaluating the strength of SI, inbreeding depression can also lead to significant reductions in seed set [1,64,65] which could confound the conclusions of these previous mating system experiments in *Coffea*. Therefore, to examine the strength of the GSI response in several species of *Coffea* (Table 1), we performed self-fertilizations and observed pollen tube growth microscopically with a standard aniline blue staining procedure [66]. Example micrographs of pollen tube growth resulting from self-compatible and incompatible fertilizations are shown in Figure 2. We were able to microscopically examine the GSI response in about 70% of the individual *Coffea* trees utilized in this study, including at least one representative of each species. All putatively SI *Coffea* species examined show a characteristic termination of self pollen tube growth with a bulbous mass of callose tissue after growing a few millimeters into the style tissue (Figure 2A)[67]. These results are consistent with a strong gametophytic self-incompatibility response as previously documented both in *C. canephora* [38] and other GSI species [9]. Figure 2B shows an example of self pollen tubes that are unhindered in their passage through the style, which was observed in the SC species *C. arabica, C. heterocalyx,* and *Psilanthus ebracteolatus*. These results are consistent with all previous reports of self-compatibility in these species [26,29].

The *Coffea* genome contains at least three Class III RNase T2 genes

In order to test if *Coffea* species utilize a GSI mechanism homologous to the canonical Eudicot S-RNase system, we cloned RNase T2 genes expressed in pistil tissues (i.e. the female floral organ) from 14 species of African and Madagascan *Coffea* and the related species *P. ebracteolatus*. Visual analysis of the sequences isolated from *C. arabica* revealed three distinct types of RNase T2 genes, which we refer to here as RNase "A", RNase "C", and S-RNase. We aligned all of the unique RNase T2 sequences (i.e. putative alleles) generated through our cloning strategy with a large sample of plant RNase T2 genes and performed a Bayesian phylogenetic analysis to examine the molecular evolution and class membership of RNase T2 genes in the *Coffea* genome. The resulting phylogenetic tree, shown in Figure 3, exhibits well-supported clades of Class I, II, and III T2 RNases. Furthermore, the placement of RNase T2 genes that we have isolated from the *Coffea* genome suggests that it contains at least three RNase T2 gene copies (i.e. RNase "A", RNase "C", and S-RNase), with unambiguous membership in the Class III RNase T2 subfamily as defined by Sassa et al. [24]. Additionally, RNase T2 cDNAs that we have isolated from *P. ebracteolatus* show clear homology to *Coffea* RNase "C" and S-RNase genes, but this species appears to lack a homolog to the *Coffea* RNase "A" gene.

All previously identified RNase genes implicated in the Eudicot S-RNase GSI system are members Class III RNase T2 family, and the proteins they encode are known to be relatively basic (i.e. isoelectric point, or $pI > 7.5$), while Class I and II RNase T2 proteins tend to be relatively acidic (i.e. $pI < 7.5$; [41]). To further evaluate the class III membership of *Coffea* RNase T2 genes, we calculated the predicted pI of all partial transcripts from conserved active site I (CAS I as defined by Irie [68]) to the end of the transcript. Consistent with class III RNase T2 membership, predicted proteins from *C. arabica* and *P. ebracteolatus* RNase "A",

Table 1. S-RNase alleles isolated from *Coffea* and *Psilanthus*.

Species	Distribution	SI/SC	Plants Sampled	S-RNase Alleles
C. andrambovatensis	E. Madagascar	SI	4	S1, S4, S8, S21, S36, S37
C. bonnieri	N. Madagascar	SI	3	S13, S25, S26, S27, S45
C. millotii	E. Madagascar	SI	8	S8, S10, S17, S18, S19, S32, S33
C. montis-sacri	E. Madagascar	SI	4	S9, S13, S18, S21, S24, S34
C. perrieri	W. Madagascar	SI	11	S11, S12, S13, S15, S16, S26
C. resinosa	E. Madagascar	SI	7	S1, S4, S10, S13, S25, S28, S29, S51
C. tricalysoides	N. Madagascar	SI	2	S11, S38, S39
C. tsirinananae	N. Madagascar	SI	5	S6, S7, S8, S9, S10, S49
C. brevipes	Tropical W. Africa	SI	1	S1
C. canephora	Tropical W. and central Africa	SI	3	S3, S17, S27, S41, S42
C. eugenoidoides	Tropical central Africa	SI	4	S1, S4, S13
C. pseudozanguebariae	Tropical E. Africa	SI	1	S7, S38
C. arabica	Tropical NE Africa	SC*	1	S1, S2**, S3
C. heterocalyx	Tropical W. Africa	SC	4	S12
P. ebracteolatus	Tropical W. Africa	SC	1	PS1

*SC likely the result of allotetraploid origin (i.e. 2n = 44).
**Short message length indicates that this is likely pseudogene copy.
doi:10.1371/journal.pone.0021019.t001

"C" genes are distinctly basic (8.85 and 8.43, respectively) and all putative S-RNase transcripts are also basic (mean pI = 8.97, standard deviation = 0.38).

Coffea S-RNase heterozygosity

Of the 60 individual *Coffea* trees sampled in our study, all but 10 specimens from GSI species (see above) were found to carry two putatively functional S-RNase alleles (i.e. assumed by the absence of premature stop codons or frame-shift mutations). It is likely that amplification efficiency will vary considerably when employing degenerate primers. The seeming homozygosity observed in 10 SI *Coffea* specimens is likely a product of over-representation of one of the two alleles actually carried by these individuals in the initial PCR, and thus simply a limitation of our cloning and amplification strategy rather than true S-RNase homozygosity. Supporting this claim, when we screened amplicon pools more deeply from one of these putatively homozygous individuals (i.e. >48 colonies

sequenced), a second S-RNase allele was found. All four specimens of the SC species *C. heterocalyx* were found to be homozygous for allele S12 despite relatively deep screening of amplicon pools (i.e. approximately 48 clones were sequenced) from two individual plants. The SC species *P. ebracteolatus* was also found to be homozygous for an S-RNase allele (PS1). The single accession of the tetraploid SC species *C. arabica* sampled in our study was found to carry three putative S-RNase alleles (i.e. S1, S2, and S3), but one of these (i.e. S2) was a short transcript that may represent a pseudogene copy.

Allelic divergence and patterns of polymorphism at the *Coffea* S-RNase

Putative S-RNase sequences isolated from samples of *Coffea* and *Psilanthus* were sorted into 36 putative allelic lineages based on nearly identical nucleotide composition, while a mean pairwise genetic distance of 0.248 (standard deviation = 0.064) was found
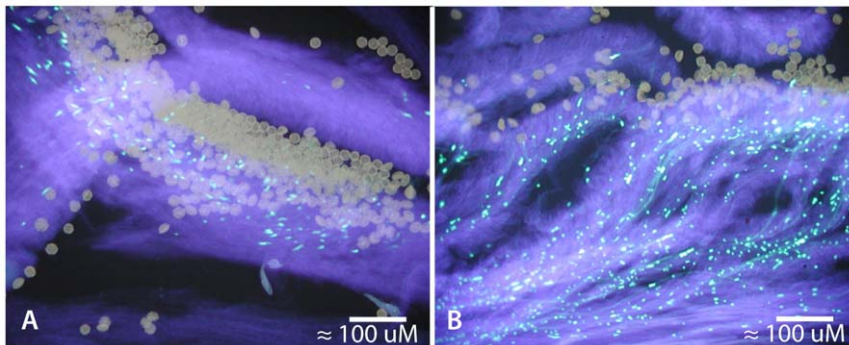
Figure 2. *Coffea* gametophytic self-incompatibility. The strength of the *Coffea* GSI response was measured by examining pollen tube growth using analine blue staining. A. Self-pollen tube growth in SI *C. perrieri* in cultivation at the Kianjavato Coffee Research Station, Madagascar. Pollen tubes terminate shortly after entering the style in bulbous callose plugs that fluoresce brightly. B. Self-pollen tube growth in SC *C. arabica* in cultivation at the Duke University Department of Biology Greenhouse Facility. Pollen tubes can be seen growing through the style tissue unhindered, leading to self-fertilization.
doi:10.1371/journal.pone.0021019.g002
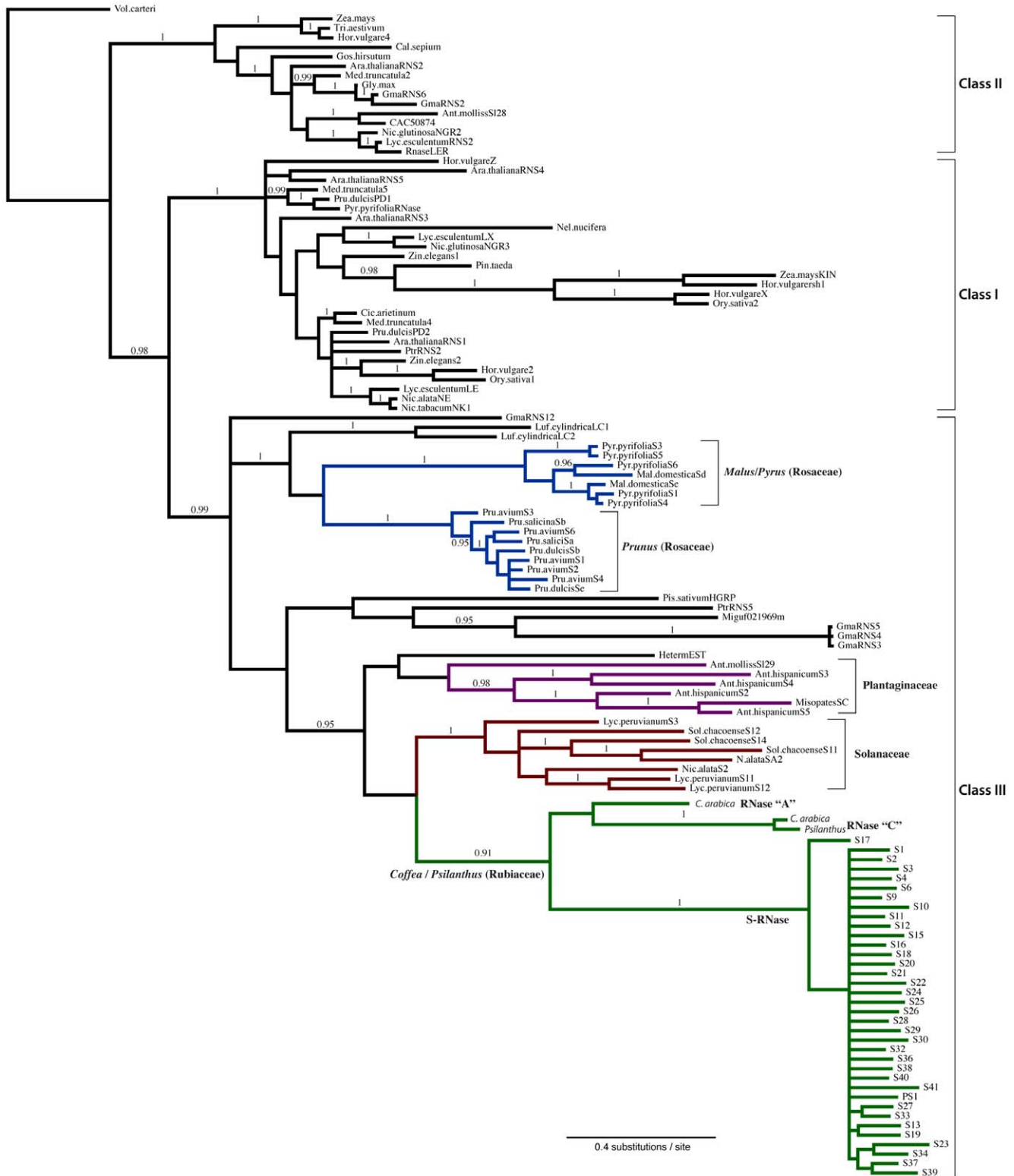
Figure 3. Phylogenetic relationships of plant RNase T2 genes. The maximum *a posteriori* tree generated from a Bayesian phylogenetic analysis of RNase T2 amino acid sequences. Branches receiving a posterior probability of 0.95 or greater have been labeled. The Class III RNase T2 genes form a well-supported clade sister to an unresolved clade of Class I and II RNase T2 genes. S-RNase alleles from other GSI lineages are labeled and color-coded (Solanaceae in red; Plantaginaceae in purple; Rosaceae in blue), and all Class III S-Like RNases as well as all Class I and II RNases are shown in black. RNase T2 genes isolated from *Coffea* and *Psilanthus* species are shown in green. The tree shows that the *Coffea* genome contains three distinct RNase T2 genes labeled "A", "C" and S-RNase. Note the significant sequence divergence (i.e. represented by long terminal branch lengths) between putative alleles of the S-RNase gene.
doi:10.1371/journal.pone.0021019.g003

between putative S-RNase allelic lineages. Using the maximum likelihood method employed in the paml software package [69], pairwise dN/dS ratios between putative *Coffea* S-RNase allelic lineages are estimated to be between 0.15 and 0.87 (mean dN/dS = 0.36; see Table S4, Supporting Information Files). We constructed a gene tree of putative *Coffea* S-RNase alleles isolated from each species sampled. Visual comparison of this gene tree with the most likely species tree of the *Coffea* species in our study and *P. ebracteolatus* (Figure 4) provides a test for trans-specific polymorphism of the putative *Coffea* sequences. The S-RNase gene tree shown in Figure 5 deviates significantly from the branching pattern represented in the species tree, showing that many S-RNase allelic lineages exist in distantly related clades of *Coffea* (i.e. African and Madagascan species), which is a pattern indicative of trans-specific polymorphism. This pattern is perhaps most striking when considering S-RNase allelic lineages shared between species of African and Madagascan *Coffea*, which have been shown to represent deeply divergent clades in the genus [30,31].

Expression patterns exhibited by RNase T2 genes in *Coffea*

While the placement of the *Coffea* "A", "C", and S-RNase genes within the class III RNase T2 subfamily is consistent with the S-RNase GSI system, we would expect that an S-RNase would also show gene expression isolated to floral tissues, which is a characteristic of all known S-RNase genes. In order to determine which, if any, of the three class III RNase T2 genes expressed in *Coffea* pistils is most likely the S-RNase gene implicated in the GSI mechanism, we examined expression patterns of each gene in various tissues from a single *C. arabica* tree using standard RT-PCR procedures. Figure 6 shows that the *Coffea* "A" and "C" class III RNase T2 genes are constitutively expressed in relatively equal proportion in all tissues examined (i.e. pistils, leaves, roots, and fruits) regardless of developmental stage. In striking contrast to the constitutive expression patterns of the RNase "A" and "C" genes, Figure 6 shows clearly that expression of the putative *Coffea* S-RNase gene is limited to pistil tissues.
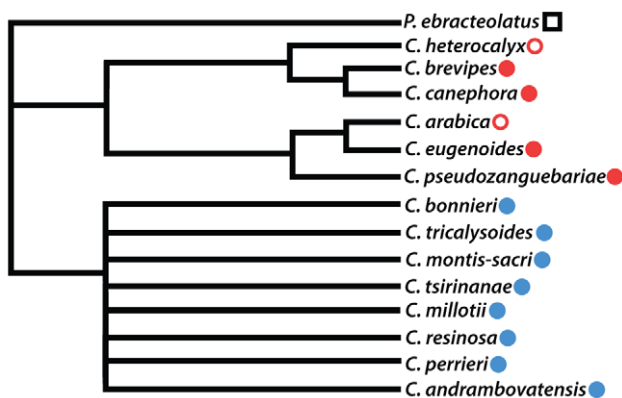
Figure 4. *Coffea* species tree. The tree is shown rooted with the outgroup representative *Psilanthus ebracteolatus*. The geographic origin of species is shown in the color-coded symbols (African species in red, Madagascan species in blue, *Psilanthus* in black). Species with filled symbols are self-incompatible (SI; e.g. *C. canephora*), and species with open symbols are self-compatible (SC; e.g. *C. arabica*). The phylogenetic relationships presented here are summarized from the results of previously published phylogenetic studies in the genus based on chloroplast and nuclear sequence data [30,31].
doi:10.1371/journal.pone.0021019.g004

Discussion

S-RNase genes that have been implicated in the S-RNase-based GSI systems of Solanaceae, Plantaginaceae, and Rosaceae are members of the class III RNase T2 gene subfamily. The results of our phylogenetic analysis of RNase T2 genes indicate the existence of at least three distinct class III RNase T2 genes (i.e. RNase "A", "C", and S-RNase) in the *Coffea* genome. We have also shown that the genome of the SC species *P. ebracteolatus* lacks a homolog to the *Coffea* RNase "A" gene. The phylogenetic relationships of the three class III RNase T2 genes supports a more recent ancestry of *Coffea* RNase "A" and "C" genes. Taken together, these data suggest that at some point after the divergence of *Coffea* from the ancestral lineage of *Coffea* and *Psilanthus* the RNase "C" gene was duplicated (i.e. producing the *Coffea* RNase "A" gene) and in the time since duplication, the two gene copies experienced considerable sequence divergence. An alternative scenario for the absence of a *Psilanthus* RNase "A" homolog is the subsequent loss of this gene following divergence from the common ancestor with *Coffea*. In the absence of more extensive sampling of *Psilanthus* and other Rubiaceae genera, especially those belonging to Coffeeae and other tribes of subfamily Ixoroideae [70,71], it is difficult to determine the directionality of gene loss or gain in this case.

Our phylogenetic analysis of RNase T2 genes included a sequence from the species *Hedyotis terminalis* (Rubiaceae) that originated from a publicly available floral EST library [72]. The genus *Hedyotis* is known to contain both SC [73] and GSI species [74], but the biochemical basis of the GSI mechanism is unknown. The placement of this *H. terminalis* RNase in our phylogenetic analysis of RNase T2 genes (see Figure 3) supports the classification by Vieira et al. [17] as a member of the class III subfamily, but amino acid motifs in this sequence are suggestive of an S-like RNase, rather than a functional S-RNase (see below). Supporting this hypothesis, our analysis places the *H. terminalis* RNase in an unresolved position within the class III RNase T2 subfamily with several other S-like RNase genes. While this placement supports the conclusions of Vieira et al. [17], this support is necessarily tentative due to the short length of this sequence in the *H. terminalis* floral EST library.

Vieira et al. [17] surveyed amino acid sequences from a large sample of known S-RNase alleles and S-like genes in order to identify motifs that might be useful in the characterization of RNase genes being mined from rapidly growing whole genome data sets. They found that S-like RNases (i.e. non-S-RNase members of the class III RNase T2 subfamily) generally exhibit an "HEW" amino acid motif directly upstream of the S-RNase conserved region 3 (c3; see Figure 1 and Figure S1, Supporting Information Files). While the *C. arabica* RNase "A" and "C" genes exhibit the HEW motif at this site, this motif also occurs in 7 of the 36 putative *Coffea/Psilanthus* S-RNase alleles (i.e. alleles S4, S16, S17, S24, S28, S32, and S37; see Figure 5). Assuming our identification of the *Coffea* S-RNase is accurate, our results suggest that the presence of the HEW amino acid motif may not provide an accurate delimitation of S-like RNases from functional S-RNases, particularly in the absence of gene expression and polymorphism data [17]. Yet, given that the two *Coffea* S-like RNase genes identified in the current study (i.e. RNase "A" and "C") were found to carry the HEW motif, the possibility of a functional role for this motif in the S-RNase GSI mechanism remains [41,67].

We have shown that *Coffea* species express three distinct class III RNase T2 genes in pistil tissues and that one of these genes (i.e. the S-RNase) is both polymorphic and heterozygous in SI species of *Coffea* (see Table 1). From the 14 *Coffea* species sampled we have identified 35 putative S-RNase allelic lineages exhibiting
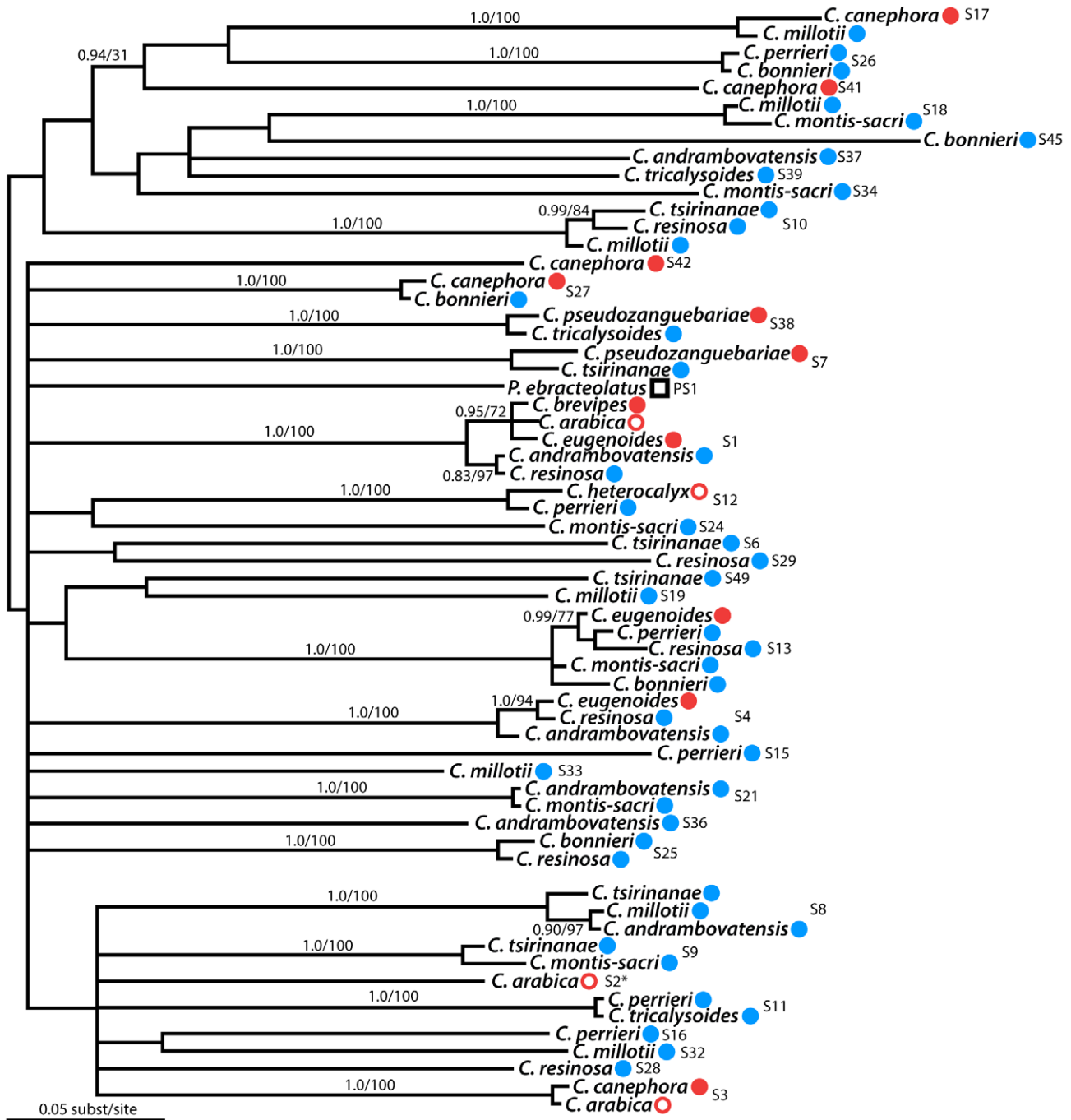
Figure 5. *Coffea* S-RNase trans-specific polymorphism. A. The maximum a posteriori phylogram of putative *Coffea* S-RNase alleles is shown with midpoint rooting and branches receiving strong support are labeled (i.e.>95% posterior probability). The *Psilanthus* S-RNase homolog is indicated by a black box. The geographic origin of putative S-RNase alleles is indicated by the colored symbols following the labeling scheme in Figure 4 (African = red; Madagascan = blue). All putative S-RNase allelic lineages (i.e. sequences that consistently group together and exhibit minimal divergence) have been labeled on the tree (e.g. S1, S4, etc.). All individuals from the self-compatible species *C. heterocalyx* carry sequences from just one putative S-RNase allelic lineage (S12). The gene tree shows a clear signature of trans-specific polymorphism in putative S-RNase alleles sequenced from a phylogenetically and geographically diverse sample of *Coffea* species.
doi:10.1371/journal.pone.0021019.g005

considerable sequence divergence (mean pairwise distance = 0.248). Furthermore, a gene tree of putative S-RNase alleles exhibits a clear pattern of trans-specific polymorphism (Figure 5) consistent with the canonical Eudicot S-RNase system. We have shown that the putative *Coffea* S-RNase exhibits expression patterns isolated to pistil tissues (Figure 6), consistent with our observations of the location of

the *Coffea* GSI response (Figure 2). Together, these lines of evidence lead to our conclusion that *Coffea* utilizes an S-RNase GSI mechanism that is homologous to the S-RNase based system of GSI that is known to be functional in the three Eudicot families Solanaceae, Plantaginaceae, and Rosaceae. This conclusion is logical given the biochemical complexity of SI mechanisms and the
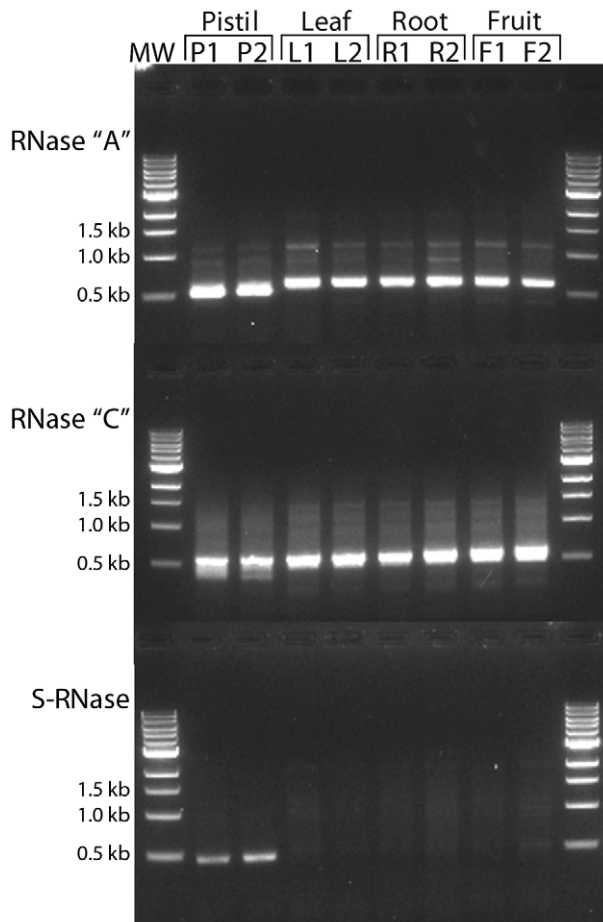
Figure 6. Expression of RNase T2 genes in *C. arabica* tissues. Agarose gel image showing RT-PCR results using gene-specific primers for *Coffea* RNase "A", "C", and S-RNase genes. cDNAs were generated from RNA extracts of various *C. arabica* tissues labeled as follows: P1 = pre-anthesis pistils dissected from flower buds; P2 = post-anthesis pistils; L1 = leaf at 1 week; L2 = leaf at >2 weeks; R1 = fine root; R2 = primary root; F1 = fruit at approximately 1 week; F2 = fruit at >3 weeks. RNase "A" and "C" genes are constitutively expressed in *C. arabica* tissues, but the S-RNase gene shows pistil-specific expression characteristic of genes involved in the S-RNase GSI system.
doi:10.1371/journal.pone.0021019.g006

relatively recent common ancestry of Rubiaceae, Solanaceae, and Plantaginaceae within the asterid clade of Eudicots (approximately 90 Ma; [75,76]). Rubiaceae is thus the fourth Eudicot family to be identified as utilizing the S-RNase GSI system, lending further support for the hypothesis of a single evolutionary origin of this SI mechanism in the ancient ancestral lineage of the Eudicots.

Several questions remain regarding *Coffea* S-RNase GSI. For example, the male determinant S-locus gene (or genes) expressed in pollen tubes has yet to be identified. Furthermore, the mechanism of self-fertility in the diploid species *C. anthonyi* and *C. heterocalyx* remains unclear. While we found no evidence for loss of function mutations in the *C. heterocalyx* S-RNase (S12), previous phylogenetic analyses utilizing samples of this species by Maurin et al. [30] and experiments performed by Coulibaly et al. [38] are a recent hybrid of *C. eugenioides* S.Moore and *C. liberica* Hiern. var. *dewevrei* (De Wild. & T.Durand) Lebrun. Hybridization has been shown to breakdown both the Brassicaceae sporophytic SI mechanism [77] and the Poaceae GSI mechanism [78], but there are also examples of recent hybrid species with fully functioning SI mechanisms [79,80].

Future work to validate our conclusions might include examining the segregation of S-RNase alleles in mapping populations of *Coffea* trees (see for example [38,81]). One would expect obligate heterozygosity of S-RNase genotypes in offspring resulting from crosses of known maternal and paternal diploid genotypes. The Rubiaceae contains many genera with heterostylous flowers, and while many of these species exhibit a heteromorphic gametophytic self-incompatibility response [82], the genetic mechanism of this GSI is unknown [83]. By identifying the *Coffea* S-RNase gene, the results presented here are an important foundation for future examinations of Rubiaceae mating systems. For example, a phylogenetic investigation of the distribution of homomorphic and heteromorphic incompatibility systems in the Rubiaceae could be fruitful in identifying the ancestral mating system condition of the family.

Materials and Methods

Plant materials and GSI mating system tests

Tissues used in this study were collected from individual plants and small populations of African (6 species sampled) and Madagascan (7 species sampled) *Coffea* maintained in cultivated collections or wild populations in Madagascar, France, and the United States. While all Madagascan species and the majority of African species sampled are thought to be SI [61], three African species (i.e. *C. arabica*, *C. anthonyi*, and *C. heterocalyx*) are known to be SC [29,38]. Additionally, a single accession of *P. ebracteolatus* was sampled, a self-compatible African species that is closely related to, or perhaps nested within *Coffea* (Davis, unpublished data). Complete information regarding the provenance of each sample used in this study is provided in Table S3 (see Supporting Information Files). Pistil tissues were sampled from each plant in this study and preserved in RNALater (Ambion) for RNA extraction, and formalin–acetic acid–alcohol (FAA) fixative for microscopic analysis.

Each plant from which tissues were collected was also tested for self-incompatibility using a standard aniline blue pollen tube staining procedure adapted from Alexander [66]. This was achieved by covering flowering branches (at least two branches per plant) with a fine nylon mesh 1–2 days before anthesis, returning to the plant when the buds have opened and vigorously shaking the bagged branch to insure self-pollination. After 12–24 hours, the bags were removed and five to ten mature styles were collected for FAA preservation. After 24 hours, FAA preserved styles were transferred to 70% ethanol for storage. The styles were subsequently softened with 1M NaOH, washed repeatedly in water and stained with 10% aniline blue in 0.5M H3PO4 for 24 hours at 4°C. The styles were transferred to a 1:1 mixture of aniline blue stain and glycerol for 4–24 hours at 4°C, and mounted onto glass slides. Gentle pressure applied to the glass cover slip flattened the stained styles tissues, and the slides were viewed through fluorescence microscopy under ultraviolet light with a blue filter.

Amplification and cloning of *Coffea* RNase T2 genes

Total RNA was extracted from RNALater-preserved styles with the Ambion RNaqueous kit, and RNA concentration was quantified with spectrophotometry (Nanodrop). Due to the level of sequence divergence presumed to exist between any putative *Coffea* RNase T2 genes and those that have been sequenced to date from eudicot species known to utilize the S-RNase system, we employed a standard 3′ RACE (rapid amplification of cDNA ends; [84]) procedure that allowed the use of just one gene-specific degenerate primer to amplify RNase T2 genes from *Coffea* pistil RNA pools.

Briefly, first strand cDNA was synthesized from heat-denatured (80°C for 5 mins) total RNA (150 ng – 1 μg) with MMLV reverse transcriptase (Promega), RNase inhibitor (RNaseOut, Invitrogen), and a 3′ RACE adapter primer at 37°C for one hour. The resulting cDNAs were diluted 1:1 with H2O, and used as template in a two-step nested PCR with reverse primers complementary to the 3′ RACE adapter sequence (3′RACE-Outer.R; 3′RACE-Inner.R), and a degenerate forward primer for conserved region c2, which was designed through an analysis of conserved motifs evident between known S-RNase T2 genes [15,16]. Reactions from the fist step were diluted 1:1 with H2O and employed as template in the second nested PCR. All reactions were performed under standard PCR conditions using Jump Start Taq polymerase (Sigma), and all primer sequences and annealing temperatures are provided in Table S1 (see Supporting Information Files).

Amplicon resulting from the second-step nested PCR was run on a 1% agarose gel stained with 0.25 mg/mL ethidium bromide and all bands larger than approximately 200 bp were extracted from the gel and eluted into 30 uL of sterile water using the UltraClean GelSpin DNA Extraction Kit (Mo-Bio). The resulting gel extracts were cloned using the TOPO TA Kit (Invitrogen) and LB agar plates with 50 μg/mL ampicillin. Approximately 32–48 colonies were screened by PCR (i.e. using the M13 primers provided in the TOPO TA Kit) for insert size and all inserts greater than approximately 200 bp were sequenced with M13 primers at the Duke IGSP DNA Sequencing Facility on an Applied Biosytems 3730 xl capillary sequencer. Preliminary sequencing results allowed the design of less-degenerate forward and reverse primers specific to each of the three *Coffea* RNase T2 gene copies (i.e. "A", "C", and S- RNases; see Table S1, Supporting Information Files). Furthermore, several less-degenerate forward/reverse primers for the C2 region of *Coffea* S- RNase T2 genes were designed due to the significant sequence divergence between putative alleles at this locus (see Table S1, Supporting Information Files). Utilizing these gene/allele-specific primers (i.e. in the second step of the nested PCR) provided considerable improvement in the efficiency of RNase T2 amplification from *Coffea* cDNAs. This improved efficiency allowed the characterization of RNase T2 transcripts by screening and sequencing fewer colonies (i.e. from 16–24) from amplicon pools generated using gene/allele-specific primers in nested 3′RACE PCR.

Analysis of *Coffea* RNase T2 gene expression

We analyzed RNase T2 expression patterns in various tissues of a single *C. arabica* plant growing in the Duke University Greenhouse Teaching Collection (see Table S3, Supporting Information Files). Total RNA was extracted from RNALater-preserved (Ambion) root, fruit, young leaf (i.e.<1 week), mature leaf (i.e.> 1 week), petals, and stamens using the methods described above. Following the preliminary cloning of RNase T2 genes described above, we designed primers specific to RNase "A", "C", and S-RNase genes known to be expressed in the pistil tissues of this plant (see Table S1, Supporting Information Files), and performed reverse transcriptase PCR and 3′ RACE following the procedures described above. The resulting amplicon was run on a 1.5% agarose gel stained with ethidium bromide and bands were visualized with ultraviolet light.

Phylogenetic analyses

To test membership in the RNase T2 gene family, we performed a phylogenetic analysis on an amino acid sequence alignment of RNase T2 sequences from Igic and Kohn [15] and several RNase sequences that grouped with Class II and Class III RNase T2 genes in the gene trees published by MacIntosh et al. [41](see Table S2, Supporting Information Files). These RNase T2 genes and *Coffea* RNase T2 genes (non-redundant with respect to putative alleles) were manually aligned using Igic and Kohn's [15] amino acid alignment as a guide. The WAG+I+G model of nucleotide sequence evolution was found to be the best fit to the data using the Akaike Information Criterion in the program ProtTest v.2.4 [85]. Bayesian phylogenetic analyses were performed in triplicate under the best-fit model in the program in MrBayes v3.1.2 [86]. Each run was composed of 4 chains that were sampled every 1000 generations for 10 million generations. Stationarity of each run and convergence of the three chains were verified by examining the lnL scores of all parameters in the program Tracer v1.5 [87]. After removing the first 1 million generations as burn-in, the 50% majority-rule consensus topology and branch lengths were estimated from the 9,001 trees sampled from the posterior distribution. Isoelectric point (pI) estimates were calculated on the ExPASy Compute pI/Mw web server [88]. Pairwise distances were estimated between putative S-RNase allelic lineages using the uncorrected p distance with the software PAUP* v.4.10b [89]. Pairwise dN/dS ratio estimates were generated using the maximum likelihood method implemented in the codeml program (i.e. by executing the command "runmode = −2") within the paml v.4.4 software package [69].

In order to test for trans-specific polymorphism in putative *Coffea* RNase T2 alleles, the gene tree of putative S-RNase alleles was estimated in MrBayes v3.1.2 [86] with model parameters assigned according to the best-fit model (GTR+G+I) estimated in the program MrModeltest v2.0 [90]. The analysis consisted of 3 independent runs, each with 4 chains that were sampled every 1000 generations for 10 million generations. Stationarity of each run and convergence of the three chains were verified by examining the lnL scores of all parameters in the program Tracer v1.5 [87]. After removing the first 1 million generations as burn-in, the 50% majority-rule consensus topology and branch lengths were estimated from the 9,001 trees sampled from the posterior distribution.

Supporting Information

Figure S1   Sequence logo for putative *Coffea* S-RNase alleles. Sequence logos for each of the *Coffea* RNase T2 genes identified in this study were generated with the WebLogo software [92]. Conserved (i.e. c3–c5) and hyper-variable (i.e. HVa and HVb) regions (see Figure 1) are labeled. The "HEW" motif discussed in the text is labeled in red.
(TIF)

Table S1   Primer sequences for amplification of Coffea RNase T2 genes. The primers listed were employed in a standard 3′ RACE procedure and all amplifications used an annealing temperature of 54°C.
(PDF)

Table S2   Plant RNase T2 sequences employed in phylogenetic analysis shown in Figure 3. Genbank accession numbers are provided for all sequences except those from *Glycine max*, but see MacIntosh et al. [41].
(PDF)

Table S3   Plant materials sampled and sequence accession information. Sample number refers to the source accession number of individual plants. Genbank sequence accession numbers are provided. Individuals carrying S-RNase alleles identical to other individuals of the same species are indicated by "-" in the sequence accession column.
(PDF)

Table S4   Pairwise dN/dS Estimates. Parameter estimates were generated with the maximum likelihood method in the paml software package [69]. A key for pairwise sequence identification is provided at the end of the table.
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: MN AY. Performed the experiments: MN. Analyzed the data: MN. Contributed reagents/materials/analysis tools: AD FA. Wrote the paper: MN.

References

1. Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. Annual Review of Ecology and Systematics 18: 237–268.
2. Stebbins GL (1970) Adaptive radiation of reproductive characteristics in angiosperms, I: pollination mechanisms. Annual Review of Ecology and Systematics 1: 307–326.
3. Igic B, Kohn JR, Lande R (2008) Loss of self-incompatibility and its evolutionary consequences. International Journal of Plant Sciences 169: 93–104.
4. Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, et al. (2010) Species Selection Maintains Self-Incompatibility. Science 330: 493–495.
5. Whitehouse HLK (1950) Multiple-allelomorph Incompatibility of Pollen and Style in the Evolution of the Angiosperms. Ann Bot 14: 199–216.
6. Allen AM, Hiscock SJ (2008) Evolution and Phylogeny of Self-Incompatibility Systems in Angiosperms. In: Franklin-Tong V, ed. Self-Incompatibility in Flowering Plants. Berlin: Springer. pp 73–101.
7. Charlesworth D, Vekemans X, Castric V, Glemin S (2005) Plant self-incompatibility systems: a molecular evolutionary perspective. New Phytologist 168: 61–69.
8. Kaothien-Nakayama P, Isogai A, Takayama S (2010) Self-Incompatibility Systems in Flowering Plants. In: Pua EC, Davey MR, eds. Plant Developmental Biology – Biotechnological Perspectives: Volume 1. Berlin Heidelberg: Springer-Verlag. pp 459–485.
9. Franklin-Tong V, Franklin FCH (2003) Gametophytic self-incompatibility inhibits pollen tube growth using different mechanisms. Trends in Plant Science 8: 598–605.
10. Bosch M, Franklin-tong VE (2008) Self-incompatibility in Papaver: signaling to trigger PCD in incompatible pollen. J Exp Bot 59: 481–490.
11. Franklin-Tong V (2007) Inhibiting Self-Pollen: Self-Incompatibility in Papaver Involves Integration of Several Signaling Events. Journal of Integrative Plant Biology 49: 1219–1226.
12. Sherman-Broyles S, Nasrallah J (2008) Self-Incompatibility and Evolution of Mating Systems in the Brassicaceae. In: Franklin-Tong V, ed. Self-Incompatibility in Flowering Plants Berlin. pp Springer.123–150.
13. Watanabe M, Suzuki G, Takayam S (2008) Milestones Identifying Self-Incompatibility Genes in Brassica Species: From Old Stories to New Findings. In: Franklin-Tong V, ed. Self-Incompatibility in Flowering Plants Berlin. pp Springer.151–172.
14. Zhang Y, Xue Y (2008) Molecular Biology Of S-Rnase-Based Self-Incompatibility. In: Franklin-Tong V, ed. Self-Incompatibility in Flowering Plants Berlin. pp Springer.193–215.
15. Igic B, Kohn JR (2001) Evolutionary relationships among self-incompatibility RNases. Proc Natl Acad Sci U S A 98: 13167–13171.
16. Steinbachs JE, Holsinger KE (2002) S-RNase-mediated Gametophytic Self-Incompatibility is Ancestral in Eudicots. Mol Biol Evol 19: 825–829.
17. Vieira J, Fonseca NA, Vieira CP (2008) An S-RNase-based gametophytic self-incompatibility system evolved only once in Eudicots. Journal of Molecular Evolution 67: 179–190.
18. McClure B, Franklin-Tong V (2006) Gametophytic self-incompatibility: understanding the cellular mechanisms involved in "self" pollen tube inhibition. Planta 224: 233–245.
19. Wheeler D, Newbigin E (2007) Expression of 10 S-Class SLF-like Genes in Nicotiana alata Pollen and Its Implications for Understanding the Pollen Factor of the S-Locus. Genetics 177: 2171–2180.
20. Lai Z, et al. (2002) An F-box gene linked to the self-incompatibility (S) locus of Antirrhinum is expressed specifically in pollen and tapetum. Plant Molecular Biology 50: 29–41.
21. Ushijima K, et al. (2003) Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. The Plant Cell Online 15: 771–782.
22. Kao T, Tsukamoto T (2004) The molecular and genetic bases of S-RNase-based self-incompatibility. The Plant Cell 16: S72–S83.
23. Lee HS, Huang S, Kao T (1994) S proteins control rejection of incompatible pollen in Petunia inflata. Nature 367: 560–563.
24. Sassa H, Nishio T, Kowyama Y, Hirano H, Koba, et al. (1996) Self-incompatibility (S) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease superfamily. Molecular and General Genetics 250: 547–557.
25. Vega FE, Rosenquist E, Collins W (2003) Global project needed to tackle coffee crisis. Nature 425: 343.
26. Davis, AP, Govaerts R, Bridson DM, Stoffelen P (2006) An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). Bot Journal Linn Soc 152: 465–512.
27. Govaerts R, Ruhsam M, Andersson L, Robbrecht E, Bridson DM, et al. (2010) World Checklist of Rubiaceae. The Board of Trustees of the Royal Botanic Gardens, Kew, <http://www.kew.org/wcsp/rubiaceae>, accessed 21 October 2010.
28. Carvalho A, Ferwerda FP, Frahm-Leliveld JA, Mdina DM, Mendes AJT, et al. (1969) Coffee. In: Ferwerda FP, Wit F, eds. Outlines of perennial crop breeding in the Tropics. Wageningen: Veenman & Zonen NV. pp 189–241.
29. Stoffelen P, Noirot M, Couturon E, Bontems S, de Block P, et al. (2009) Coffea anthonyi, a new self-compatible Central African coffee species, closely related to an ancestor of Coffea arabica. Taxon 58: 133–140.
30. Maurin O, Davis AP, Chester M, Mvungi EF, Jauferally-Fakim Y, et al. (2007) Towards a phylogeny of Coffea (Rubiaceae): Identifying well-supported lineages based on nuclear and plastid DNA sequences. Ann Bot 100: 1565–1583.
31. Anthony F, Diniz LEC, Combes M, Lashermes P (2010) Adaptive radiation in Coffea subgenus Coffea L. (Rubiaceae) in Africa and Madagascar. Plant Syst Evol 285: 51–64.
32. Berthou F, Mathieu C, Vedel F (1983) Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus Coffea. Theor Appl Genet. 65: 77–84.
33. Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, et al. (1999) Molecular characterization and origin of the Coffea arabica L. genome. Mol Genet Genomics 261: 259–266.
34. Chawla B, Bernatzky R, Liang W, Marcotrigiano M (1997) Breakdown of self-incompatibility in tetraploid Lycopersicon peruvianum: inheritance and expression of S-related proteins. Theoretical and Applied Genetics 95: 992–996.
35. Miller JS, Venable DL (2000) Polyploidy and the evolution of gender dimorphism in plants. Science 289: 2335–2338.
36. Yeung K, Miller JS, Savage AE, Husband BC, Igic B, et al. (2005) Association of ploidy and sexual system in Lycium californicum (Solanaceae). Evolution 59: 2048–2055.
37. Robertson K, Goldberg EE, Igic B (2010) Comparative evidence for the correlated evolution of polyploidy and self-compatibility in Solanaceae. Evolution.
38. Coulibaly I, Noirot M, Lorieux M, Charrier A, Hamon, S, et al. (2002) Introgression of self-compatibility from Coffea heterocalyx to the cultivated species C. canephora. Theoretical and Applied Genetics 105: 994–999.
39. Clark AG (1992) Evolutionary inferences from molecular characterization of self-incompatibility alleles. In: Takahata N, Clark AG, eds. Mechanisms of Molecular Evolution. Sunderland, MA: Sinauer. pp 79–108.
40. Luhtala N, Parker R (2010) T2 Family ribonucleases: ancient enzymes with diverse roles. Trends in Biochemical Sciences 35: 253–259.
41. MacIntosh GC, Hillwig MS, Meyer A, Flagel L (2010) RNase T2 genes from rice and the evolution of secretory ribonucleases in plants. Mol Genet Genomics 283: 381–396.
42. Green PJ (1994) The ribonucleases of higher plants. Annual Review of Plant Biology 45: 421–445.
43. Itagaki T, Koyama H, Daigo S, Kobayashi H, Koyama T, et al. (2006) Primary structure and properties of ribonuclease Bm2 (RNase Bm2) from Bryopsis maxima. Biological & Pharmaceutical Bulletin 29: 875–883.
44. Anderson MA, Cornish EC, Mau SL, Williams EG, Hoggart R, et al. (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in Nicotiana alata. Nature vol.321: 38–44.
45. Anderson MA, McFadden GI, Bernatzky R, Atkinson A, Orpin T, Dedman H, et al. (1989) Sequence variability of three alleles of the self-incompatibility gene of Nicotiana alata. The Plant Cell Online 1: 483–491.

46. Matton DP, Maes O, Laublin G, Xike Q, Bertrand C, et al. (1997) Hypervariable Domains of Self-Incompatibility RNases Mediate Allele-Specific Pollen Recognition. Plant Cell 9: 1757–1766.

47. Sassa H, Nishio T, Kowyama Y, Hirano H, Koba, et al. (1996) Self-incompatibility (S) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease superfamily. Molecular and General Genetics 250: 547–557.

48. Vieira J, Morales-Hojas R, Santos RA, Vieira CP (2007) Different positively selected sites at the gametophytic self-incompatibility pistil S-RNase gene in the Solanaceae and Rosaceae (Prunus, Pyrus, and Malus). Journal of Molecular Evolution 65: 175–185.

49. Miller JS, Levin RA, Feliciano NM (2008) A tale of two continents: Baker's Rule and the maintenance of self-incompatibility in Lycium (Solanaceae). Evolution 62: 1052–1065.

50. Vekemans X, Slatkin M (1994) Gene and Allelic Genealogies at a Gametophytic Self-Incompatibility Locus. Genetics 137: 1157–1165.

51. Foxe JP, Slotte T, Stahl EA, Neuffer B, Hurka H, et al. (2009) Recent speciation associated with the evolution of selfing in Capsella. P Natl Acad Sci USA 106: 5241–5245.

52. Guo YL, Bechsgaard JS, Slotte T, Neuffer B, Lascoux M, et al. (2009) Recent speciation of Capsella rubella from Capsella grandiflora, associated with loss of self-incompatibility and an extreme bottleneck. P Natl Acad Sci USA 106: 5246–5251.

53. Bateman AJ (1952) Self-incompatibility systems in the angiosperms. I. Theory. Heredity 6: 285–310.

54. Charlesworth D (1995) Multi-allelic self-incompatibility polymorphisms in plants. BioEssays 17: 31–38.

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

56. Uyenoyama MK (2003) Genealogy-dependent variation in viability among self-incompatibility genotypes. Theoretical Population Biology 63: 281–293.

57. Richman AD, Kohn JR (1999) Self-incompatibility alleles from Physalis: Implications for historical inference from balanced genetic polymorphisms. Proc Natl Acad Sci U S A 96: 168–172.

58. Igic B, Bohs L, Kohn JR (2003) Historical inferences from the self-incompatibility locus. New Phytologist 161: 97–105.

59. Figueroa F, Gunther E, Klein J (1988) MHC polymorphism pre-dating speciation. Nature 335: 265–267.

60. Devier B, Aguileta G, Hood ME, Giraud T (2008) Ancient Trans-specific Polymorphism at Pheromone Receptor Genes in Basidiomycetes. Genetics 181: 209–223.

61. Berthaud J (1980) L'incompatibilité chez Coffea canephora: méthode de test et déterminisme génétique. Café Cacao Thé 24: 267–274.

62. Charrier A (1979) La structure genetique des cafeiers spontanes de la region malgache (Mascarocoffea). Memoires ORSTOM (87). Paris: ORSTOM.

63. Charrier A, Berthaud J (1985) Botanical classification of coffee. In: Clifford MN, Willson KC, eds. Coffee botany, biochemistry and production of beans and beverage. London: Croom Helm. pp 13–47.

64. Schemske, DW, Lande R (1985) The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution 39: 41–52.

65. Husband BC, Schemske DW (1996) Evolution of the magnitude and timing of inbreeding depression in plants. Evolution 50: 54–70.

66. Alexander MP (1987) A method for staining pollen tubes in pistil. Stain Technology 62: 107–112.

67. de Nettancourt D (1997) Incompatibility in angiosperms. Sex Plant Reprod 10: 185–199.

68. Irie M (1999) Structure-Function Relationships of Acid Ribonucleases: Lysosomal, Vacuolar, and Periplasmic Enzymes. Pharmacology & Therapeutics 81: 77–89.

69. Yang Z (2007) PAML 4: a program package for phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24: 1586–1591.

70. Davis, AP, Chester M, Maurin O, Fay MF (2007) Searching for the relatives of Coffea (Rubiaceae, Ixoroideae): The circumscription and phylogeny of Coffeeae based on plastid sequence data and morphology. Amer J Bot 94: 313–329.

71. Tosh J, Davis AP, Dessein S, de Block P, Huysmans S, et al. (2009) Phylogeny of Tricalysia (Rubiaceae) and its relationships with allied genera based on plastid DNA data: resurrection of the genus Empogona. Ann Missouri Bot Gard 96: 194–213.

72. Plechakova O, Tranchant-Dubreuil C, Benedet F, Couderc M, Tinaut A, et al. (2009) MoccaDB - an integrative database for functional, comparative and diversity studies in the Rubiaceae family. BMC Plant Biology 9: 123–134.

73. Riveros GM, Riveros OR, Barria, Humana PAM (1995) Self-compatibility in distylous Hedyotis salzmannii (Rubiaceae). Plant Systematics and Evolution 194: 1–8.

74. Wu X, Li A, Zhang D (2010) Cryptic self-incompatibility and distyly in Hedyotis acutangula Champ. (Rubiaceae). Plant Biology 12: 484–494.

75. Anderson, CL, Bremer K, Friis EM (2005) Dating phylogenetically basal eudicots using rbcL sequences and multiple fossil reference points. American Journal of Botany 92: 1737–1748.

76. Magallon S, Castillo A (2009) Angiosperm diversification through time. American Journal of Botany 96: 349–365.

77. Nasrallah JB, Liu P, Sherman-Broyles S, Schmidt R, Nasrallah ME (2007) Epigenetic mechanisms for breakdown of self-incompatibility in interspecific hybrids. Genetics 175: 1965–1973.

78. Sloop CM, Ayres DR, Strong DR (2009) The rapid evolution of self-fertility in Spartina hybrids (Spartina alterniflora × foliosa) invading San Francisco Bay, CA. Biological Invasions 11: 1131–1144.

79. Brandvain Y, Haig D (2005) Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. American Naturalist 166: 330–338.

80. Paetsch M, Mayland-Quellhorst S, Hurka H, Neuffer B (2011) Evolution of the Mating System in the Genus Capsella (Brassicaceae). In: Glaubrecht M, ed. Evolution in Action. Berlin Heidelberg: Springer-Verlag. pp 77–100.

81. Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M et al (2001) Genetic linkage map of Coffea canephora: effect of segregation distortion and analysis of recombination rate in male and female meioses. Genome 44: 589–595.

82. Anderson WR (1973) A morphological hypothesis for the origin of heterostyly in the Rubiaceae. Taxon 22: 537–542.

83. Klein DE, Freitas EL, Da Cunha EM (2009) Self-incompatibility in a distylous species of Rubiaceae: is there a single incompatibility response of the morphs? Sexual Plant Reproduction 22: 121–131.

84. Frohman MA, Dush MK, Martin GR (1988) Rapid Production of Full-Length cDNAs from Rare Transcripts: Amplification Using a Single Gene-Specific Oligonucleotide Primer. Proc Natl Acad Sci U S A 85: 8998–9002.

85. Abascal F, Zardoya R, Posada, D (2005) ProtTest: Selection of best-fit models of protein evolution. Bioinformatics: 21(9): 2104–2105.

86. Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

87. Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

88. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM, ed. The proteomics protocols handbook. Totowa, NJ: Humana Press. pp 571–607.

89. Swofford, DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

90. Nylander JAA (2004) MrModeltest v2. Program distributed by the author Evolutionary Biology Centre, Uppsala University.

91. Vieira J, Charlesworth D (2002) Molecular variation at the self-incompatibility locus in natural populations of the genera Antirrhinum and Misopates. Heredity 88: 172–181.

92. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Research 14: 1188–1190.