JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12396

The founding of Mauritian endemic coffee trees by a synchronous long-distance dispersal event

M. D. NOWAK*, B. C. HALLER†[1] & A. D. YODER‡

*National Centre for Biosystematics, Natural History Museum, University of Oslo, Oslo, Norway
†Department of Biology and Redpath Museum, McGill University, Montreal, QC, Canada
‡Department of Biology, Duke University, Durham, NC, USA

Keywords:

gametophytic self-incompatibility;
genetic bottleneck;
long-distance dispersal;
negative frequency-dependent selection;
trans-specific polymorphism.

Abstract

The stochastic process of long-distance dispersal is the exclusive means by which plants colonize oceanic islands. Baker's rule posits that self-incompatible plant lineages are unlikely to successfully colonize oceanic islands because they must achieve a coordinated long-distance dispersal of sufficiently numerous individuals to establish an outcrossing founder population. Here, we show for the first time that Mauritian *Coffea* species are self-incompatible and thus represent an exception to Baker's rule. The genus *Coffea* (Rubiaceae) is composed of approximately 124 species with a paleotropical distribution. Phylogenetic evidence strongly supports a single colonization of the oceanic island of Mauritius from either Madagascar or Africa. We employ Bayesian divergence time analyses to show that the colonization of Mauritius was not a recent event. We genotype S-RNase alleles from Mauritian endemic *Coffea*, and using S-allele gene genealogies, we show that the Mauritian allelic diversity is confined to just seven deeply divergent *Coffea* S-RNase allelic lineages. Based on these data, we developed an individual-based model and performed a simulation study to estimate the most likely number of founding individuals involved in the colonization of Mauritius. Our simulations show that to explain the observed S-RNase allelic diversity, the founding population was likely composed of fewer than 31 seeds that were likely synchronously dispersed from an ancestral mainland species.

Introduction

The dispersal of seed or other diaspores is a fundamental process influencing the geographic distribution of plant populations (Cain *et al.*, 2000). While the majority of seed from a given generation will be dispersed in close proximity to maternal plants, there is some small probability that any given seed will experience long-distance dispersal (LDD; Nathan, 2006). Despite the presumed rarity of such events, the impact of LDD on the spatial and genetic structure of plant populations is thought to be profound (Gillespie *et al.*, 2012). The importance of this mechanism is perhaps most evident when considering processes that are entirely dependent on LDD, such as the assembly and evolution of oceanic island biotas (Carlquist, 1967; Cowie & Holland, 2006).

Empirical evidence and theoretic expectations highlight the importance of plant mating systems in the success or failure of island colonization by LDD. Baker (1955) was the first to recognize the pattern of over-representation of self-compatible (SC) species relative to self-incompatible (SI) species in the floras of oceanic islands compared with terrestrial floras. This observation later became formalized as Baker's rule (Baker, 1967) and has since been verified repeatedly with empirical data in a diversity of lineages (Rambuda & Johnson, 2004; Crawford *et al.*, 2008; van Kleunen *et al.*, 2008; Hao *et al.*, 2010). Baker theorized that this pattern results because SI lineages are considerably less likely to colonize islands due to the requirements of their breeding system. Maintenance of the minimal

Correspondence: Michael D. Nowak, National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, 0318 Oslo, Norway. Tel.: +41(0)797500962; fax: +41(0)446348403; e-mail: michaeldnowak@gmail.com
[1]Present address: Centre d'Ecologie Fonctionelle et Evolutive, CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, France

genetic diversity required for the functioning of an SI breeding system in a founding population necessitates the dispersal of many more individuals than would be necessary for colonization by SC plants. As seeming exceptions to Baker's rule, some plant lineages appear to have colonized oceanic islands despite evidence of an SI mating system in the colonizing ancestor. These exceptions include the Hawaiian Silversword alliance (Carr *et al.*, 1986; Baldwin, 2007), North American Yellow Starthistle (*Centaurea solstitialis*; Sun & Ritland, 1998), *Senecio squalidus* in the British Isles (Brennan *et al.*, 2006) and Old World *Lycium* (Miller *et al.*, 2008, 2011).

Exceptions to Baker's rule must logically represent colonization by a founding population that was sufficiently large to meet the minimum requirements of the SI system. A suitable founding population could be assembled via a single dispersal event carrying multiple seeds from a source population, or via multiple temporally distinct dispersal events, particularly in perennial species capable of waiting for sufficient colonists to accumulate a population with sufficient genetic diversity to maintain itself in perpetuity (Pannell & Barrett, 1998). Regardless of the different mechanisms for assembling a founding population (i.e. single vs. multiple LDD), estimating the size of a founding population with empirical data is quite challenging, and direct observations of LDD events are extremely rare (Nathan, 2006). Nonetheless, the clear importance of this process has prompted attempts to use genetic data to estimate the geographic origin and absolute size of founding populations resulting from ancient LDD processes. When a given LDD event occurred relatively recently, it is possible to estimate the size of the founding population by analysing patterns of polymorphism at rapidly evolving and presumably neutral genetic loci (e.g. microsatellites, AFLP; Nathan *et al.*, 2003; Alsos *et al.*, 2007), but this becomes increasingly intractable with increasing time since the LDD event, or with substantial *in situ* diversification of the island lineage. As the founding population is likely to be much smaller than the source population, genetic drift should be particularly strong in fixing or eliminating polymorphisms segregating at the time of island colonization (Mayr, 1942). Consequently, when there is a strong bottleneck following colonization, allelic diversity at neutral loci will be maintained for only a short time, making it increasingly difficult to estimate the size of a founding population with the increasing antiquity of the LDD event. An important exception is encountered, however, when considering loci at which polymorphism is maintained in populations through the action of balancing selection (Takahata & Nei, 1990; Vekemans & Slatkin, 1994).

Classic examples of the effects of balancing selection on observed patterns of polymorphism include the major histocompatibility complex (MHC) locus in vertebrates (Figueroa *et al.*, 1988), the mating-type locus in fungi (Devier *et al.*, 2008) and self-incompatibility loci in plants (Richman *et al.*, 1996; Roux *et al.*, 2013). Negative frequency-dependent selection (NFDS) is a specific form of balancing selection in which the individual fitness effect of an allele is inversely related to the frequency of that allele, resulting in a large fitness advantage to individuals carrying rare alleles (Ayala & Campbell 1974). Theory predicts that a large population at equilibrium should carry a relatively large number of alleles at a locus under NFDS and that allele frequencies will be approximately equal if no other source of selection is operating on the locus (Lawrence, 2000). This is because the probability that an allele will drift to either fixation or elimination is vanishingly small (Clark, 1992). Polymorphisms at loci under NFDS tend to be maintained much longer than polymorphisms at neutral loci – even resulting in shared ancestral polymorphisms in descendent populations or species, a phenomenon known as trans-specific polymorphism (Vekemans & Slatkin, 1994; Richman & Kohn, 2000; Igic *et al.*, 2004; Bos *et al.*, 2008).

Plant self-incompatibility loci (S-loci) are among the classic examples of NFDS. Among the various mechanisms of self-incompatibility that have been described in plants, the most phylogenetically widespread is the S-RNase gametophytic self-incompatibility (GSI) mechanism. This type of SI mechanism has been found in at least four angiosperm families, including the Solanaceae, Rosaceae, Plantaginaceae and Rubiaceae (Igic & Kohn, 2001; Nowak *et al.*, 2011). It is based on the biochemical interaction of self-incompatibility RNases (S-RNases), expressed in the stigma and style, with F-box proteins expressed in the growing pollen tube. This interaction leads to the destruction of self- and incompatible pollen tubes as they grow through the style (Kao & Tsukamoto, 2004). This self-incompatibility mechanism is gametophytic because its function is based on the interaction of the haploid pollen genotype and the diploid genotype of the maternal plant. SI populations are known to carry numerous (i.e. dozens) functional S-RNase alleles, which exhibit extensive trans-specific polymorphism (Ioerger *et al.*, 1990; Clark 1992). Trans-specific polymorphism at loci under NFDS has been employed to detect and measure the impact of ancient demographic events (Prugnolle *et al.*, 2005). For example, polymorphism at the vertebrate MHC locus has been used to estimate the size of the wolf population that gave rise to domestic dogs (Niskanen *et al.*, 2013), the size of the founding population of Darwin's finches (Vincek *et al.*, 1997) and the demographics of human origins (Ayala *et al.*, 1994). Furthermore, S-RNase polymorphism has been used to reconstruct an ancient population bottleneck in the Iochrominae/*Physalis*/*Witheringia* clade (Paape *et al.*, 2008) and the demographic impacts of LDD from North America to Africa and Asia in *Lycium* (Miller *et al.*, 2008, 2011).

The genus *Coffea* (Rubiaceae) is composed of approximately 124 species distributed in Africa, Madagascar, tropical Asia, Australia, Grande Comore Island and the Mascarene Islands of Mauritius and Réunion (Davis *et al.*, 2006, 2011). It was recently shown that GSI in the genus *Coffea* functions through the classic Eudicot S-RNase system (S-GSI; Asquini *et al.*, 2011; Nowak *et al.*, 2011). Accordingly, SI populations of *Coffea* have been shown to harbour high allelic diversity at the S-locus, and gene trees of S-RNase alleles isolated from multiple African and Malagasy species show a clear pattern of trans-specific polymorphism (Nowak *et al.*, 2011) likely resulting from the action of balancing selection (Reboiro-Jato *et al.*, 2012). Results of several molecular phylogenetic studies in *Coffea* strongly support the monophyly of the three Mascarene species, thus indicating a single colonization event of the volcanic Mascarene Islands (Davis *et al.*, 2011; Nowak *et al.*, 2012; Fig. 1). While it is currently unclear whether Africa or Madagascar was the source of colonization, it is very likely that Mauritius (the older of the two islands, approximately 8 Ma; McDougall & Chamalaun, 1969; Moore *et al.*, 2011) was the first to be colonized, with subsequent diversification of the lineage into three species: *C. myrtifolia* and *C. macrocarpa* are endemic to Mauritius, while *C. mauritiana* is also found on Réunion (the younger of the two islands, approximately 1 Ma; Moore *et al.*, 2011). The vast majority of African and Malagasy *Coffea* species exhibit strong gametophytic SI (Coulibaly *et al.*, 2002; Nowak *et al.*, 2011), suggesting that the Mascarene *Coffea* could represent an exception to Baker's rule.

Here, we test the self-incompatibility response in Mauritian *Coffea* to determine whether this clade is an exception to Baker's rule, and attempt to estimate specific parameters of the colonization event that led to the establishment of *Coffea* on the Mascarene Islands. To estimate the age of the colonization of Mauritius, we perform a Bayesian molecular divergence time analysis with plastid DNA sequence data. To determine whether Mauritian *Coffea* represent an exception to Baker's rule, we test the strength of GSI by examining pollen tube growth following controlled self-pollination in the field. S-RNase allelic diversity is compared between Mauritian and African/Malagasy *Coffea* species to evaluate the bottleneck in allelic diversity that likely accompanied island colonization. Using this observed S-RNase allelic diversity, we employ an individual-based model to estimate the size of the ancestral *Coffea* population that colonized the island of Mauritius.

Materials and methods

Divergence time estimation

Bayesian divergence time estimation was performed in the software package BEAST v1.7.5 (Drummond *et al.*, 2012) to estimate the age of the colonization of Mauritius. We supplemented the plastid sequence data set from Nowak *et al.* (2012) composed of six loci (*atpF*, *rpolb1*, *rpolb2*, *trnL-F*, *rpl16* and *psa1*) with sequence data from nine species that were previously placed in the genus *Psilanthus*, which was recently subsumed into *Coffea* by Davis *et al.* (2011). The final data matrix consisted of 69 taxa and a total of 6791 aligned nucleotides that were partitioned by locus, and each locus was assigned a unique substitution model (see Table S1; Dryad accession doi: 10.5061/dryad.6r52k). The BEAST analysis was calibrated temporally using priors based on island age and clock rate. First, the age of the most recent common ancestor (MRCA) of the three Mascarene *Coffea* species (i.e. *C. mauritiana*, *C. macrocarpa* and *C. mytrifolia*) was constrained to be no older than 8 Ma. This constraint is consistent with the age of the island

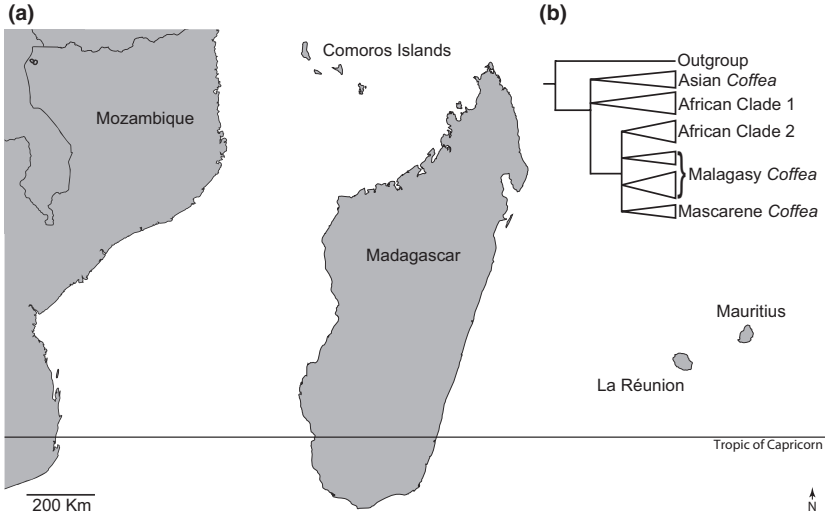

Fig. 1 (a) Islands of the western Indian Ocean. The East African coast (Mozambique), Madagascar, the Comoros Islands and the Mascarene Islands (Mauritius and Réunion) are shown to indicate the distances between landmasses and their relative sizes. (b) Summary phylogenetic tree showing the relationship of several well-supported geographic clades in the genus *Coffea*. These relationships are summarized from the phylogenetic analyses of plastid and nuclear sequence data of Davis *et al.* (2011) and Nowak *et al.* (2012).

of Mauritius based on the oldest volcanic rocks of which the island is composed (McDougall & Chamalaun, 1969; Moore *et al.*, 2011). The molecular clock assumption was relaxed by employing an uncorrelated lognormal distribution on branch rates, and following Paape *et al.* (2008) and Miller *et al.* (2011), we fixed the mean of the lognormal distribution to 0.0007 substitutions/million years (Palmer, 1991; Schnabel & Wendell, 1998). The prior on the standard deviation of the lognormal distribution was left as the default setting (i.e. exponential distribution with mean 0.33). This had the effect of constraining the branch rates while avoiding a strict clock. The monophyly of the Mascarene *Coffea* calibration node was the only topological constraint enforced in the BEAST analysis. The BEAST analysis was provided with a starting tree with branch lengths estimated using the default neighbour-joining algorithm in PAUP* v4.0B10 (Swofford, 2003). The analysis was performed in triplicate for 10 million generations sampling the Markov chain Monte Carlo (MCMC) distribution every 1000 generations. The resulting parameter log files were analysed in Tracer v1.5 (Rambaut & Drummond, 2007) for stationarity, convergence and sufficient effective sample sizes from the posterior distribution of all parameters. The first 1000 trees were removed as burn-in from each posterior sample of trees, and the remaining 9000 trees from each run were combined using the LogCombiner v1.7.5 software (Drummond *et al.*, 2012). This combined sample of 27 000 trees was summarized in a single maximum clade credibility tree annotated with parameter estimates using the TreeAnnotator v1.7.5 software (Drummond *et al.*, 2012). This annotated tree was viewed and manipulated in the tree editing software FigTree v1.4 (Rambaut, 2012). A single run was performed using identical settings, but with an empty data matrix to ensure, there were no spurious effects on the joint prior distribution due to the interaction of different temporal constraints (Heled & Drummond, 2012).

Self-incompatibility tests

Tests of GSI strength were performed through controlled self-pollinations conducted in January 2009 in the Plaine Champagne (PCH) population, which is located within the Black River Gorges National Park of Mauritius. The habitat at Plaine Champagne is native upland dwarf forest, and *C. macrocarpa* and *C. mauritiana* can be found there growing within 100 metres of each other. Four individual plants of each species were selected for breeding system experiments. For each of these plants, all open flowers were removed from one flowering branch and a fine mesh fabric was used to construct a pollinator exclusion barrier around that branch. After 48 h, many flowers had opened, and each branch was vigorously shaken to stimulate self-pollination. After a further 24 h, the fabric was removed and 5–10 styles were removed from each plant and preserved in formalin–acetic acid–alcohol (FAA) fixative for transport. We were unable to perform similar breeding system tests in *C. myrtifolia* because this species is confined to three small and protected populations and the plants were not flowering on the days that we were allowed access to these populations. Microscopic analysis of pollen tube growth was performed following the methods described by Nowak *et al.* (2011).

S-RNase data and gene tree estimation

To ensure maximum allelic diversity from 'mainland' Malagasy and African species, we supplemented the data set of Nowak *et al.* (2011) with S-RNase sequences from small samples (i.e. 1–2 individuals) of several additional species. In total, 21 species of African and Malagasy *Coffea* were sampled in the current study, and S-RNase alleles were genotyped in 84 individual plants (see Table S2). This expanded sampling includes four individuals from the Comoros Islands endemic species *C. humblotiana*, which in the current study is considered a member of Malagasy *Coffea* based on the geographic proximity of Madagascar and the Comoros (see Fig. 1) and the fact that molecular phylogenetic results consistently place this species within a clade of Malagasy *Coffea* (Maurin *et al.*, 2007; Nowak *et al.*, 2012). We also included four putative S-RNase alleles from *C. canephora* (CcanCc2a, CcanCc2b, CcanCc3a, CcanCc3b and CcanCc4b) and one from *C. arabica* (CaraCA1a) from Asquini *et al.* (2011). S-RNase alleles were sequenced from a total of 13 Mauritian *Coffea* individuals comprising six *C. macrocarpa* from three populations, six *C. mauritiana* from three populations and one *C. myrtifolia*. S-RNase amplification, cloning and sequencing followed methods described by Nowak *et al.* (2011). Briefly, total RNA was extracted from style tissues preserved in RNALater (Ambion, Grand Island, NY, USA) using the RNAqueous Kit (Ambion). First-strand cDNA was synthesized using MMLV reverse transcriptase (Promega, Madison, WI, USA) and a poly(T) adapter primer following the manufacturer's protocol. S-RNase alleles were amplified using the 3′ RACE (rapid amplification of cDNA ends; Frohman *et al.*, 1988) procedure with a nested PCR using reverse RACE primers (RACE-outer.R; RACE-inner.R) and a forward primer for conserved region c2 of the *Coffea* S-RNase gene (CoffBall-3.F; see Nowak *et al.*, 2011 for all primer sequences). The resulting amplicon was cloned using the TOPO TA Kit (Invitrogen, Grand Island, NY, USA) and between 8 and 12 positive colonies were sequenced to characterize S-RNase alleles.

Cloned partial S-RNase sequences were assembled and edited using the Sequencher v4.8 software (GeneCodes, Anne Arbor, MI, USA), and the resulting contigs were aligned by eye in the program MacClade v4.08 (Sinaur, Sunderland, MA, USA). Cloned sequences

were examined carefully for evidence of chimeric sequences, and when identified, these sequences were removed from the data matrix. The final data matrix containing 137 S-RNase sequences was then partitioned by codon position, and MRMODELTEST v2.3 (Nylander, 2008) was used to identify the best-fit model for each partition (position 1 = SYM + G; position 2 = GTR + I + G; position 3 = HKY + I + G). A gene tree was estimated from the final data matrix with the software MrBayes v3.2.1 (Ronquist *et al.*, 2012), and parameters of the substitution model were unlinked across the three partitions. The analysis consisted of three independent runs, each with four 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. After removing the first 1000 trees as burn-in, the 50% majority-rule consensus topology and branch lengths were estimated from the 27 000 trees sampled from the posterior distribution. All sequence alignments presented in this study are archived at Dryad and can be accessed using the study number doi: 10.5061/dryad.6r52k.

Individual-based model and simulations

With inspiration from Vincek *et al.* (1997), our model, written in Objective-C, is based on a simple island-colonization process in which a sample of diploid individuals (i.e. seeds) is drawn from an infinitely large source population with a specified number of S-alleles (A_s) of equal frequency to form an island-founding population (of size N_f) with a resulting associated S-allele diversity (A_f) that depends stochastically on the sampling of the source population. By constraining the source population to carry S-alleles at equal frequency, our model is conservative with respect to estimating the minimum founding population size needed to explain the observed S-allelic diversity in Mauritian *Coffea*. The founding population is then allowed to grow, with overlapping generations, governed by a growth rate parameter (b), with age- and mortality-related characteristics parameterized to fit *Coffea* life-history characteristics (detailed below). S-allele genotypes of pollen and of recipient plants determined fertility following the gametophytic SI system of *Coffea*, with offspring S-allele genotypes determined accordingly. Each simulation realization is considered complete when the population either goes extinct or reaches a specified maximum population size (N_{max}), set to be sufficiently large that S-alleles are extremely unlikely to be lost due to genetic drift after that point. For the simulation realizations reported in this study, N_{max} was set to 5000 (which typically meant a run length of approximately 1000–1500 generations, for $b = 0.05$), but values of $N_{max} \geq 500$ produce essentially identical results because the effects of genetic drift on the S-allelic diversity are realized almost entirely during the founding bottleneck and early population growth. The surviving S-allele diversity in the island population (A) is recorded when the realization is complete.

In the first phase of each modelled year, age-based mortality occurs. The annual probability of mortality for a given plant is calculated using the logistic function

$$P_{\text{mortality}}(x) = \frac{1}{1 + e^{-0.2(x-37)}}$$

where x is the age of the plant. The logistic function is a logical choice because the probability of mortality increases with the age of the plant. Figure S1 shows the resulting annual and cumulative probabilities of mortality. The shape parameters of the logistic function were set to 0.2 and 37 such that plants at the end of our simulations had a mean age of 11.1 years and the majority of plants had died within 37 years. These features are consistent with the expected average longevity in wild *Coffea* populations (Lashermes & Anthony, 2007; Wintgens, 2009). Modest changes to the shape parameters of the logistic function resulted in minor changes to the average lifespan, which slightly shifted the threshold value of b at which populations tended to go extinct (data not shown), but these differences do not significantly affect our conclusions. The age of every plant is incremented by one at the beginning of each year.

The second phase of each modelled year comprises pollination and the generation of seedlings. Following typical *Coffea* life history (Lashermes & Anthony, 2007), plants 5 years old and younger are considered juveniles and cannot make pollen or set seed, whereas older plants are adults. The number of seedlings that establish in a given year is determined by a sample from the binomial distribution $B(N, b)$, where N is the number of adult plants and b is the probability, per adult, that a new seedling will be generated. The model parameter b thus represents the per-individual birth rate, which, together with a given mortality function (see above), determines the population growth rate. Crosses are then conducted to generate seedlings as follows. First, two parental plants are drawn randomly from the adult population. The first parent is the pollen donor; the pollen genotype is chosen randomly from its two S-alleles. The other parent is the pollen recipient; if the pollen genotype matches either of the recipient's S-alleles, fertilization is blocked. Otherwise, an S-allele is chosen at random from the recipient (the ovule genotype), and a seedling, inheriting its S-alleles from the pollen and ovule, is added to the population with an age of zero. Crosses continue in this manner until all seedlings have been generated; an infertile cross thus does not decrease the number of seedlings ultimately generated.

This crossing procedure generates frequency-dependent selection on the S-locus, as plants with a common

S-allele will be more likely to be blocked by pollen incompatibility and will thus be less likely to be the parent of a seedling. This model design thus mechanistically implements the gametophytic SI system of *Coffea*. This SI system is also followed in the generation of the initial founding population (in particular, individuals are never homozygous at the S-locus, as that is prevented by pollen incompatibility). The generation of new S-alleles through mutation was not modelled, as six of the seven S-RNase alleles found in Mauritian *Coffea* can be confidently assigned to trans-specific allelic lineages identified by Nowak *et al.* (2011; discussed in more detail below).

S-allele diversity in the source population (A_s) was varied between 10 and 34 to reflect the S-allele diversity observed in SI plant populations of various sizes (reviewed by Castric & Vekemans, 2004); the maximum value, 34, is further supported by the number of trans-specific S-RNase allelic lineages identified in *Coffea* by Nowak *et al.* (2011). It is notoriously difficult to estimate population growth rates (Alvarez-Buylla *et al.*, 1996), and it was not possible to base the range of values explored for the population growth parameter (b) on empirical data for *Coffea* or other SI species. Thus, to achieve a representative range of estimates of founding population size, we chose a minimal value of b as the lowest rate that would yield a successful colonization in most simulation replicates ($b = 0.05$), and a maximal value of b that represents a very rapidly growing tropical understory tree population ($b = 0.1$; Alvarez-Buylla *et al.*, 1996). For the number of founding seeds (N_f), simulations were performed for every value from 1 to 50 (for each value of A_s used) with the specific intent of estimating the number of founders that would be likely to have generated the observed S-allelic diversity in Mauritian *Coffea*. The phylogenetic evidence for a monophyletic Mauritian *Coffea* suggests that the founding individuals were likely from a single ancestral species (Maurin *et al.*, 2007; Nowak *et al.*, 2012). While it is theoretically possible that multiple LDD events from the same ancestral species were involved in assembling the founding population, we feel that the geographic isolation of the island of Mauritius makes a single LDD event carrying multiple seeds the most parsimonious scenario.

Results

To test SI in the Mascarene *Coffea* lineage, we performed controlled self-pollinations in wild populations of *C. macrocarpa* and *C. mauritiana* and microscopically examined pollen tube growth. Comparing the results of this experiment with similar tests performed in the self-compatible *C. arabica* (Nowak *et al.*, 2011), it was found that none of the plants sampled showed evidence of self-compatibility (Table S3). All of the individuals sampled showed the termination of pollen tube growth shortly after entering the style tissue, characteristic of gametophytic self-incompatibility in *Coffea* (Coulibaly *et al.*, 2002; Nowak *et al.*, 2011) and other Rubiaceae (Bawa & Beach, 1983).

We performed a Bayesian divergence time analysis in the genus *Coffea* to estimate the age of colonization of the island of Mauritius. The maximum *a posteriori* ultrametric tree summarizing the results of our BEAST analysis is generally consistent with the results presented by Davis *et al.* (2011), in that many of the oldest nodes in the *Coffea* phylogeny have low support (i.e. < 0.95 posterior probability; Fig. S2). This result might be due in part to the limited sequence divergence observed at plastid compared with nuclear loci (Nowak *et al.*, 2012). The median age for the most recent common ancestor (MRCA) of the genus *Coffea* is found to be 10.55 Ma (95% HPD: 6.96–14.37 Ma; Fig. S3). Despite the limited resolution deep in the tree, several clades receive relatively high support including the Africa/Indian Ocean (A/IO) clade (PP = 0.98), which is found to be 5.37 Ma (95% HPD: 3.74–7.09 Ma). A minimum limit on the age of the colonization of Mauritius can be interpreted as the age of the MRCA of the Mascarene *Coffea*, which is found to be 1.80 Ma (95% HPD: 0.82–2.82 Ma). Similarly, a maximum limit on the age of colonization can be interpreted as the age of the stem lineage of the Mascarene *Coffea*, and this is found to be 2.64 Ma (95% HPD: 1.68–3.68 Ma). Taken together, these results suggest that the colonization of Mauritius likely occurred sometime between 1.80 and 2.64 Ma (95% HPD: 0.82–3.68 Ma). Due to low support in this part of the tree, it is still impossible to determine whether the Mascarene *Coffea* clade is more closely related to Malagasy or African species, and thus, the likely geographic source of the founding population cannot be determined.

Results from our Bayesian phylogenetic analysis allowed the identification of 43 deeply divergent and putatively functional S-RNase allelic lineages (Fig. 2). Consistent with the results of Nowak *et al.* (2011), at least 21 of these allelic lineages show evidence of trans-specific polymorphism given that they are carried by at least two distinct *Coffea* species. There is remarkable consistency in the level of sequence divergence between S-RNase allelic lineages (Fig. 2). Furthermore, each of the named allelic lineages shown in Fig. 2 that is not represented by a single sequence is strongly supported as monophyletic (PP = 1.0). The Mauritian *Coffea* lineage is found to carry S-RNase alleles belonging to seven distinct allelic lineages. The two individual *C. mauritiana* that were originally collected from the Mondrain population in Mauritius (Cmau13 and Cmau15; Table S2) appear to be fixed for S-RNase allelic lineage S3 (Fig. 2). Both of these plants carry two S3 alleles that differ from one another by 2 and 3 non-synonymous mutations, respectively. Similarly, the four individuals sampled from the Malagasy species *C. ankaranensis* appear homozygous for a version of the S30

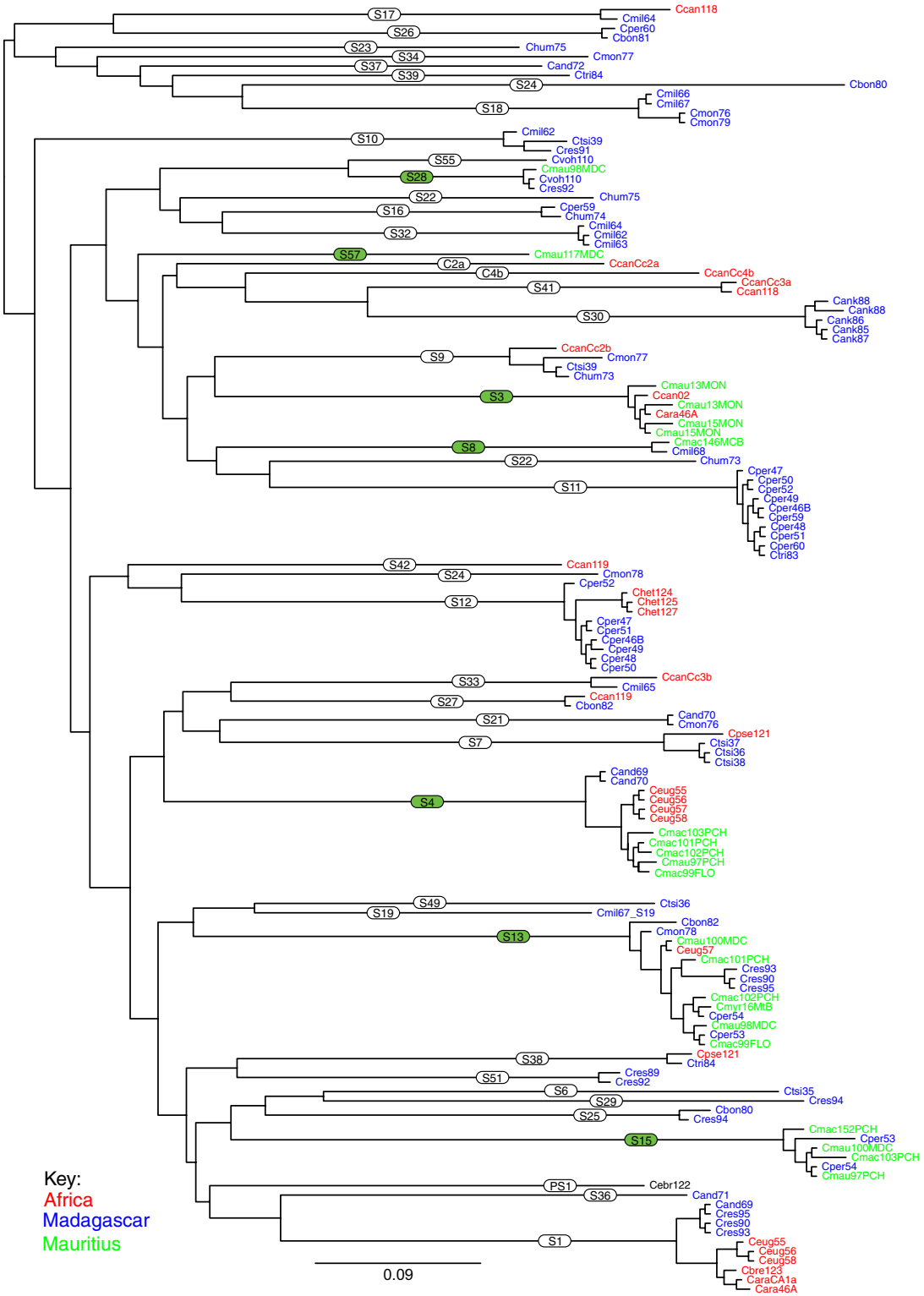

Fig. 2 *Coffea* S-RNase gene tree. This tree is a midpoint-rooted maximum *a posteriori* phylogram estimated in MrBayes v1.3.2. All deeply divergent S-RNase allelic lineages are labelled with a unique number (i.e. S17) and supported by a posterior probability of 1.0. S-RNase alleles are colour coded by geographic origin following the key provided, and the seven allelic lineages found in Mauritian *Coffea* are shown in green.

allelic lineage. This result is similar to that observed by Nowak *et al.* (2011) in the self-compatible species *C. heterocalyx* and thus could indicate that the gametophytic SI system in these populations has been lost due to the evolution of a nonfunctional allele that has swept to fixation. Evaluating this possibility will require more detailed study of the strength of the SI system in these species and a larger sample of the population.

To examine how the size of the genetic bottleneck associated with island colonization would affect the number of surviving S-alleles, we developed an individual-based model and performed a large number of simulations under varying input parameters. Experimenting with various values for the per-individual birth rate parameter b indicated that values <0.05 usually led to population extinction (Fig. S4). This threshold effect at per-individual birth rates below 0.05 is due to the fact that the population is not producing seedlings at a rate sufficient to overcome the rate of mortality. Increasing this parameter above 0.05 has the effect of reducing the impact of genetic drift on the surviving S-allele diversity (A). This is expected because as b increases, the probability that each individual in the founding population will successfully mate and produce viable offspring in the next generation also increases. Given a diverse source population carrying 30 S-alleles (Castric & Vekemans, 2004), a founding population expanding at close to the minimum population growth rate that avoids near-certain extinction ($b = 0.05$) would likely have to be composed of 14–21 individuals to produce the observed S-allelic diversity of Mauritian *Coffea* (Fig. 3b). Higher growth rates ($b > 0.05$) decrease the loss of alleles due to drift and thus decrease the likely size of the founding population (Fig. S4). For higher population growth rates ($b > 0.06$), the mean number of seeds resulting in the observed S-allelic diversity of Mauritian *Coffea* converges towards a minimum of about five; however, *Coffea* is a slow-growing understory tree, and thus, there is no evidence that the ancestral colonizing population expanded at such an exceptional rate. A somewhat surprising result of our simulations is that altering the allelic diversity of the source population (A_s) had only a limited effect on the number of surviving S-alleles in the island population when b is held constant (Fig. 3a). The simulation replicates in which a founding population is drawn from a source population of low S-allele diversity ($A_s = 10$) and grows slowly ($b = 0.05$) provide a conservative upper limit of 31 founding individuals to yield the seven surviving S-alleles observed in our sample of extant Mauritian *Coffea* populations.

Discussion

Our results suggest that the ancestral lineage of *Coffea* that colonized the island of Mauritius approximately 1.8–2.6 Ma was very likely self-incompatible and thus this lineage's successful colonization represents an exception to Baker's rule. Notably, our divergence time estimates within the genus *Coffea* are about 10 times older than the preliminary estimates reported by Anthony *et al.* (2010). This discrepancy is likely due to the fact that their estimates were based on a molecular clock analysis of the plastid *trn*L-F region alone, which shows very little sequence divergence in the genus (Anthony *et al.*, 2010; Nowak *et al.*, 2012). Future

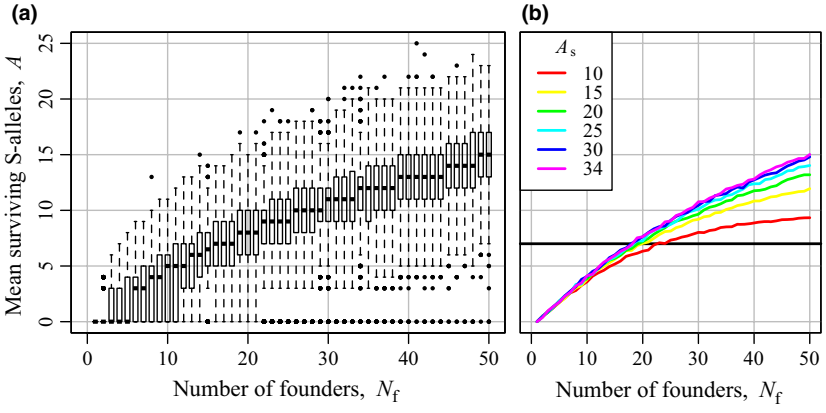

Fig. 3 (a) Results of simulations assuming a source population carrying 30 S-alleles ($A_s = 30$), performed with $b = 0.05$ and $N_{max} = 5000$. Each boxplot summarizes the number of surviving S-alleles (A) resulting from 1000 model replicates using a given number of founding individuals (N_f,) from 1 to 50. The central line of each box shows the median, the box top and bottom show the first and third quartiles, and the whiskers extend to the most extreme data point no more than 1.5 times the interquartile range from the box. (b) Plot of the mean number of surviving alleles following colonization, as a function of different levels of allelic diversity in the source population (A_s). The vertical scale is identical to that of panel A. The black horizontal line shows the observed S-allelic diversity in Mauritian *Coffea* (seven distinct S-alleles). Boxplots similar to panel A for each of these sets of simulations can be found in Fig. S5.

phylogenetic studies of *Coffea* including a more comprehensive taxon sampling and more informative molecular markers are likely to offer improved phylogenetic resolution, particularly at the base of the tree, and thus more accurate and precise divergence time estimates.

By examining allelic diversity at the *Coffea* S-RNase gene, which is preserved by NFDS (Nowak *et al*., 2011), we can make inferences about the ancient bottleneck associated with the colonization of Mauritius. The genealogy of *Coffea* S-RNase alleles shows that the bottleneck associated with the colonization of Mauritius resulted in seven deeply divergent S-RNase allelic lineages (Fig. 2). While this clearly does not represent a massive bottleneck, some perspective on its magnitude can be gained by comparing S-RNase allelic diversity isolated from the 13 individuals sampled from the Mauritian *Coffea* clade with Malagasy and African *Coffea* species. For example, we were able to sample 7 individuals from both *C. canephora* and *C. resinosa*, and each species was found to carry nine S-RNase allelic lineages. Evidence for a bottleneck in Mauritian *Coffea* is perhaps more stark when one considers that nearly all of the Malagasy and African *Coffea* species sampled in this study were collected from *ex situ* germplasm resources and have been shown to harbour reduced genetic diversity compared with wild populations (Krishnan *et al*., 2012), while all but two Mauritian samples were from wild populations (Table S2). Our estimate of seven allelic lineages assumes that allele S57 did not evolve *in situ* on Mauritius (Fig. 2), but rather has not yet been sampled from mainland species. We feel that this is the most conservative assumption in the light of our limited population-level sampling of mainland species, the substantial sequence divergence of S57 compared with other alleles and the relatively low sequence divergence observed in other S-RNase allelic lineages found in Mauritian *Coffea*.

The maintenance of allelic diversity expected at the S-locus under NFDS makes it possible to estimate the size of the founding population that was likely responsible for the colonization of Mauritius. Fruits are the unit of dispersal in *Coffea*, and each fruit carries at most two seeds (coffee 'beans'). Assuming that no deeply divergent S-allelic lineages evolved *in situ* on Mauritius, the original population could have been founded by a minimum of two fruits from carrying a total of four seeds. This represents the logical minimum number of seeds possible to initiate a founding population that has seven S-alleles. The results of our individual-based model suggest that it is most likely that more than 14 seeds were involved in the founding event and that even with a very diverse source population and the fastest population growth rate, at minimum five seeds were likely required. Furthermore, our results show that a founding population of >31 individuals would almost assuredly have resulted in S-allelic diversity greater than that observed in our sample of Mauritian *Coffea*.

Mauritius is separated from the closest native populations of *Coffea* on the Eastern Madagascar coast by approximately 800 km of ocean. It is tempting to speculate a synchronous LDD of between five and 31 seeds to this remote island accomplished via single flock avian dispersal, perhaps facilitated by a storm and accompanying high winds. While such dispersal events are expected to be rare, the global distribution of island biota shows that such events nonetheless must have occurred, with dramatic impacts on the geographic distribution of plant lineages.

Acknowledgments

Fieldwork in Madagascar was carried out under collaborative agreements between the Parc Botanique et Zoologique de Tsimbazaza (PBZT), the University of Antananarivo (Madagascar), the Association Nationale de Gestion des Aires Protegées (ANGAP) and the Royal Botanic Gardens, Kew. Field experiments at the Kianjavato Coffee Research Station were conducted with permission from the Madagascar Ministère de la Recherche Scientifique Centre Nationale de Recherche Appliquée au Développement Rural (FOFIFA). Field collections and experiments in Mauritius were conducted with permission from the Mauritian Forestry Service and the Mauritian Wildlife Foundation. Special thanks are extended to Yasmina Jaufeerally-Fakim (University of Mauritius) and Aaron Davis (Kew) for assistance in the field, and to David Lorence, Jay Horn and François Anthony for collecting tissues that were used in this study. This work was funded in part by US National Science Foundation Doctoral Dissertation Enhancement Project grant No. 0849186 to ADY and graduate student research grants from the American Society of Plant Taxonomists and Duke University Graduate School to MDN. BCH was supported by US National Science Foundation Graduate Research Fellowship Grant No. 1038597 and the ContempEvol grant (ANR 11 PDOC 005 01) from the Agence National de la Recherche of France.

References

Alsos, I.G., Eidesen, P.B., Ehrich, D., Skrede, I., Westergaard, K., Jacobsen, G.H. *et al.* 2007. Frequent long-distance plant colonization in the changing Arctic. *Science* **316**: 1606–1609.

Alvarez-Buylla, E.R., Garcia-Barrios, R., Lara-Moreno, C. & Martinez-Ramos, M. 1996. Demographic and genetic models in conservation biology: applications and perspectives for tropical rain forest tree species. *Annu. Rev. Ecol. Syst.* **27**: 387–421.

Anthony, F., Diniz, L.E.C., Combes, M.-C. & Lashermes, P. 2010. Adaptive radiation in *Coffea* subgenus *Coffea* L. (Rubiaceae) in Africa and Madagascar. *Plant Syst. Evol.* **285**: 51–64.

Asquini, E., Gerdol, M., Gasperini, D., Igic, B., Graziosi, G. & Pallavicini, A. 2011. S-RNase-like sequences in styles of *Coffea* (Rubiaceae). Evidence for S-RNase based gametophytic self-incompatibility? *Trop. Plant Biol.* **4**: 237–249.

Ayala, F.J. & Campbell, C.A. 1974. Frequency-Dependent Selection. *Annu. Rev. Ecol. Syst.* **5**: 115–138.

Ayala, F.J., Escalante, A., O'Huigin, C. & Klein, J. 1994. Molecular genetics of speciation and human origins. *Proc. Natl. Acad. Sci. USA* **91**: 6787–6794.

Baker, H. 1955. Self-compatibility and establishment after "long-distance" dispersal. *Evolution* **9**: 347–349.

Baker, H. 1967. Support for Baker's law-as a rule. *Evolution* **21**: 853–856.

Baldwin, B. 2007. Adaptive radiation of shrubby tarweeds (Deinandra) in the California Islands parallels diversification of the Hawaiian silversword alliance (Compositae-Madiinae). *Am. J. Bot.* **94**: 237.

Bawa, K. & Beach, J. 1983. Self-incompatibility systems in the Rubiaceae of a tropical lowland wet forest. *Am. J. Bot.* **70**: 1281–1288.

Bos, D.H., Gopurenko, D., Williams, R.N. & DeWoody, J.A. 2008. Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. *Evolution* **62**: 1458–1468.

Brennan, A., Harris, S. & Hiscock, S. 2006. The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L.(Asteraceae): the number, frequency, and dominance interactions of S alleles across its British range. *Evolution* **60**: 213–224.

Cain, M.L., Milligan, B.G. & Strand, A.E. 2000. Long-distance seed dispersal in plant populations. *Am. J. Bot.* **87**: 1217–1227.

Carlquist, S. 1967. The biota of long-distance dispersal. V. Plant dispersal to Pacific Islands. *Bull. Torrey Bot. Club* **94**: 129–162.

Carr, G., Powell, E. & Kyhos, D. 1986. Self-incompatibility in the Hawaiian Madiinae (Compositae): an exception to Baker's rule. *Evolution* **40**: 430–434.

Castric, V. & Vekemans, X. 2004. Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol. Ecol.* **13**: 2873–2889.

Clark, A.G. 1992. Evolutionary inferences from molecular characterization of self-incompatibility alleles. In: *Mechanisms of Molecular Evolution* (N. Takahata & A.G. Clark, eds), pp. 79–108. Sinauer Sunderland, MA.

Coulibaly, I., Noirot, M., Lorieux, M., Charrier, A., Hamon, S. & Louarn, J. 2002. Introgression of self-compatibility from *Coffea heterocalyx* to the cultivated species *Coffea canephora*. *Theor. Appl. Genet.* **105**: 994–999.

Cowie, R.H. & Holland, B.S. 2006. Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *J. Biogeogr.* **33**: 193–198.

Crawford, D.J., Archibald, J.K., Stoermer, D., Mort, M.E., Kelly, J.K. & Santos-Guerra, A. 2008. A test of Baker's law: breeding systems and the radiation of tolpis (Asteraceae) in the Canary Islands. *Int. J. Plant Sci.* **169**: 782–791.

Davis, A.P., Govaerts, R., Bridson, D.M. & Stoffelen, P. 2006. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot. J. Linn. Soc.* **152**: 465–512.

Davis, A.P., Tosh, J., Ruch, N. & Fay, M.F. 2011. Growing coffee: Psilanthus (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot. J. Linn. Soc.* **167**: 357–377.

Devier, B., Aguileta, G., Hood, M.E. & Giraud, T. 2008. Ancient trans-specific polymorphism at pheromone receptor genes in basidiomycetes. *Genetics* **181**: 209–223.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**: 1969–1973.

Figueroa, F., Gúnther, E. & Klein, J. 1988. MHC polymorphism pre-dating speciation. *Nature* **335**: 265–267.

Frohman, M.A., Dush, M.K. & Martin, G.R. 1988. Rapid Production of Full-Length cDNAs from Rare Transcripts: Amplification Using a Single Gene-Specific Oligonucleotide Primer. *Proc. Natl. Acad. Sci. USA* **85**: 8998–9002.

Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R. & Roderick, G.K. 2012. Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol. Evol.* **27**: 47–56.

Hao, J.H., Qiang, S., Chrobock, T., Kleunen, M. & Liu, Q. 2010. A test of baker's law: breeding systems of invasive species of Asteraceae in China. *Biol. Invasions* **13**: 571–580.

Heled, J. & Drummond, A.J. 2012. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Syst. Biol.* **61**: 138–149.

Igic, B. & Kohn, J. 2001. Evolutionary relationships among self-incompatibility RNases. *Proc. Natl. Acad. Sci. USA* **98**: 13167.

Igic, B., Bohs, L. & Kohn, J. 2004. Historical inferences from the self-incompatibility locus. *New Phytol.* **161**: 97–105.

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

Kao, T.-H. & Tsukamoto, T. 2004. The molecular and genetic bases of S-RNase-based self-incompatibility. *Plant Cell* **16** (Suppl): S72–S83.

van Kleunen, M., Manning, J.C., Pasqualetto, V. & Johnson, S.D. 2008. Phylogenetically independent associations between autonomous self-fertilization and plant invasiveness. *Am. Nat.* **171**: 195–201.

Krishnan, S., Ranker, T.A., Davis, A.P. & Rakotomalala, J.J. 2012. An assessment of the genetic integrity of *ex situ* germplasm collections of three endangered species of *Coffea* from Madagascar: implications for the management of field germplasm collections. *Genet. Resour. Crop Evol.* **60**: 1021–1036.

Lashermes, P. & Anthony, F. 2007. Coffee. In: *Genome Mapping and Molecular Breeding in Plants, Volume 6 Technical Crops* (C. Kole, ed.), pp. 109–118. Springer-Verlag, Berlin, Heidelberg.

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

Maurin, O., Davis, A.P., Chester, M., Mvungi, E.F., Jaufeerally-Fakim, Y. & Fay, M.F. 2007. Towards a phylogeny for *Coffea* (Rubiaceae): identifying well-supported lineages based on nuclear and plastid DNA sequences. *Ann. Bot.* **100**: 1565–1583.

Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.

McDougall, I. & Chamalaun, F.H. 1969. Isotopic dating and geomagnetic polarity studies on volcanic rocks from Mauritius, Indian Ocean. *Geol. Soc. Am. Bull.* **80**: 1419–1442.

Miller, J.S., Levin, R.A. & Feliciano, N.M. 2008. A tale of two continents: Baker's Rule and the maintenance of self-incompatibility in *Lycium* (Solanaceae). *Evolution* **62**: 1052–1065.

Miller, J.S., Kamath, A., Damashek, J. & Levin, R.A. 2011. Out of America to Africa or Asia: inference of dispersal histories using nuclear and plastid DNA and the S-RNase self-incompatibility locus. *Mol. Biol. Evol.* **28**: 793–801.

Moore, J., White, W.M., Paul, D., Duncan, R.A., Abouchami, W. & Galer, S.J.G. 2011. Evolution of shield-building and rejuvenescent volcanism of Mauritius. *J. Volcanol. Geoth. Res.* **207**: 47–66.

Nathan, R. 2006. Long-distance dispersal of plants. *Science* **313**: 786–788.

Nathan, R., Perry, G., Cronin, J.T., Strand, A.E. & Cain, M.L. 2003. Methods for estimating long-distance dispersal. *Oikos* **103**: 261–273.

Niskanen, A.K., Lohi, H., Ruokonen, M., Esparza-Salas, R., Aspi, J. & Savolainen, P. 2013. MHC variability supports dog domestication from a large number of wolves: high diversity in Asia. *Heredity* **110**: 80–85.

Nowak, M.D., Davis, A.P., Anthony, F. & Yoder, A.D. 2011. Expression and trans-specific polymorphism of self-incompatibility RNases in *Coffea* (Rubiaceae). *PLoS One* **6**: e21019.

Nowak, M.D., Davis, A.P. & Yoder, A.D. 2012. Sequence data from new plastid and nuclear COSII regions resolves early diverging lineages in *Coffea* (Rubiaceae). *Syst. Bot.* **37**: 995–1005.

Nylander, J.A.A. 2008. *MrModeltest v2.3*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Paape, T., Igic, B., Smith, S.D., Olmstead, R., Bohs, L. & Kohn, J.R. 2008. A 15-Myr-Old genetic bottleneck. *Mol. Biol. Evol.* **25**: 655–663.

Palmer, J.D. 1991. Plastid chromosome: structure and evolution. In: *The Molecular Biology of Plastids* (L. Bogorad & I.K. Vasil, eds), pp. 5–53. Academic Press, San Diego, CA.

Pannell, J. & Barrett, S. 1998. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* **52**: 657–668.

Prugnolle, F., Manica, A., Charpentier, M., Guégan, J.F., Guernier, V. & Balloux, F. 2005. Pathogen-driven selection and worldwide HLA class I diversity. *Curr. Biol.* **15**: 1022–1027.

Rambaut, A. 2012. FigTree v1.4. http://tree.bio.ed.ac.uk/software/figtree.

Rambaut, A. & Drummond, A.J. 2007. Tracer v1.5. http://beast.bio.ed.ac.uk/Tracer.

Rambuda, T.D. & Johnson, S.D. 2004. Breeding systems of invasive alien plants in South Africa: does Baker's rule apply? *Divers. Distrib.* **10**: 409–416.

Reboiro-Jato, D., Reboiro-Jato, M., Fdez-Riverola, F., Vieira, C.P., Fonseca, N.A. & Vieira, J. 2012. ADOPS – automatic detection of positively selected sites. *J. Integr. Bioinform.* **9**: 200.

Richman, A. & Kohn, J. 2000. Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Mol. Biol.* **42**: 169–179.

Richman, A., Uyenoyama, M. & Kohn, J. 1996. Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. *Science* **273**: 1212.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S. *et al.* 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**: 539–542.

Roux, C., Pauwels, M., Ruggiero, M.V., Charlesworth, D., Castric, V. & Vekemans, X. 2013. Recent and ancient signature of balancing selection around the S-Locus in *Arabidopsis halleri* and A. lyrata. *Mol. Biol. Evol.* **30**: 435–447.

Schnabel, A. & Wendell, J.F. 1998. Cladistic biogeography of Gleditsia (Leguminosae) based on ndhF and rpl16 chloroplast gene sequences. *Am. J. Bot.* **85**: 1753–1765.

Sun, M. & Ritland, K. 1998. Mating system of yellow starthistle (*Centaurea solstitialis*), a successful colonizer in North America. *Heredity* **80**: 225–232.

Swofford, D.L. 2003. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland, MA.

Takahata, N. & Nei, M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**: 967–978.

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

Vincek, V., O'Huigin, C., Satta, Y., Takahata, N., Boag, P.T., Grant, P.R. *et al.* 1997. How large was the founding population of Darwin's finches? *Proc. R. Soc. Lond. B Biol.* **264**: 111–118.

Wintgens, J.N. 2009. *Coffee: Growing, Processing, Sustainable Production. A Guidebook for Growers, Processors, Traders and Researchers*. Wiley-VCH, Weinheim.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1 The annual probability of mortality used in the simulation model (black curve), and the resulting cumulative probability of mortality (red curve).

Figure S2 Maximum *a posteriori* chronogram produced from the BEAST v1.7.5 analysis. Branches are labelled with posterior probability.

Figure S3 Maximum *a posteriori* chronogram produced from the BEAST v1.7.4 analysis.

Figure S4 Summary of supplemental simulations run using different values for the population growth rate b (see Table S4 for the parameter values used).

Figure S5 Complete simulation results for various values of the source population S-allelic diversity (A_s), ranging from 10 to 34 alleles, for $b = 0.05$ and $N_{max} = 5000$.

Table S1 Sequence data used in BEAST divergence time analyses.

Table S2 Plant materials used in the study of *Coffea* S-RNase allelic diversity.

Table S3 Results of self-incompatibility tests performed in a wild sympatric population of *C. macrocarpa and C. mauritiana* growing at the Plaine Champagne (PCH) locality in the Black River Gorges National Park of Mauritius.

Table S4 Parameters, with their symbols and values, for the supplementary model realizations shown in Fig. S4.

Data deposited at Dryad: doi:10.5061/dryad.6r52k

Received 1 November 2013; revised 14 March 2014; accepted 1 April 2014