npg

ORIGINAL ARTICLE

Host-associated divergence and incipient speciation in the yucca moth *Prodoxus coloradensis* (Lepidoptera: Prodoxidae) on three species of host plants

CS Drummond[1,3], H-J Xue[1,2,3], JB Yoder[1] and O Pellmyr[1]

[1]*Department of Biological Sciences, University of Idaho, Moscow, ID, USA and* [2]*Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China*

A wide range of evolutionary processes have been implicated in the diversification of yuccas and yucca moths, which exhibit ecological relationships that extend from obligate plant–pollinator mutualisms to commensalist herbivory. *Prodoxus coloradensis* (Lepidoptera: Prodoxidae) is a yucca moth, which feeds on the flowering stalks of three *Yucca* species as larvae, but does not provide pollination service. To test for evidence of host-associated speciation, we examined the genetic structure of *P. coloradensis* using mitochondrial (cytochrome oxidase I) and nuclear (elongation factor 1 alpha) DNA sequence data. Multilocus coalescent simulations indicate that moths on different host plant species are characterized by recent divergence and low levels of effective migration, with large effective population sizes and considerable retention of shared ancestral polymorphism. Although geographical distance explains a proportion of the mitochondrial and nuclear DNA variation among moths on different species of *Yucca*, the effect of host specificity on genetic distance remains significant after accounting for spatial isolation. The results of this study indicate that differentiation within *P. coloradensis* is consistent with the evolution of incipient species affiliated with different host plants, potentially influenced by sex-biased dispersal and female philopatry.
Heredity (2010) **105,** 183–196; doi:10.1038/hdy.2009.154; published online 16 December 2009

Introduction

Host plant specialization is a key mechanism for the diversification of phytophagous insects at both macro-evolutionary and microevolutionary time scales (Ehrlich and Raven, 1964; Diehl and Bush, 1984; Mitter *et al.*, 1988; Farrell, 1998; Drès and Mallet, 2002; Nason *et al.*, 2002; Miller *et al.*, 2003; Ferrari *et al.*, 2006). Under the Dobzhansky-Muller model of speciation, host shifts can result in genetic differentiation as a consequence of assortative mating and reduced gene flow, ultimately leading to isolation and speciation due to the accumulation of mutations that cause reproductive incompatibility among populations (Dobzhansky, 1951; Coyne and Orr, 2004). However, evidence for host-related speciation can be overlooked when recent divergence leads to the retention of shared ancestral polymorphism (Maddison, 1997; Rosenberg, 2002; Degnan and Rosenberg, 2006) or confounded when genetic differentiation occurs as a consequence of biogeography rather than host specificity (Becerra and Venable, 1999). Because these circumstances are not mutually exclusive, teasing apart the population-level relationships between host use and speciation remains a principal challenge for understanding the evolution of insect–plant associations.

Yuccas (*Yucca*: Agavaceae) and yucca moths (*Prodoxus, Parategeticula, Tegeticula*: Prodoxidae) have provided extraordinary opportunities for the study of divergence and speciation, illustrating the importance of host plant specificity for diversification (reviewed in Pellmyr, 2003; see Althoff *et al.*, 2006; Leebens-Mack and Pellmyr, 2004; Segraves and Pellmyr, 2004; Segraves *et al.*, 2005; Svensson *et al.*, 2005; Althoff, 2008; Godsoe *et al.*, 2008; Smith *et al.*, 2008a, b for more recent work). Yucca moths in the genus *Prodoxus* are a derived lineage within the Prodoxidae, comprising the sister lineage of the pollinating yucca moths *Parategeticula* and *Tegeticula* (Pellmyr and Leebens-Mack, 1999; Pellmyr, 2003; Pellmyr *et al.*, 2006). The pollinating yucca moths have coevolved as obligate mutualists with various host species of *Yucca*, providing exclusive pollination service and ovipositing so that emerging larvae feed on seeds. In contrast, members of *Prodoxus* do not pollinate their host plants, and instead oviposit so that larvae typically feed on tissues other than seeds (Pellmyr, 2003; Pellmyr *et al.*, 2006).

Non-pollinating *Prodoxus* always coexist with pollinating yucca moths, as the latter are required for successful reproduction of yuccas; moreover, up to three different

Correspondence: *Dr CS Drummond, Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA.*
E-mail: cdrummon@uidaho.edu
[3]*These authors contributed equally to this work.*

species of *Prodoxus* may coexist on the same plant, respectively specializing on fruit, flowering stalks or leaves (Pellmyr *et al.*, 2006). Within *Prodoxus*, stalk feeding is the most common mode of host plant use, from which a more limited number of fruit or leaf-feeding taxa have evolved (Pellmyr *et al.*, 2006). Although obligate mutualisms are expected to promote coevolutionary divergence in both yuccas and yucca moths (Pellmyr and Krenn, 2002; Pellmyr and Segraves, 2003; Leebens-Mack and Pellmyr, 1998, 2004; Althoff *et al.*, 2006; Smith *et al.* 2008a; but see Smith *et al.* 2008b), the extent to which commensalist host specialization might lead to divergence among non-pollinating yucca moths is less clear. In the best-studied example to date, *Prodoxus quinquepunctellus* and *Prodoxus decipiens* have been shown to exhibit rapid phenological and morphological divergence among populations on different host species, particularly the evolution of ovipositor shape, a feature that is critical for successful oviposition and reproduction of yucca moths (Groman and Pellmyr, 2000; Althoff *et al.*, 2001; Althoff and Pellmyr, 2002; Svensson *et al.*, 2005). These results suggest that other members of *Prodoxus* are capable of highly specialized reproductive

behavior, which could promote both host specificity and speciation.

Prodoxus coloradensis (Prodoxidae: Lepidoptera) is a stalk-feeding yucca moth endemic to southwestern North America, extending from coastal California to the Rio Grande in Texas (Figure 1; Pellmyr *et al.*, 2006). Earlier taxonomic treatments of *P. coloradensis* indicated that this taxon was a geographically widespread generalist affiliated with multiple species of *Yucca* (Riley, 1880; Davis, 1967). However, a more recent revision of the Prodoxidae based on morphological and molecular data has narrowed the current status of *P. coloradensis* to a complex associated with three host species: *Yucca baccata*, *Yucca schidigera* and *Yucca treculeana* (Pellmyr *et al.*, 2006). These host plants are parapatric, with limited areas of sympatry at their range margins (Figure 1; Matuda and Lujan, 1980; Hess and Robbins, 2002); within these zones, altitudinal and/or ecological segregation may further limit contact between moths on different host plants. Although *P. coloradensis* is known to be variable in wing coloration, which provides a useful diagnostic character for discriminating between other closely related species
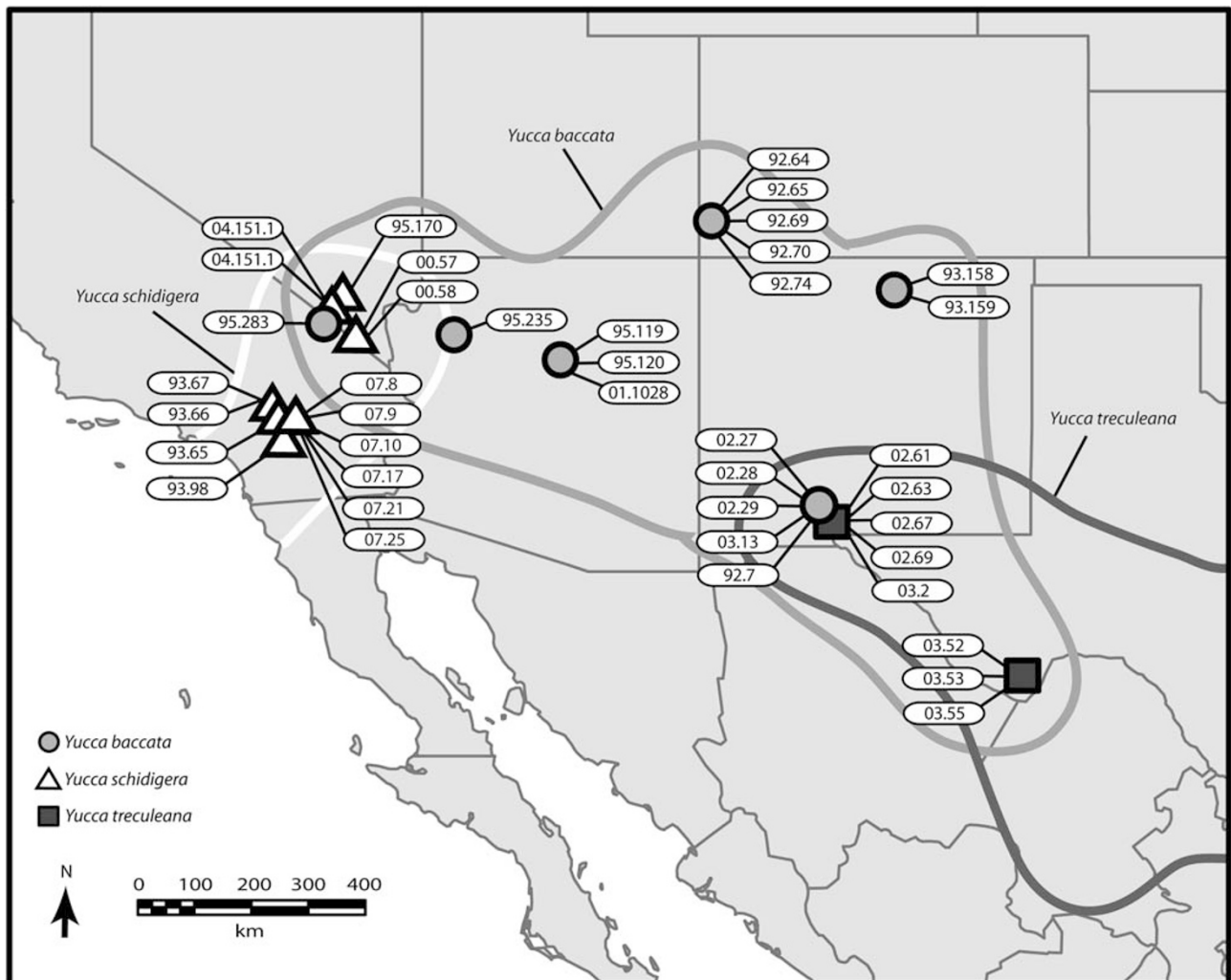

Figure 1 Geographical distribution of *P. coloradensis* specimens sampled for this study, with approximate ranges for associated host species of *Yucca* in southwestern North America. Numbers in ovals refer to accession numbers for each specimen (cf. Figures 2 and 3).

of *Prodoxus* (Pellmyr *et al.*, 2006), there has been no clear morphological evidence in support of separate taxonomic status for host-associated populations within the current delimitation of *P. coloradensis*.

For this study, we tested the hypothesis that *P. coloradensis* represents a complex of three different host-associated species. On the basis of mitochondrial DNA (mtDNA) and nuclear DNA (nucDNA) sequence data, we evaluated the genealogical relationships of stalk-feeding moths sampled from *Y. baccata*, *Y. schidigera* and *Y. treculeana*. As shared ancestral polymorphism and recent divergence can lead to the rejection of strict phylogenetic criteria for speciation (for example, reciprocal monophyly, Knowles and Carstens, 2007), we adopted a population genetic approach to test for significant differentiation among host groups and estimate effective migration rates using multilocus coalescent simulations. Because geographical ranges of the host species for *P. coloradensis* are not widely overlapping, we also examined whether evidence for genetic structure could be attributed to host use after accounting for spatial isolation. Finally, we evaluated differential rates of male versus female dispersal on the basis of relative rates of gene flow for maternally inherited mtDNA compared with biparentally inherited nucDNA.

Methods

Field collections

Adult ($n = 26$) and larval specimens ($n = 14$) were sampled from 15 locations across the range of *P. coloradensis* in southwestern North America during 1992–2007, then flash-frozen in liquid nitrogen and placed in permanent storage at $-80\,^{\circ}\text{C}$. Sample sites and the approximate ranges of each host *Yucca* species are shown in Figure 1. Specific collection localities and GPS coordinates are listed in Supplementary Table S1). Because neither adult moths nor larvae are readily identified as distinct morphological species, sampling was designed to estimate migration between populations of moths on different host plants. Consequently, throughout the remainder of the paper, for brevity we often refer to moths collected from *Y. baccata*, *Y. schidigera* and *Y. treculeana* by the name of the host species.

DNA sequencing

Adult moths were dissected to remove the thorax for DNA extraction, retaining the wings, head and abdomen as voucher specimens. Larvae were used to obtain the remaining DNA samples. Tissues dissected from each adult moth (or whole larvae) were macerated in liquid nitrogen, and genomic DNA was isolated using DNeasy extraction kits (Qiagen, Valencia, CA, USA) for blood and animal tissue. Standard PCR protocols for yucca moths were used to amplify 1482 bp of subunit I from the mitochondrial cytochrome oxidase I (*COI*) gene using three overlapping PCR amplifications (Smith *et al.*, 2008a). A modified touchdown PCR protocol was used to amplify 502 bp of the nuclear elongation factor 1 alpha (*EF1α*) gene (Smith *et al.*, 2008a). PCR products were purified using QIAquick PCR purification kits (Qiagen), followed by cycle sequencing according to standard protocols for BigDye 3.1 (Applied Biosystems, Foster City, CA, USA). Dye terminator fragments were cleaned

using Sephadex and electrophoresed on an ABI 3730 capillary instrument (Applied Biosystems). Trace files were aligned and edited using CodonCode Aligner 2.02 (CodonCode, Dedham, MA, USA). We were unable to obtain complete reads of EF1α for one individual, which was excluded from subsequent analyses. Variable sites in the remaining sequences were identified using the automated mutation detection function in CodonCode Aligner, followed by manual inspection of electropherograms and additional resequencing of ambiguous bases. Successful PCR amplification and sequencing of target genes (rather than paralogs or PCR artifacts) were confirmed by aligning the resulting sequences to other closely related species in the Prodoxidae, verifying the correct reading frame without stop codons.

As the resolution of nucDNA sequences into allelic haplotypes provides additional information for coalescent analyses of population structure (see below), the gametic phase of heterozygous sites (that is, double peaks) in EF1α sequences were defined in PHASE v. 2.1.2 (Stephens *et al.*, 2001; Li and Stephens, 2003; Stephens and Donnelly, 2003) using 10 000 Markov chain Monte Carlo (MCMC) iterations with 1000 burn-in iterations and a sampling interval of 10 iterations. Convergence was assessed by comparing the posterior probability (PP) of all haplotype resolutions between three replicate PHASE runs, increasing the posterior sampling until PPs differed by no more than 0.01 between runs. Heterozygous sites were coded as distinct alleles when resolved with > 0.90 PP, otherwise these sites were treated as the missing data at that position. We also estimated the background recombination rate ($\rho = 4N_e c$) and deviations from ρ between each variable site.

After assembling aligned DNA sequence data matrices, we used DNaSP 5.0 (Librado and Rozas, 2009) to calculate the following summary statistics for each gene in the total sample of individuals: (i) number of haplotypes (h); haplotypic diversity (H_d; Nei, 1987); average number of nucleotide differences (K; Tajima, 1983); (iv) nucleotide diversity (π; Nei, 1987). The distribution of COI haplotypes and EF1α alleles among hosts was also summarized using statistical parsimony networks inferred in TCS 1.2.1 (Clement *et al.*, 2000). GenBank accessions for DNA sequences are listed in Supplementary Table S1).

Gene tree estimation

Phylogenies for COI and EF1α were estimated using MCMC in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003), including seven additional outgroup taxa in the Prodoxidae (see Supplementary Table S1 for GenBank accessions of outgroup species). Nucleotide substitution models ($COI = GTR + I + \Gamma$; $EF1\alpha = GTR + \Gamma$) were selected based on congruent results from hierarchical likelihood-ratio tests (Posada and Crandall, 1998) and AIC scores (Posada and Buckley, 2004) among competing models in MrModeltest 2.3 (Nylander, 2008). Because the default branch length priors in MrBayes can lead to unreasonable estimates of substitution rates (for example, > 1.0 substitution per site from root to tip) and upward bias of PPs (Yang and Rannala, 2005; Marshall *et al.*, 2006), empirical priors for this parameter were defined based on the mean branch lengths recovered from maximum likelihood topologies inferred using

GARLI 0.951 (Zwickl, 2006). MCMC analyses were executed for 200 000 burn-in iterations followed by 1 800 000 iterations of posterior sampling using two independent runs with four chains each and a sampling interval of 100 iterations. After posterior sampling was completed, we verified that the average s.d. of split frequencies for log-likelihood scores across paired runs was <0.01, with potential scale reduction factors for all parameters ~1.000. Log-likelihood scores and parameter estimates were examined in TRACER v 1.4.1 (Rambaut and Drummond, 2008) to confirm that independent runs had reached stationarity and convergence, with effective sample sizes >200 for all parameters. MCMC posterior distributions of tree topologies and branch lengths were summarized in TreeAnnotator 1.4.8 (Drummond and Rambaut, 2007) as the maximum clade credibility tree, which represents the single tree recovered from the posterior distribution that maximizes the product of PPs across clades.

To obtain empirical estimates of COI and EF1α substitution rates on an absolute time scale for mutation rate priors in the isolation-with-migration coalescent analyses (see below), we analyzed both genes under a log-normal relaxed clock model in BEAST 1.4.8 (Drummond et al., 2002, 2006; Drummond and Rambaut, 2007) using a representative set of sequences from the Prodoxid moths, comprising a single taxon from each of the species used for phylogenetic inference of the COI and EF1α gene trees (Supplementary Table S1). We constrained the ages of two key nodes using a normal distribution with 95% confidence intervals (CI) approximating the means and s.e. of estimated divergence times from Pellmyr and Leebens-Mack (1999): most recent common ancestor (MRCA) of Lampronia and yucca moths = 78.2 million years (Myr), 95% CI = 59.8–96.6 Myr; MRCA yucca moths = 41.5 Myr, 95% CI = 31.7–51.3 Myr. MCMC analyses in BEAST were executed under a simple Yule birth model and allowed to run for 100 000 burn-in iterations followed by 10 000 000 iterations of posterior sampling, using two independent runs with a sampling interval of 1000 iterations. Stationarity and convergence of independent runs were confirmed by examining PPs, parameter estimates and effective sample sizes of MCMC output in TRACER (see above). Posterior distributions of time-scaled substitution rates for each locus were estimated as the mean and 95% highest posterior density (HPD) of this parameter.

Population structure

We examined the degree of population structure among moths on different host species using analysis of molecular variance (AMOVA) (Excoffier et al., 1992) in Arlequin 3.11 (Excoffier et al., 2005). For each gene, we estimated the proportion of genetic variation partitioned between host plants, calculated fixation indices (Φ_{ST}), and tested for significant differentiation using 10 000 nonparametric permutations. These analyses did not include substructure within host groups due to limited sampling at any given collection locality (see below). AMOVAs were performed under the Tamura–Nei substitution model (that is, the closest approximation to GTR available in Arlequin), with empirical estimates from MrBayes for the Γ parameter of each gene (COI:

$\alpha = 0.191$; EF1α: $\alpha = 0.192$). Because differentiation between a subset of populations can indicate significant departures from panmixia, even when other populations are not genetically distinct, we also tested comparisons between each pair of host plants to determine whether global patterns of differentiation were supported by pairwise comparisons between host groups.

As AMOVA does not account for the spatial effects of gene flow across geographical landscapes (for example, isolation by distance; Rousset, 1997), we used a series of Mantel and partial Mantel tests (Mantel, 1967; Smouse et al., 1986) to examine correlations between pairwise genetic distance, geographical distance and host use. Preliminary analyses based on pairwise genetic distances between localities were difficult to interpret, particularly for sites with ⩽2 individuals, which commonly yielded pairwise estimates of $F_{ST} = 1.0$ (that is, no shared haplotypes or alleles) or $F_{ST} = 0.0$ (that is, identical haplotypes or alleles). Because these results are more likely to represent an artifact of limited sampling within each collection locality (Figure 1) rather than true patterns of biological differentiation, we instead calculated pairwise genetic distances between individual sequences for each gene. Genetic distances between COI haplotypes and EF1α alleles were estimated using maximum composite likelihood (Tamura et al., 2004) in MEGA 4.0.2 (Kumar et al., 2008) under the same model of nucleotide substitution used for AMOVA (that is, Tamura–Nei with α values as described above). Geographical distances were determined from GPS coordinates using Geographic Distance Matrix Generator 1.2.3 (Ersts, 2009). Host use was coded as a binary indicator variable, such that any two individuals either share the same host species (0) or use different host species (1). Mantel and partial Mantel tests were performed in zt 1.0 (Bonnet and Van de Peer, 2002) using 10 000 nonparametric permutations for significance testing.

Coalescent simulations

Estimates of demographic values for divergence times (T_{div}), effective migration rates (M_e) and effective population sizes (N_e) among populations of moths on different host plants were obtained using multilocus coalescent simulations in IM and IMa (Hey and Nielsen, 2004, 2007). This isolation-with-migration approach permits the estimation of demographic parameters while evaluating the likelihood of more or less complex models of historical divergence (reviewed in Hey, 2006). First, we used IMa to simulate genealogies under the full model ($\theta_1, \theta_2, \theta_A, m_1, m_2, t$) and perform likelihood-ratio tests to evaluate the fit of nested models with fewer parameters (for example, (i) $\theta_1 = \theta_2 = \theta_A$, m_1, m_2; (ii) $\theta_1, \theta_2, \theta_A$, $m_1 = m_2$; (iii) $\theta_1 = \theta_2 = \theta_A$, $m_1 = m_2$; and so on). Although IMa generates point estimates from the joint posterior distributions of nested models, it does not calculate Bayesian posterior credibility intervals for these functions. Therefore, we reanalyzed the data in IM under the least complex nested model that could not be rejected, both to reduce the variance in parameter estimates and to obtain 90% HPD intervals as a measure of uncertainty.

The isolation-with-migration model assumes that sampled taxa represent sister lineages, which do not experience gene flow from other unsampled lineages. As there is often no clear a priori approach for ordering

pairwise IM and IMa analyses among three or more populations, for completeness we examined all first-order pairwise comparisons among host plant groups ($Y.$ $baccata$ versus $Y.$ $schidigera$, $Y.$ $baccata$ versus $Y.$ $treculeana$, $Y.$ $schidigera$ versus $Y.$ $treculeana$). Initial ranges on prior probabilities for coalescent parameters scaled by the population mutation rate were established using estimates of $\theta = 4 N_e \mu$ for the total sample of $P.$ $coloradensis$ and the subsamples of moths on different host plants. These estimates were obtained using a series of simple IMa models assuming no divergence (that is, $t = 0$), describing panmictic populations at equilibrium ($P.$ $coloradensis$, $\theta = 10.2$–24.0 90% HPD; $Y.$ $baccata$, $\theta = 6.7$–16.6 90% HPD; $Y.$ $schidigera$, $\theta = 3.3$–12.6 90% HPD; $Y.$ $treculeana$, $\theta = 2.9$–10.6 90% HPD). We conservatively set the upper bound for a uniform prior on θ to exceed the highest value from initial estimates for panmictic populations (that is, $\theta = 0$–30). The upper bound on m was set using a uniform prior distribution that allowed extremely high levels of effective migration between populations ($m = 0$–10, $M_e = \theta m / 2 = 30 \times 10 / 2 = 150$). On the basis of the posterior distributions for the rate of substitutions per site per million year from BEAST (Supplementary Figure S1; COI: mean $= 1.894 \times 10^{-3}$, 95% HPD $= 1.341 \times 10^{-3}$–2.488×10^{-3}; EF1α: mean-3.128×10^{-3}, 95% HPD $= 1.616 \times 10^{-3}$–4.955×10^{-3}), we set uniform priors on the means and ranges for the substitutions per locus per year, discounting ambiguous or missing sites excluded by IMa (COI, 1454 bp: mean-2.753×10^{-6}, lower $= 1.950 \times 10^{-6}$, upper $= 3.617 \times 10^{-6}$; EF1α, 485 bp: mean $= 1.517 \times 10^{-6}$, lower $= 0.783 \times 10^{-6}$, upper $= 2.403 \times 10^{-6}$). Priors for t were determined using the earliest median date for the MRCA of $P.$ $coloradensis$ and $P.$ $sordidus$ for COI or EF1α (Supplementary Figure S2; 15.6 Myr for COI), then calculating the maximum coalescent scaled bound on t as the divergence time multiplied by the geometric mean (μ_g) of substitutions/locus/year ($t = 0$–31.9, $T_{div} = t / \mu_g = 31.9 / 2.044 \times 10^{-6} = 15.6$ Myr), permitting considerable freedom for divergence times within $P.$ $coloradensis$.

On the basis of these initial prior distributions, we ran a series of preliminary IMa analyses to optimize MCMC settings for the number of chains and heating scheme. A small number of individuals with missing data at variable sites (that is, six individuals comprising 12 nucDNA alleles) were removed from the EF1α data matrix, as IMa excludes ambiguous sites across the complete set of sequences. Subsequent IMa runs were executed using a geometric heating scheme with 20–50 chains and heating parameters $g_1 = 0.95$–0.99 and $g_2 = 0.80$–0.50 under the HKY model of nucleotide substitution for both loci. Posterior sampling began after 500 000 burn-in iterations and was allowed to continue for 5 000 000–15 000 000 iterations. To evaluate stationarity and convergence, we monitored effective sample sizes and autocorrelation values, inspected trend plots for all parameters, and compared the results of at least three independent runs to confirm that marginal posterior distributions had reached similar solutions.

Following the completion of MCMC simulations in IMa, we used the load trees mode to subsample 20 000 genealogies across replicate runs and conduct likelihood-ratio tests to evaluate the relative fit of reduced models versus the full model, following a χ^2 approximation in which the test statistic is twice the difference in log-probability scores (-2Λ) between nested models, with degrees of freedom equal to the difference in the number of parameters (Hey and Nielsen, 2007). When one or both migration parameters were fixed at zero, we calculated significance where the expected distribution is a mixture, such that -2Λ has a value of zero with probability 0.5 and takes a value from the χ^2 distribution with probability 0.5 (Hey and Nielsen, 2007). In addition, when both migration parameters were set to zero, we evaluated this model against the corresponding model in which both migration parameters were equal, rather than the full model with asymmetric migration rates, as this represents a better approximation when m_1 and m_2 are correlated (Hey and Nielsen, 2007).

Following the evaluation of nested models, we then executed a series of IM runs under the least complex model that could not be rejected based on the likelihood-ratio tests described above, using the same prior probability distributions and MCMC settings. This approach has the advantage of estimating 90% HPD intervals for each parameter (which are not available for nested models in IMa), as well as reducing variance due to overparameterization of the full model. Run times varied from 5 000 000–25 000 000 iterations, and stationarity and convergence across independent IM runs were assessed as described above. For the final results, we used the output from the longest run for each pairwise comparison between populations. Coalescent parameters were converted to demographic values using the following formulae, incorporating the geometric mean of substitutions per locus per year ($\mu_g = 2.044 \times 10^{-6}$) and assuming a generation time of 1 year: (i) divergence time, $T_{div} = t / \mu_g$; (ii) effective population size, $N_e = \theta / 4\mu_g$; (iii) effective number of migrants ($M_e = \theta m / 2$, calculated from the marginal posterior distribution of m conditioned on the mode of θ). We also estimated the coalescent time to most recent common ancestry (TMRCA) of each locus, as well as locus-specific distributions for the mean number of migration events in simulated genealogies (M_{sim}). To verify that marginal posterior distributions for these parameters were informed by the data, rather than representing an artifact of the prior distributions, each pairwise analysis was repeated with the $-j0$ run option in IM, which sets all likelihood functions to a value of one, so that the posterior distributions are equal to the prior distributions.

Results

Summary statistics for molecular variation in COI and EF1α indicated substantially higher levels of genetic diversity in mtDNA versus nucDNA (COI: $h = 33$, $H_d = 0.979$, $K = 27.612$, $\pi = 0.019$; EF1α: $h = 15$, $H_d = 0.642$, $K = 4.167$, $\pi = 0.010$). The mode of the posterior distribution for the background recombination rate in EF1α was extremely low ($\rho = 2.0 \times 10^{-8}$), indicating that the frequency of recombination (scaled to N_e) was negligible, with little deviation from ρ between variable sites (none > 1.444 relative to ρ). The parsimony network for COI (Figure 2) showed that no mtDNA haplotypes were shared among host groups, with the highest frequency haplotype comprising 12.5% of the total number of inferred haplotypes ($n = 28$). In contrast, the parsimony network for EF1α (Figure 2) indicated that a
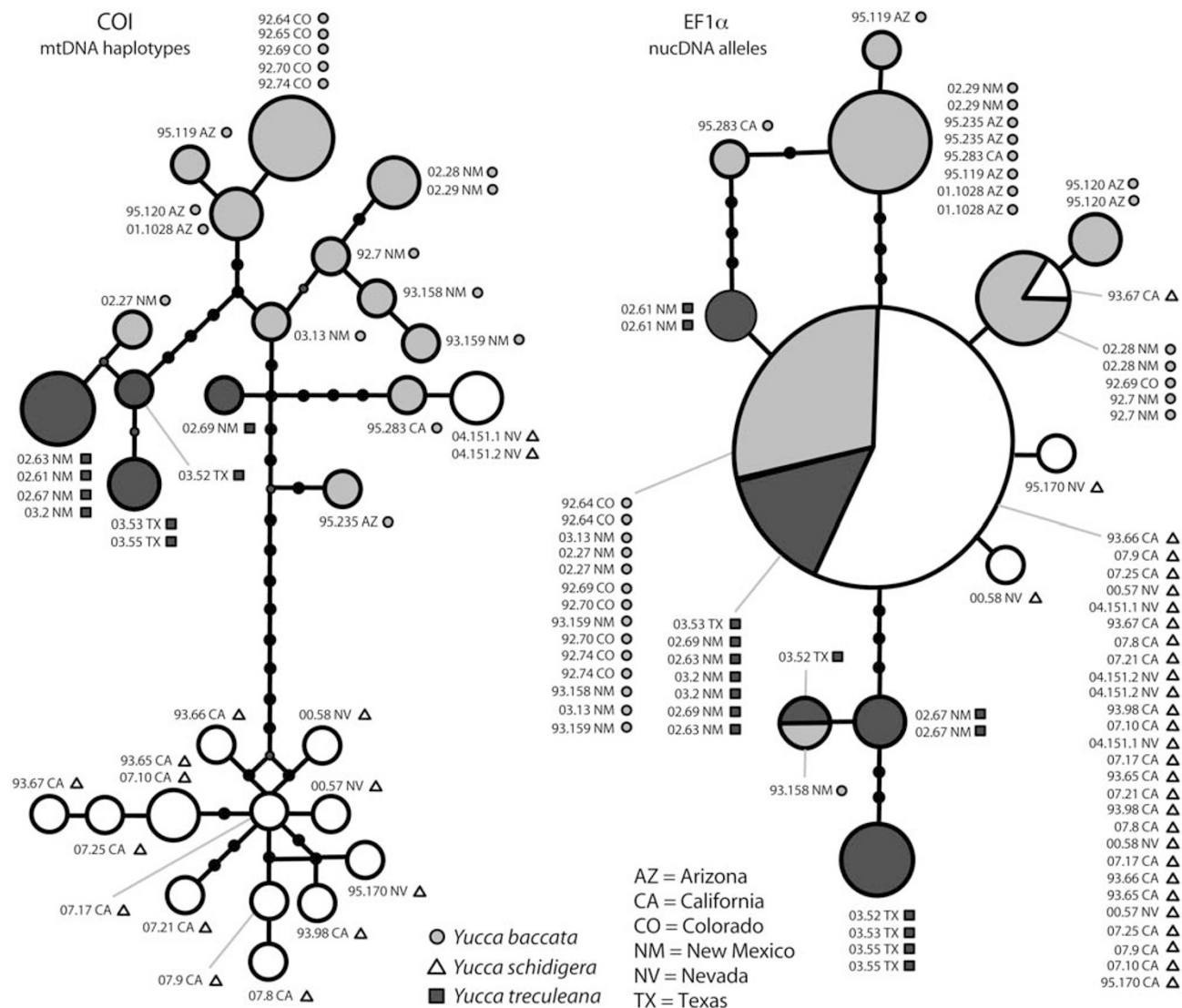
Figure 2 Statistical parsimony networks for cytochrome oxidase I (COI) and elongation factor 1 alpha (EF1α) from *P. coloradensis* on three species of *Yucca*. Circles are proportional to the frequency of unique haplotypes; branches connecting haplotypes indicate the number of inferred mutations. Labels indicate accession, sampling location, and host plant (cf. Figure 1). COI has 28 unique haplotypes, none of which are shared among host plants. In contrast, EF1α has 12 unique alleles, 3 of which are shared among host plants, including a single allele accounting for 61.5% of the total number of sampled alleles.

single nucDNA allele accounted for 61.5% of inferred alleles ($n = 12$), and this allele was shared among all the three host groups. In addition, a single nucDNA allele was shared by *Y. schidigera* and *Y. baccata*, and a single nucDNA allele was shared by *Y. baccata* and *Y. treculeana*. Note that the number of inferred haplotypes and alleles varied slightly in the results from DNaSP and TCS due to the manner in which each program handles the missing and/or ambiguous sites.

Posterior probabilities from MCMC consensus trees for COI and EF1α (Figure 3) showed support for the monophyly of *P. coloradensis* (COI: PP = 1.0; EF1α: PP = 0.91). Within *P. coloradensis*, the posterior distribution of the COI gene trees recovered a series of lineages corresponding approximately to the three host species of *Yucca*. However, these clades were not monophyletic with respect to host plants. For example, although all mtDNA haplotypes collected from *Y. baccata* and *Y. treculeana* were placed in a single clade with strong

support (PP = 1.0), *Y. baccata* formed a paraphyletic assemblage within this clade, including a single haplotype from two individuals sampled from *Y. schidigera*, as well as a single haplotype from *Y. treculeana*. Remaining haplotypes from *Y. schidigera* were placed in a clade sister to the *Y. baccata*–*Y. treculeana* group, albeit with low support (PP = 0.59). Remaining haplotypes from *Y. treculeana* were placed in a derived clade (PP = 1.0) that included a single haplotype from *Y. baccata*. In contrast to COI, support for clades in the posterior distribution of EF1α gene trees was uniformly low, with most clades having negligible support (PP ⩽ 0.01) among a large number of equally likely topologies. Very few nucDNA lineages were recovered with PP > 0.90, including: (i) a single allele from *Y. treculeana* and a single allele from *Y. baccata* (PP = 0.95); (ii) a single allele shared by three individuals of *Y. treculeana* (PP = 1.0); and (iii) two alleles each found in two different individuals of *Y. schidigera*, (PP = 0.93, 1.0), respectively.
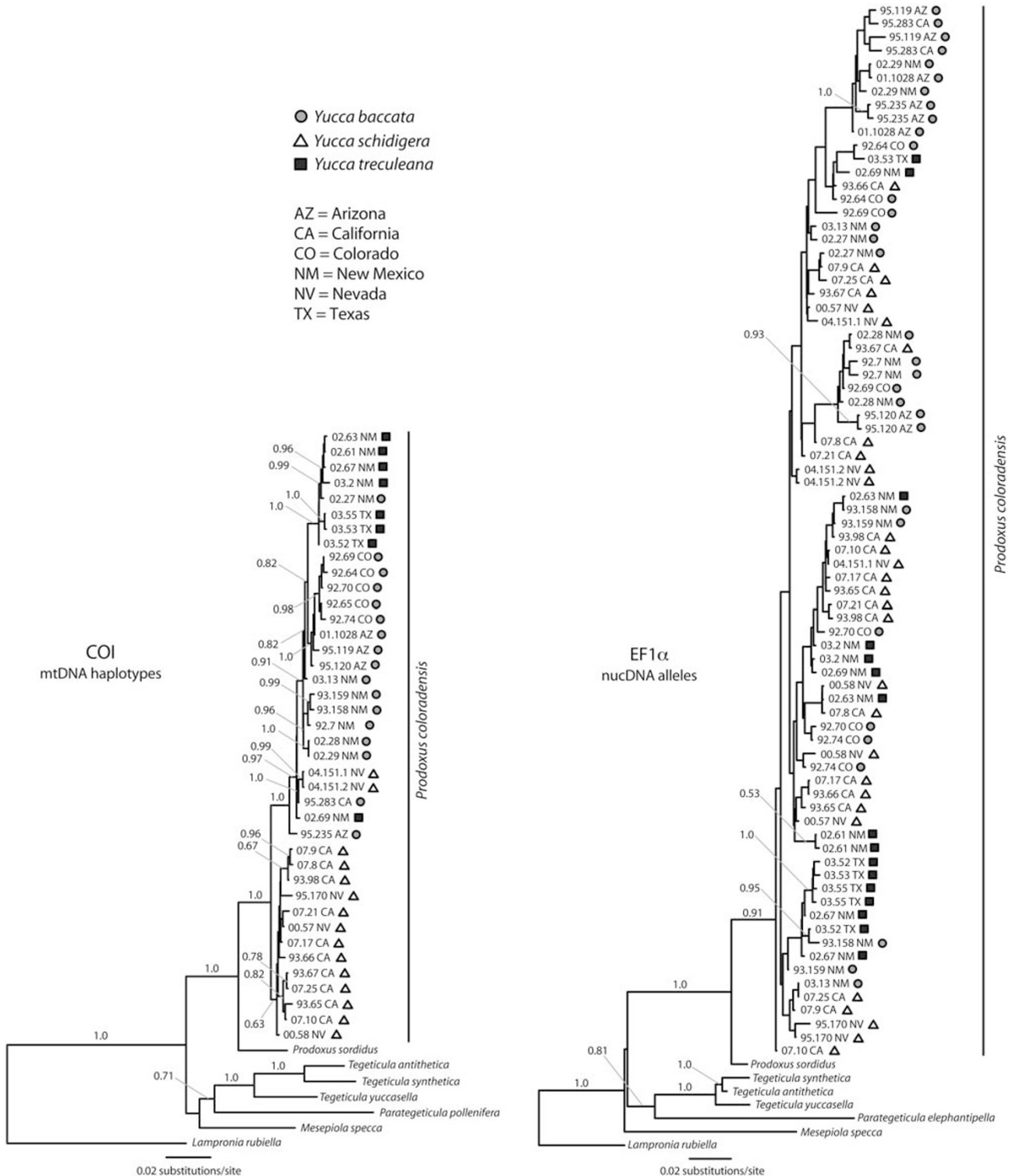
Figure 3 Bayesian Markov chain Monte Carlo (MCMC) maximum clade credibility trees for cytochrome oxidase I (COI) and elongation factor 1 alpha (EF1α) from *P. coloradensis* on three species of *Yucca*. Tip node labels indicate accession, sampling location and host plant (cf. Figure 1). Posterior probabilities (>0.50) are shown next to the branches. These trees represent a single realization from the posterior distribution of genealogies for each locus. COI recovered many nodes with strong support (≥0.90), approximately corresponding to host plant use. In contrast, EF1α recovered a large number of equally probable arrangements with very low support (≤0.01), consistent with lower levels of host structure for nuclear DNA (nucDNA) versus mitochondrial DNA (mtDNA).

The AMOVAs for genetic differentiation among all three host populations showed substantial levels of structure within *P. coloradensis* (Table 1), with strong support for rejection of the null hypothesis of panmixia among host groups (COI: $\Phi_{ST} = 0.625$, $P < 0.001$; EF1α: $\Phi_{ST} = 0.261$, $P < 0.001$). Pairwise AMOVAs between host

Table 1 Analysis of molecular variance for COI and EF1α from *Prodoxus coloradensis* on three species of *Yucca*

Source of variation	Percentage of variation	Fixation index (Φ_{ST})	P-value
COI			
Among host plants	62.53	0.625	<0.001
Within host plants	37.47		
EF1α			
Among host plants	26.08	0.261	<0.001
Within host plants	73.92		

Abbreviations: COI, cytochrome oxidase I; EF1α, elongation factor 1 alpha.
Both loci showed evidence for significant departures from panmixia among moths on different host species, although less variation was partitioned among hosts for nuclear DNA versus mitochondrial DNA. Tests for differentiation among each pair of host populations were all significant at a threshold of $P < 0.05$ (results not shown), indicating that the global pattern of differentiation among host plants was not driven by structure between a subset of host populations.

Table 2 Mantel and partial Mantel tests for COI and EF1α from *Prodoxus coloradensis* on three species of *Yucca*, examining the relationship between genetic distance, geographic distance and host use

Comparison	r	P-value
COI		
Genetic vs geographic	0.650	<0.001
Genetic vs host	0.623	<0.001
Genetic vs geographic (control host)	0.445	<0.001
Genetic vs host (control geographic)	0.386	<0.001
EF1α		
Genetic vs geographic	0.174	<0.001
Genetic vs host	0.155	<0.001
Genetic vs geographic (control host)	0.101	<0.001
Genetic vs host (control geographic)	0.064	0.021

Abbreviations: COI, cytochrome oxidase I; EF1α, elongation factor 1 alpha.
Both loci showed evidence for a significant correlation between genetic distance and host use, even after controlling for the relationship between genetic and geographical distances, with evidence of greater host structuring for mitochondrial DNA versus nuclear DNA. Partial Mantel tests for correlations between genetic distance and host use (controlling for geographic distance) among each pair of host populations were all significant at a threshold of $P < 0.05$ (results not shown), indicating that the global pattern of differentiation among host plants was not driven by a subset of host populations.

plants recovered significant differentiation for each comparison at a threshold of $P < 0.05$ (results not shown), indicating that significant results for the global test were not driven by genetic structure between a subset of the three host groups. Mantel tests showed a similar pattern, with significant correlations between genetic distance and both geographical distance and host use (Table 2, Figure 4). Partial Mantel tests for correlation between genetic distance and host use (controlling for geographical distance) were also significant (COI: $r = 0.386$, $P < 0.001$; EF1α: $r = 0.064$, $P = 0.021$). Likewise, partial Mantel tests between each pairwise comparison of host groups were significant (results not shown), indicating that the results of the global test between all three host populations was not an artifact of host structure between any two host plants.

Likelihood-ratio tests for nested models in IMa are shown in Supplementary Table S2. The least complex model that could not be rejected for *Y. baccata* versus *Y. schidigera* or *Y. baccata* versus *Y. treculeana* was $\theta_1 = \theta_2 = \theta_A$, $m_1 = m_2$ (that is, equal effective population sizes with equal migration rates). For these two pairwise comparisons, we were able to reject the null hypothesis that $m_1 = 0$ and/or $m_2 = 0$ in almost all cases. The least complex model that could not be rejected for *Y. schidigera* versus *Y. treculeana* was $\theta_1 = \theta_2 = \theta_A$, $m_1 = m_2 = 0$ (that is, equal effective population sizes with no migration). For this pairwise comparison, we were unable to reject any null models in which $m_1 = 0$ and/or $m_2 = 0$. To evaluate estimates for *m* between *Y. schidigera* versus *Y. treculeana* with respect to the same model used for other comparisons between host groups, subsequent IM analyses for all three pairwise comparisons were conducted using $\theta_1 = \theta_2 = \theta_A$, $m_1 = m_2$.

Marginal posterior distributions for estimates of demographic parameters from IM are shown in Figure 5 and Table 3, representing results from the longest replicate run for each pairwise comparison. Posterior distributions for T_{div} had sharp peaks at <1.0 Myr (*ca.* 0.5–0.9 Myr) but exhibited long upper tails that failed to reach zero within the limits of the prior distribution (15.6 Myr). Convergent results from replicate runs with large ESS values (>300 minimum, typically >10 000)

and no evidence of trends in run-time plots indicated that failure to achieve complete posterior distributions for T_{div} was not a consequence of poor mixing or inadequate run times. Likewise, increasing the upper bounds on *t* did not recover any evidence for decreasing probabilities at larger values, suggesting that the data were compatible with a wide range of divergence times at low but finite probabilities. The mode of N_e was consistently large for all comparisons among host plant groups ($N_e = 952\,000$–$1\,334\,000$) with well-shaped posterior distributions. The mode of M_e was <1.0 for all pairwise comparisons ($M_e = 0.260$–0.963), although posterior distributions for this parameter encompassed a range of values exceeding $M_e > 1.0$. Locus-specific estimates for TMRCA and the mean number of simulated migration events are shown in Supplementary Figure S3 and Table S3. EF1α exhibited longer coalescence times compared with COI (4.0 Myr versus 1.0–2.0 Myr, respectively). Likewise, there was some evidence for higher rates of gene flow in EF1α versus COI (Supplementary Table S3), although there was substantial overlap in the posterior distributions for this parameter (Supplementary Figure S3).

Discussion

Discriminating between the recent divergence of species and contemporary gene flow among populations is a principal difficulty for understanding the evolutionary processes that drive the diversification of phytophagous insects on different host plants. At deeper time scales, fixation of haplotypes or alleles leads to the coalescence of gene copies within species, permitting the application of simple phylogenetic models based on bifurcating lineages (Maddison, 1997). At more recent time scales, determining the balance between competing evolutionary forces (that is, mutation, selection, genetic drift and
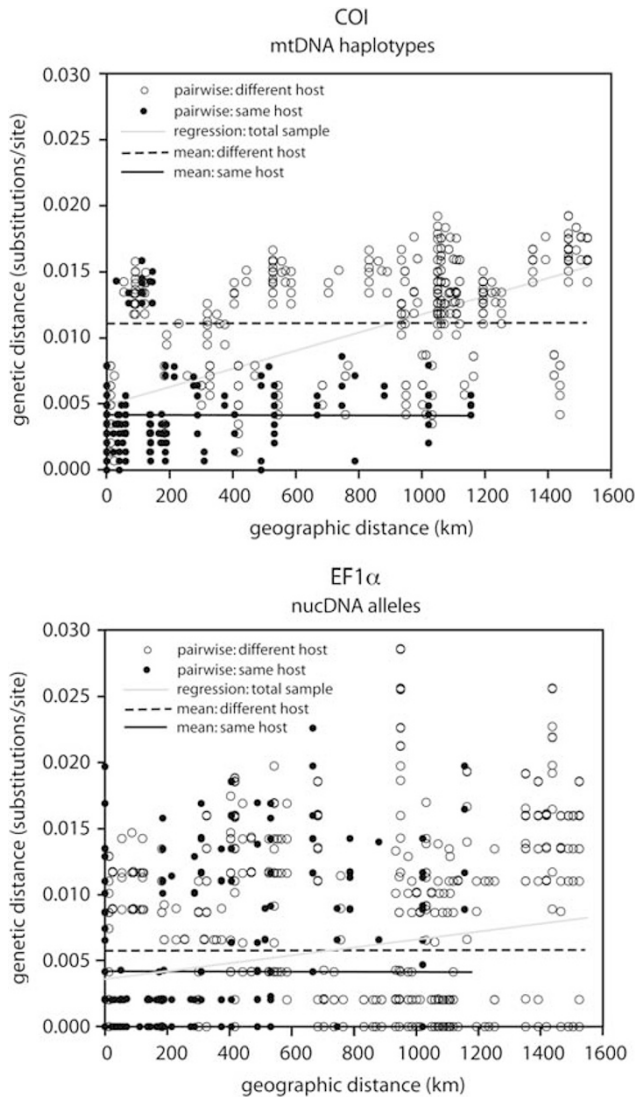
COI
mtDNA haplotypes

EF1α
nucDNA alleles

Figure 4 Scatter plots illustrating the relationship between genetic distance and geographical distance for cytochrome oxidase I (COI) and elongation factor 1 alpha (EF1α) from *P. coloradensis* on three species of *Yucca*. There is an overall pattern of isolation by distance for both loci. However, the correlation between genetic distance and host use remains significant after controlling for geographical distance (see Table 2 for correlation coefficients and *P*-values).

gene flow) remains a challenging proposition, despite advances in the application of coalescent theory to population genetics (reviewed in Rosenberg and Nordborg, 2002; Wakeley, 2008). Moreover, the recognition of random coalescence among different loci as a consequence of genetic drift has led to the increased emphasis on multilocus models for unraveling the historical demography of species and populations (Hey and Machado, 2003).

These issues are compounded when one considers the continuum of divergence from host races with appreciable gene flow to distinct, reproductively isolated taxa (Drès and Mallet, 2002; Mallet, 2008; Peccoud *et al.*, 2009). This scenario of recent divergence at the boundary between species and populations is well represented by the results from *P. coloradensis*. Although we did not find evidence for monophyletic host plant clades, the

posterior distribution of COI gene trees indicated that individuals from the same host species were more likely to cluster together than individuals from different hosts (Figure 3). In contrast, EF1α provided little support for clades within *P. coloradensis*, yielding a large number of equally probable topologies with no evidence for host structuring. Likewise, statistical parsimony networks for COI and EF1α showed that none of the sampled mtDNA haplotypes were shared among hosts, whereas a single widespread nucDNA allele was distributed across all the three hosts (Figure 2).

The absence of reciprocal monophyly among recently diverged species is unsurprising, as the probability of fixation is proportional to the strength of genetic drift (inversely related to N_e) and the time since divergence (Rosenberg and Nordborg, 2002; Wakeley, 2005). Under a Fisher–Wright model at equilibrium, populations will diverge due to the combined effects of mutation and genetic drift when the effective number of migrants (that is, $M_e = 2 N_e m$) is < 1.0 (Wright, 1931; see Hey *et al.*, 2004) for an empirical example using IM). Estimates of large N_e and recent T_{div} from coalescent simulations (Table 3, Figure 5) supported the hypothesis of retained ancestral polymorphism. Furthermore, TMRCA values for both COI and EF1α extended past the estimated time of divergence (*ca.* 0.5–0.9 Myr), indicating that ancestral mtDNA haplotypes and nucDNA alleles have persisted in descendant populations (Supplementary Figure S3). In particular, TMRCA values showed that coalescence times for EF1α (*ca.* 4.0 Myr) were earlier than COI (*ca.* 1.0–2.0 Myr), consistent with larger N_e for diploid, biparentally nucDNA genes compared with the haploid, uniparentally inherited mtDNA genome (Birky, 2001; Ballard and Whitlock, 2004).

Likelihood-ratio tests rejected the null hypothesis of zero migration between the parapatric pairwise comparisons of *Y. baccata* and *Y. schidigera* or *Y. baccata* and *Y. treculeana*, but were unable to reject zero migration between the allopatric pair of *Y. schidigera* and *Y. treculeana* (Supplementary Table S2), indicating that low levels of gene flow may persist through sympatric contact zones (Figure 1). However, point estimates on the basis of the mode of M_e exhibited a clear pattern of M_e < 1.0 for all pairwise comparisons among host plants (Table 3), suggesting that limited migration is insufficient to prevent divergence. Although 90% HPDs for M_e included a range of values for which $M_e > 1.0$, we note that calculating the posterior density of m conditioned on the modal value of θ represents an extremely conservative interpretation of the data, yielding substantially wider confidence intervals than the alternative parameterization of the posterior density of θ conditioned on the modal value of m.

Although the general trends for coalescent estimates of N_e, T_{div} and M_e were similar across comparisons among host populations, *post hoc* interpretation of pairwise IM results can present formidable challenges (Hoelzel *et al.*, 2007; Lucas *et al.*, 2009; Pinho *et al.*, 2008; King and Roalson, 2009). When two populations have received immigrants from a third population not represented in a given pairwise analysis, estimates of N_e can be inflated; if both populations are exchanging migrants with this third population, estimates of M_e can be similarly biased (Won *et al.*, 2005). Likewise, as $M_e = 2N_e m$, any upward bias in the estimate of N_e would also inflate the estimates
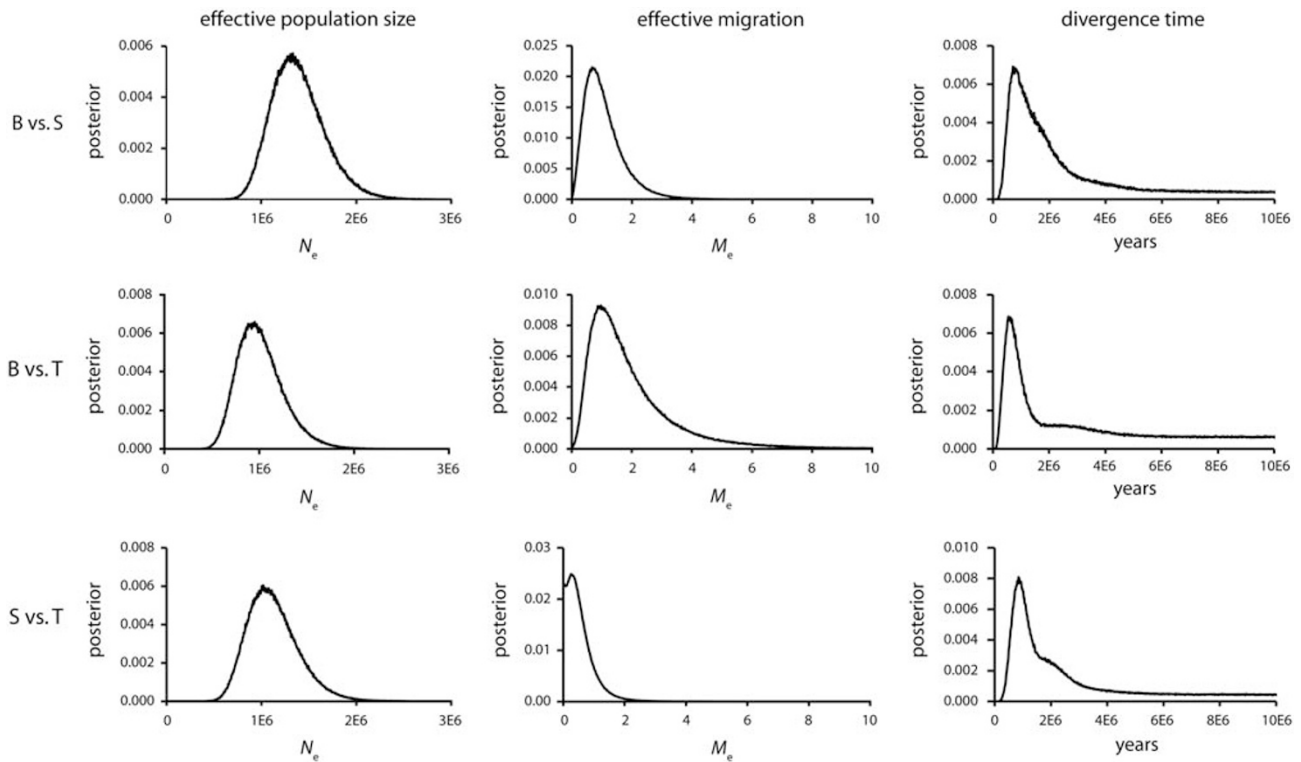
Figure 5 Marginal posterior distributions for effective population size (N_e), divergence time scaled to years (T_{div}) and effective migration rates (M_e) for *P. coloradensis* on three host species of *Yucca* (B = *Y. baccata*, S = *Y. schidigera*, T = *Y. treculeana*). See Table 3 and the Methods section for details on calculating point estimates and 90% posterior densities of demographic values from coalescent parameters.

Table 3 Marginal posterior distributions of multilocus coalescent estimates for effective population size (N_e), divergence time scaled to years (T_{div}) and effective migration rates (M_e) for pairwise comparisons of *Prodoxus coloradensis* on three host species of *Yucca*

Comparison	N_e	M_e	T_{div}
B vs S	1 334 000 (937 000–1 822 000)	0.668 (0.123–1.977)	710 000 (226 000–15 599 000)
B vs T	952 000(626 000–1 396 000)	0.963 (0.165–3.902)	554 000 (164 000–15 599 000)
S vs T	1 018 000 (692 000–1 528 000)	0.260 (0.010–1.093)	851 000 (273 000–15 599 000)

Abbreviations: B, *Y. baccata*; S, *Y. schidigera*; T, *Y. treculeana*.
Estimates have been converted to demographic values based on the mode and 90% highest posterior density (HPD) (in parentheses) for each coalescent parameter (cf. Figure 4 and Methods). The results were consistent with large N_e and recent T_{div}, although 90% HPDs for T_{div} approach the upper bound for divergence times (*ca.* 15.6 Myr) due to incomplete posterior distributions. Modal values for M_e <1.0 suggest that migration is insufficient to prevent divergence among moths on different host species, whereas 90% HPDs for this parameter encompass a range of values that include higher rates of migration.

of M_e. In part, the large values for N_e and wide posterior distributions for M_e from pairwise analyses of *P. coloradensis* may reflect these patterns (Figure 5), suggesting that our estimates are likely to overstate the true levels of effective migration among host plants.

The conclusion of low effective migration rates based on multilocus coalescent simulations was also supported by significant partitioning of genetic structure among host plants on the basis of AMOVA (Table 1). It is noteworthy that the levels of mtDNA differentiation were more than twice as large as nucDNA differentiation (COI: $\Phi_{ST} = 0.625$; EF1α: $\Phi_{ST} = 0.261$). This result was consistent with the greater phylogenetic resolution of host plant use based on COI versus EF1α (Figure 3). Likewise, the partial Mantel test for correlation between COI genetic distance and host use (controlling for geographical distance) was highly significant ($r = 0.386$, $P < 0.001$), whereas the analogous correlation between EF1α genetic distance and host use was much weaker ($r = 0.064$, $P = 0.021$).

As mtDNA is maternally inherited, these patterns may represent the signature of sex-biased dispersal. *Prodoxus* larvae typically pupate in the host tissue at the original site of oviposition. On pupal eclosion, we might expect that female moths remain on their natal host or within the same patch of host individuals, where oviposition sites are assured, whereas males disperse farther in search of unrelated mates (Gandon, 1999). This behavior parallels sex-biased dispersal seen in other phytophagous insects (Kuussaari *et al.*, 1996; Albrectsen and

Nachman, 2001; Petit et al., 2001; Caudill, 2003; Bowler and Benton, 2005; but see Mallet, 1986; Lawrence, 1988; Hill et al., 1996) and should lead to higher rates of nucDNA versus mtDNA gene flow (Salle et al., 2007).

The distribution of locus-specific migration events in the IMa coalescent simulations, which explicitly account for mutation rates and organellar versus autosomal inheritance, indicated that the peak values for inferred migration were similar for both COI and EF1α (Supplementary Figure S3). However, locus-specific posterior distributions for EF1α tended to be wider, encompassing higher rates of migration at low but finite probabilities, with the mean number of nucDNA migration events approximately two to three times larger compared with mtDNA (Supplementary Table S3). These posterior distributions are distinct from their respective prior distributions (see Methods), showing that differences in locus-specific gene flow are not simply an artifact resulting from the parameterization of nucDNA versus mtDNA inheritance models. Nonetheless, although female philopatry could represent a critical stage in the evolution of reproductive isolation associated with host shifts, the determination of sex-biased dispersal remains a difficult task (Prugnolle and de Meeus, 2002; Hedrick, 2005). Further understanding of the potential relationship between female philopatry and host-driven speciation is likely to require more detailed ecological studies of the fine-scale fitness consequences of host selection and oviposition in sympatric populations.

Given the parapatric distribution of Y baccata with respect to both Y. schidigera and Y. treculeana (Figure 1), we also asked whether high levels of genetic differentiation among moths on different hosts could be explained under a simple isolation-by-distance model, as distances between host populations can vary from immediate sympatry to hundreds of kilometers. Mantel tests indicated significant correlations between genetic distance versus both geographical distance and host use (Table 2, Figure 4). However, the relationship between genetic distance and host use remained significant after controlling for geographical distance (Table 2), suggesting that differentiation between host plants is not simply an artifact of spatial isolation. Conversely, geographical distance remains significant after controlling for host use. Noting the sister position of moths on Y. schidigera to Y. baccata and Y. treculeana in the COI phylogeny, as well as the general trend from basally branching western lineages to derived eastern lineages (Figure 3), this residual signal of spatial structure could reflect the historical signature of range expansion from west to east, leading to colonization and specialization on different hosts (for example, Segraves and Pellmyr, 2004). Alternatively, despite recent divergence, host populations may have reached equilibrium with limited gene flow between Y. baccata and geographically proximate flanking populations of Y. schidigera and Y. treculeana, with no direct gene flow between the latter, as suggested by the failure to reject the null model of $\theta_1 = \theta_2 = \theta_A$, $m_1 = m_2 = 0$ for these two widely separated host groups (Supplementary Table S2).

Despite the opportunity for gene flow in narrow zones of sympatry, the results of this study support the hypothesis that P. coloradensis comprises a complex of incipient host-specialized species. These results are consistent with high levels of differentiation and low rates of effective migration among host plants, although with considerable variation in the width of posterior densities for M_e, accompanied by a strong signal of recent divergence (ca. 0.5–0.9 Myr) and the retention of shared ancestral polymorphism. Although the precise mechanisms remain unclear, there is also evidence for higher rates of nucDNA versus mtDNA gene flow, suggesting that female philopatry might contribute to divergence among host populations. Furthermore, many of these patterns are analogous to the results from previous study on pollinating yucca moths affiliated with Y. baccata and Y. schidigera (Leebens-Mack and Pellmyr, 1998), indicating parallel trajectories of host-specific divergence in both pollinating and non-pollinating species.

Considering the numerous closely related species of Prodoxus found on similar host plants with overlapping biogeographical ranges (Pellmyr et al., 2006), future studies will benefit from a better understanding of evolutionary relationships at the interface between species and populations, not just within P. coloradensis but across the entire genus. Deep coalescence of mtDNA haplotypes and nucDNA alleles suggests that shared ancestral polymorphism may persist across other recently diverged lineages of Prodoxus, complicating efforts to distinguish incomplete lineage sorting from the ongoing gene flow. However, the development of highly polymorphic nucDNA microsatellite markers in the closely related genus Tegeticula (Drummond et al., 2009) provides new molecular methods for dissecting the fine-scale population dynamics of prodoxid yucca moths distributed among sympatric or parapatric host plants. Likewise, coordinated estimates of the phylogenetic relationships among Yucca (Pellmyr et al., 2007; Smith et al., 2008b) have created an expanded framework for evaluating the biogeographical history and diversification of these taxa with respect to host plant specialization. Coupled with these tools, coalescent models for multilocus phylogenetic inference (Edwards et al., 2007; Liu et al., 2008; Kubatko et al., 2009) and extensions of the basic isolation-with-migration coalescent model to multiple populations offer novel opportunities to examine the extent to which host specificity influences the divergence of phytophagous insects.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

References

Albrectsen B, Nachman G (2001). Female-biased density-dependent dispersal of a tephritid fly in a fragmented

194

habitat and its implications for population regulation. *Oikos* **94**: 263–272.

Althoff DM (2008). A test of host-associated differentiation across the 'parasite continuum' in the tri-trophic interaction among yuccas, bogus yucca moths, and parasitoids. *Mol Ecol* **17**: 3917–3927.

Althoff DM, Groman JD, Segraves KA, Pellmyr O (2001). Phylogeographic structure in the bogus yucca moth *Prodoxus quinquepunctellus* (Prodoxidae): comparisons with coexisting pollinator yucca moths. *Mol Phylogenet Evol* **21**: 117–127.

Althoff DM, Pellmyr O (2002). Examining genetic structure in a bogus yucca moth: a sequential approach to phylogeography. *Evolution* **56**: 1632–1643.

Althoff DM, Segraves KA, Leebens-Mack J, Pellmyr O (2006). Patterns of speciation in the yucca moths: parallel species radiations within the *Tegeticula yuccasella* species complex. *Syst Biol* **55**: 398–410.

Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. *Mol Ecol* **13**: 729–744.

Becerra JX, Venable DL (1999). Macroevolution of insect-plant associations: the relevance of host biogeography to host affiliation. *Proc Natl Acad Sci USA* **96**: 12626–12631.

Birky CW (2001). The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet* **35**: 125–148.

Bonnet E, Van de Peer Y (2002). zt: a software tool for simple and partial Mantel tests. *J Stat Softw* **7**: 1–12.

Bowler DE, Benton TG (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biol Rev* **80**: 205–225.

Caudill CC (2003). Measuring dispersal in a metapopulation using stable isotope enrichment: high rates of sex-biased dispersal between patches revealed in a mayfly metapopulation. *Oikos* **101**: 624–630.

Clement M, Posada D, Crandall K (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1660.

Coyne JA, Orr HA (2004). *Speciation*. Sinauer Associates: Sunderland, MA.

Davis DR (1967). A revision of the moths of the sub-family Prodoxinae (Lepidoptera: Incurviidae). *Bull US Nat Hist Mus* **255**: 1–170.

Degnan JH, Rosenberg NA (2006). Discordance of species trees with their most likely gene trees. *PLoS Genet* **2**: 762–768.

Diehl SR, Bush GL (1984). An evolutionary and applied perspective of insect biotypes. *Ann Rev Entomol* **29**: 471–504.

Dobzhansky T (1951). *Genetics and the Origin of Species*. Columbia University Press: New York.

Drès M, Mallet J (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Phil Trans R Soc Lond B* **357**: 471–492.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol* **4**: 699–710.

Drummond AJ, Nicolls GK, Rodrigo AG, Solomon W (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**: 1307–1320.

Drummond AJ, Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.

Drummond CS, Smith CI, Pellmyr O (2009). Species identification and sibship assignment of sympatric larvae in the yucca moths *Tegeticula synthetica* and *Tegeticula antithetica* (Lepidoptera: Prodoxidae). *Mol Ecol Resour* **9**: 1369–1372.

Edwards SV, Liu L, Pearl DK (2007). High-resolution species trees without concatenation. *Proc Natl Acad Sci USA* **104**: 5036–5041.

Ehrlich PR, Raven PH (1964). Butterflies and plants: a study of coevolution. *Evolution* **18**: 586–608.

Ersts PJ (2009). *Geographic Distance Matrix Generator 1.2.3*. Available at http://biodiversityinformatics.amnh.org/open_source/gdmg.

Excoffier L, Laval G, Schneider S (2005). Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from matrix distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Farrell BD (1998). Inordinate fondness explained: why are there so many beetles. *Science* **281**: 555–559.

Ferrari J, Godfray HCJ, Faulconbridge AS, Prior K, Via S (2006). Population differentiation and genetic variation in host choice among pea aphids from eight host plant genera. *Evolution* **60**: 1574–1584.

Gandon S (1999). Kin competition, the cost of inbreeding and the evolution of dispersal. *J Theor Biol* **200**: 345–364.

Godsoe W, Yoder JB, Smith CI, Pellmyr O (2008). Coevolution and divergence in the Joshua tree/yucca moth mutualism. *Am Nat* **171**: 816–823.

Groman JD, Pellmyr O (2000). Rapid evolution and specialization following host colonization in a yucca moth. *J Evol Biol* **13**: 223–236.

Hedrick PW (2005). A standardized genetic differentiation measure. *Evolution* **59**: 1633–1638.

Hess WJ, Robbins RL (2002). *Yucca*. In: Flora of North America Editorial Committee (eds). *Flora of North America North of Mexico*, Vol. 26. New York and Oxford. pp 413–440.

Hey J (2006). Recent advances in assessing gene flow between diverging species and populations. *Curr Opin Genet Dev* **16**: 592–596.

Hey J, Machado C (2003). The study of structured populations—new hope for a difficult and divided science. *Nat Rev Genet* **4**: 535–543.

Hey J, Nielsen R (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D persimilis*. *Genetics* **167**: 747–760.

Hey J, Nielsen R (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci USA* **104**: 2785–2790.

Hey J, Won Y-J, Sivasundar A, Nielsen R, Markert JA (2004). Using nuclear haplotypes with microsatellites to study gene flow between recently separated Cichlid species. *Mol Ecol* **13**: 909–919.

Hill JK, Thomas CD, Lewis OT (1996). Effects of habitat patch size and isolation on dispersal by *Hesperia comma* butterflies: implications for metapopulation structure. *J Anim Ecol* **65**: 725–735.

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, Black N (2007). Evolution of population structure in a highly social top predator, the killer whale. *Mol Biol Evol* **24**: 1407–1415.

King MG, Roalson EH (2009). Discordance between phylogenetics and coalescent-based divergence modelling: exploring phylogeographic patterns of speciation in the *Carex macrocephala* species complex. *Mol Ecol* **18**: 468–482.

Knowles LL, Carstens BC (2007). Delimiting species without monophyletic gene trees. *Syst Biol* **56**: 887–895.

Kubatko LS, Carstens BC, Knowles LL (2009). STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* **25**: 971–973.

Kumar S, Dudley J, Nei M, Tamura K (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinf* **9**: 299–306.

Kuussaari M, Nieminen M, Hanski I (1996). An experimental study of migration in the Glanville fritillary butterfly *Melitaea cinxia*. *J Anim Ecol* **65**: 791–801.

Lawrence WS (1988). Movement ecology of the red milkweed beetle in relation to population size and structure. *J Anim Ecol* **57**: 21–35.

Leebens-Mack J, Pellmyr O (2004). Patterns of genetic structure among populations of an oligophagous pollinating yucca moth (*Tegeticula yuccasella*). *J Hered* **95**: 127–135.

Leebens-Mack J, Pellmyr O (1998). Host specificity and the genetic structure of two yucca moth species in a yucca hybrid zone. *Evolution* **52**: 1376–1382.

Li N, Stephens M (2003). Modelling linkage disequilibrium and identifying recombination hotspots using snp data. *Genetics* **165**: 2213–2233.

Liu L, Pearl DK, Brumfield RT, Edwards SV (2008). Estimating species trees using multiple-allele DNA sequence data. *Evolution* **62**: 2080–2091.

Librado P, Rozas J (2009). DNaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Lucas L, Gompert Z, Ott J, Nice C (2009). Geographic and genetic isolation in spring-associated Eurycea salamanders endemic to the Edwards Plateau region of Texas. *Conserv Genet* **10**: 1309–1319.

Maddison WP (1997). Gene trees in species trees. *Syst Biol* **46**: 523–536.

Mallet J (1986). Dispersal and gene flow in a butterfly with home range behavior: *Heliconius erato* (Lepidoptera: Nymphalidae). *Oecologia* **68**: 210–217.

Mallet J (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Phil Trans R Soc Lond B* **363**: 2971–2986.

Mantel N (1967). The detection of disease clustering as a generalized regression approach. *Cancer Res* **27**: 209–220.

Marshall DC, Simon C, Buckley TR (2006). Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. *Syst Biol* **55**: 993–1003.

Matuda E, Lujan IP (1980). *Las Plantas Mexicanas del Genero Yucca*. Serie Fernando de Alva Ixtlilxochitl, Collecion miscelanea estado de México, Toluca, México.

Miller NJ, Birley AJ, Overall ADJ, Tatchell GM (2003). Population genetic structure of the lettuce root aphid, *Pemphigus bursarius* (L.), in relation to geographic distance, gene flow and host plant. *Heredity* **91**: 217–223.

Mitter C, Farrell B, Wiegmann B (1988). The phylogenetic study of adaptive zones: has phytogeography promoted plant diversification? *Am Nat* **132**: 107128.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.

Nason JD, Heard SB, Williams FR (2002). Host-associated genetic differentiation in the goldenrod elliptical-gall moth, *Gnorimoschema gallaesolidaginis* (Lepidoptera: Gelechiidae). *Evolution* **56**: 1475–1488.

Nylander J: MrModeltest 2.3. Available at http://www.abc.se/~nylander/.

Peccoud J, Ollivier A, Plantegenest M, Simon J-C (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proc Natl Acad Sci USA* **106**: 7495–7500.

Pellmyr O (2003). Yuccas, yucca moths, and coevolution: a review. *Ann Mo Bot Gard* **90**: 35–55.

Pellmyr O, Balcázar-Lara M, Althoff DM, Segraves KA, Leebens-Mack J (2006). Phylogeny and life history evolution of *Prodoxus* yucca moths (Lepidoptera: Prodoxidae). *Syst Entomol* **31**: 1–20.

Pellmyr O, Krenn HW (2002). Origin of a complex key innovation in an obligate insect–plant mutualism. *Proc Natl Acad Sci USA* **99**: 5498–5502.

Pellmyr O, Leebens-Mack J (1999). Forty million years of mutualism: evidence for Eocene origins of the yucca–yucca moth association. *Proc Natl Acad Sci USA* **96**: 9178–9183.

Pellmyr O, Segraves KA (2003). Pollinator divergence within an obligate mutualism: two yucca moth species (Lepidoptera; Prodoxidae: *Tegeticula*) on the Joshua tree (*Yucca brevifolia*; Agavaceae). *Ann Entomol Soc Am* **96**: 716–722.

Pellmyr O, Segraves KA, Althoff DM, Balcázar-Lara M, Leebens-Mack J (2007). The phylogeny of yuccas. *Mol Phylogenet Evol* **43**: 493–501.

Petit S, Moilanen A, Hanski I, Baguette, M (2001). Metapopulation dynamics of the Bog Fritillary butterfly: movements between habitat patches. *Oikos* **92**: 491–500.

Pinho C, Harris DJ, Ferrand N (2008). Non-equilibrium estimates of gene flow inferred from nuclear genealogies suggest that Iberian and North African wall lizards (*Podarcis* spp.) are an assemblage of incipient species. *BMC Evol Biol* **8**: 63.

Posada D, Buckley TR (2004). Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. *Syst Biol* **53**: 793–808.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Prugnolle F, de Meeus T (2002). Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**: 161–165.

Rambaut A, Drummond AJ (2008). Tracer 1.4.1. Available at http://tree.bio.ed.ac.uk/software/tracer/.

Riley CV (1880). The true and bogus yucca moth, with remarks on the pollination of *Yucca*. *Am Entomol* **3**: 141–145.

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

Rosenberg NA (2002). The probability of topological concordance of gene trees and species trees. *Theor Popul Biol* **61**: 225–247.

Rosenberg NA, Nordborg M (2002). Genealogical trees, coalescent theory, and the analysis of genetic polymorphisms. *Nat Rev Genet* **3**: 380–390.

Rousset F (1997). Genetic differentiation and estimation of gene flow from *F* statistics under isolation by distance. *Genetics* **145**: 1219–1228.

Salle A, Arthofer W, Lieutier F, Stauffer C, Kerdelhue C (2007). Phylogeography of a host-specific insect: genetic structure of *Ipstypographus* in Europe does not reflect past fragmentation of its host. *Biol J Linn Soc* **90**: 239–246.

Segraves KA, Pellmyr O (2004). Testing the out-of-Florida hypothesis on the origin of cheating in the yucca–yucca moth mutualism. *Evolution* **58**: 2266–2279.

Segraves KA, Althoff DM, Pellmyr O (2005). Limiting cheaters in mutualism: evidence from hybridization between mutualist and cheater yucca moths. *Proc R Soc Lond B* **272**: 2195–2202.

Smith CI, Godsoe WKW, Tank S, Yoder JB, Pellmyr O (2008a). Distinguishing coevolution from covicariance in an obligate pollination mutualism: asynchronous divergence in Joshua tree and its pollinators. *Evolution* **62**: 2676–2687.

Smith CI, Pellmyr O, Althoff DM, Balcázar-Lara M, Leebens-Mack J, Segraves KA (2008b). Pattern and timing of diversification in *Yucca* (Agavaceae): specialized pollination does not escalate rates of diversification. *Proc R Soc Lond B* **275**: 249–258.

Smouse PE, Long JC, Sokal RR (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* **35**: 727–732.

Stephens M, Donnelly P (2003). A comparison of Bayesian methods for haplotype reconstruction. *Am J Hum Genet* **73**: 1162–1169.

Stephens M, Smith N, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**: 979–989.

Svensson GP, Althoff DM, Pellmyr O (2005). Replicated host-race formation in bogus yucca moths: genetic and ecological divergence of *Prodoxus quinquepunctellus* on yucca hosts. *Evol Ecol Res* **7**: 1139–1151.

196

Tajima F (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**: 437–460.

Tamura K, Nei M, Kumar S (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci USA* **101**: 11030–11035.

Wakeley J (2005). The limits of theoretical population genetics. *Genetics* **169**: 1–7.

Wakeley J (2008). *Coalescent Theory: an Introduction*. Roberts & Company: Greenwood, Colorado.

Won YJ, Sivasundar A, Wang Y, Hey J (2005). On the origin of Lake Malawi cichlid species: a population genetic analysis of divergence. *Proc Natl Acad Sci USA* **102**: 6581–6586.

Wright S (1931). Evolution in mendelian populations. *Genetics* **16**: 97–159.

Yang Z, Rannala B (2005). Branch-length prior influences Bayesian posterior probability of phylogeny. *Syst Biol* **54**: 455–470.

Zwickl DJ (2006). Garli 0.96. Software available from http://garli.nescent.org.