Supporting Information

Moore and Donoghue 10.1073/pnas.0807230106

SI Text

Evaluating the impact of morphological/biogeographic events on rates of diversification in Adoxaceae and *Lupinus* basically entails estimating (and subsequently combining) the marginal posterior densities of parameters for the timing of each event $\Pr(t_k|X)$ and the corresponding background rate of diversification $\Pr(\lambda_{(k)}|X_{(k)})$. Operationally, this is achieved via an analytical pipeline with 5 main steps (Fig. 1). Step I involves preliminary analyses to select appropriate models for nucleotide-substitution and substitution-rate variation. Step II involves estimating the joint posterior probability density of the phylogeny and divergence times from the entire data set of nucleotide sequences. In step III, the history of each event is estimated from the resulting posterior probability distribution of trees: this involves estimating ancestral character states of fruit morphology in Adoxaceae and inferring ancestral geographic ranges and episodes of dispersal in *Lupinus*. In step IV, the training data partition is defined by removing the subset of species associated with the event and then subjecting the remaining data set to a second round of phylogeny/divergence-time estimation. Finally, we draw from the marginal posterior densities of t_k and $\lambda_{(k)}$ (estimated in steps III and IV, respectively) to generate the cross-validation predictive diversity density (using Eq. **3**), which permits us to assess whether the realized species diversity is significantly greater or less than predicted (using Eq. **4**). The series of analyses described below have been detailed in two spreadsheets, Datasets S1 (Adoxaceae.xls) and S2 (Lupinus.xls).

I. Preliminary Analyses. Sequences for Adoxaceae and *Lupinus* were aligned using MUSCLE (1) and then adjusted manually as necessary; the resulting sequence alignments are provided as NEXUS formatted text files. Various properties of these data sets are summarized in Table S1. Model selection for each gene region was based on ΔAIC scores for the set of substitution models evaluated by Modeltest v.3.7 (2). Models selected for each gene region are listed in Table S1. The degree of substitution rate variation was investigated by means of likelihood-ratio tests using PAUP* (3): we estimated the likelihood for each partition under the selected model both with and without enforcing substitution rate constancy and then compared the resulting test statistic to a χ^2 distribution with $(N - 2)$ degrees of freedom, where N is the number of taxa. These molecular-clock tests detected high levels of substitution rate variation in all data partitions (Table S1).

II. Divergence-Time Estimation. ***Relaxed Clock Model.*** Owing to the high level of substitution rate variation, divergence times were estimated under a relaxed molecular clock. We chose to model substitution-rate variation using the uncorrelated lognormal (UCLN) model (4), as this represents a more general relaxed clock model that does not assume that substitution rates are autocorrelated along ancestor-descendant branches. Moreover, the implementation of the UCLN model in BEAST v.1.4.7 (5) is attractive as it enables uncertainty in the age of calibrations to be represented as informative prior distributions rather than committing to point estimates or inappropriately vague uniform priors.

Fossil Calibrations. Following Moore and Donoghue (6), we constrained the crown group of Adoxaceae to a lognormal distribution with an upper bound of 86 My and a lower bound of 45 My (Table S2), on the basis of the earliest macrofossil occurrence (see ref. 7) and palynological record (8). Within Adoxaceae, the crown group of *Viburnum* was constrained to a lognormal distribution with an upper bound of 85 My and a lower bound of 45 My [on the basis of the oldest unambiguous crown-group fossil (ref. 9)]. Following Hughes and Eastwood (10), the stem of *Lupinus* was constrained to a normal distribution with a mean of 18.5 My [on the basis of results of a more inclusive analysis of legumes that used *matK* sequence data and multiple fossil calibrations (ref. 11)].

Analyses. We approximated the posterior probability density of divergence times for each group under a mixed model (i.e., comprising the partition-specific substitution model selected for each gene region) and assumed that substitution rates evolved under the UCLN relaxed molecular clock model and that branching process conformed to a Yule prior (i.e., the prior distributions of topologies and divergence times were generated by a pure-birth stochastic branching process). For each group, we initiated 4 independent Markov chain Monte Carlo (MCMC) analyses from starting trees that satisfied the respective priors on divergence times described above (Table S2). Convergence of each chain to the target distribution was inferred by plotting time series of the marginal probabilities of sampled parameter values using Tracer v.1.4 (5). After achieving convergence, each chain was sampled every 10^3 steps until 10^5 samples were obtained, and the adequacy of sampling intensity was assessed by means of the estimated sample size (ESS) diagnostic.

Results. The UCLN analyses corroborated findings of the likelihood-ratio test of substitution rate constancy (Table S1): the estimated coefficient of variation identified substantial levels of substitution rate variation in both groups (Adoxaceae $\sigma = 0.364$, 95% HPD, [0.271, 0.463]; *Lupinus* $\sigma = 0.429$, 95% HPD, [0.264, 0.614]). Substitution rates across ancestor-descendant lineages were effectively uncorrelated in both groups (Table S1: Adoxaceae covariance = 0.0299, 95% HPD, [−0.139, 0.201]; *Lupinus* covariance = 0.005, 95% HPD, [−0.140, 0.0.153]), such that these data are ill suited to methods that assume substitution-rate autocorrelation. For each group, we combined the post-burn-in samples from the 4 independent MCMC analyses to provide a composite posterior of 6×10^4 dated trees. Large ESS values ($\gg 10^4$) for all parameters for both groups suggest adequate sampling intensity. The composite posterior probability density for Adoxaceae had a mean of −16887.136 (95% HPD: [−16915.623, −16861.762]), and that for *Lupinus* had a mean of −9091.605 (95% HPD: [−9138.836, −9040.371]). The posterior probability density of each group was summarized using TreeAnnotator v.1.4.7 (5): the resulting chronograms for Adoxaceae and *Lupinus* are depicted in Figs. 2 and 4, respectively.

IIIA. Inferring the Evolution of Fruit Morphology in Adoxaceae. ***Bayesian Inference of Morphological History.*** The evolution of fruit form in Adoxaceae was inferred by estimating the marginal posterior probability of ancestral states under a Bayesian model-averaging approach (12, 13). Like other methods based on stochastic mutational mapping (14–17), this is a fully hierarchical Bayesian approach for inferring ancestral character states that exploits MCMC sampling to integrate over uncertainty in the phylogeny and parameters of a stochastic model of trait evolution. However, rather than conditioning inferences on a specific model of trait evolution, the approach developed by Pagel and Meade (13) uses reversible-jump (RJ) MCMC to integrate over a set of plausible models. The evolution of binary traits (i.e., discrete characters that occur in 1 of 2 states, i and j) involves 5 continuous-time Markov models that differ with respect to the parameters describing the instantaneous rate of change between

states, q_{ij} and q_{ji}. These include three 1-rate parameter models: the first, *00*, constrains the rate of forward and backward change to be equal [i.e., $q_{ij} = q_{ji}$, which is equivalent to the *Mk*1 model of Lewis (ref. 18)]; the other two 1-rate models, *Z0* and *0Z*, constrain the rate of forward or backward change to be zero (i.e., $q_{ij} = 0$, $q_{ji} > 0$ and $q_{ij} > 0$, $q_{ji} = 0$, respectively). The 2-rate parameter models allow an asymmetry in the rate of forward and backward change: the *10* model assumes a higher rate of forward change (i.e., $q_{ij} > q_{ji}$), whereas the *01* model assumes an opposite rate bias (i.e., $q_{ij} < q_{ji}$). At stationarity, these 5 models will be visited by the RJMCMC in proportion to their posterior probabilities, effectively selecting among (or averaging over) the models of trait evolution while simultaneously accommodating uncertainty in the phylogeny and rate parameters of the models.

Analyses. Analyses were performed using the program BayesTraits v.1.0 (19). Posterior probabilities of rate coefficients and ancestral states for the fruit trait were estimated by integrating over the composite post-burn-in sample of chronograms for Adoxaceae. The posterior distribution of chronograms was thinned to 10^3 trees to reduce the autocorrelation of successive samples. Ancestral states for nodes with $Pr < 1.0$ were estimated using the most recent common ancestor approach of Pagel *et al.* (12). We used a uniform prior on models of trait evolution and an exponential prior on the rate coefficients. Because there is little information on which to specify the mean of the exponential prior, we chose to seed this parameter from a uniform hyperprior, which allows values of the prior to be estimated from the data. We used an empirical approach to specify the range of the hyperprior. Specifically, we used BayesMultistate (19) to estimate the rate coefficients on each of the trees under maximum likelihood. The maximum-likelihood estimates of the rate coefficients were used to inform the range of the hyperprior. To ensure adequate mixing, we ran a series of preliminary chains, iteratively adjusting the magnitude of the rate-coefficient proposals (i.e., the value of the "ratedev" parameter) until acceptance rates reached 20−40%. To assess convergence, we performed 4 independent RJMCMC analyses. Convergence of each chain to the target distribution was inferred by plotting time series of the marginal posterior probability of sampled parameter values using Tracer v.1.4 (5). After achieving convergence, each chain was sampled every 10^3 steps until 10^4 samples were obtained, and the adequacy of sampling intensity was assessed by means of the ESS diagnostic.

Results. The inferred history of fruit type in Adoxaceae is depicted in Fig. 2, and the corresponding marginal log likelihoods and posterior probabilities of parameter estimates are summarized in Table S3. Note that the posterior probabilities under various models suggest that the *00* (i.e., *Mk*1) model is inappropriate for the fruit trait.

Inferring the Biogeographic History of *Lupinus*. *Likelihood Inference of Biogeographic History.* To locate dispersal events in *Lupinus* we used a maximum-likelihood approach (24, 25) for inferring the evolution of geographic range. The objective of this method is to identify the biogeographic history that maximizes the likelihood of realizing the observed geographic distributions of extant and fossil taxa under the specified biogeographic model. The biogeographic model comprises 2 main components. The phylogenetic component specifies the tree topology, the vector of branches with estimated durations/divergence times, and the intrinsic probabilities of lineage dispersal between areas and extinction within an area. The paleogeographic component includes the set of areas that encompass the ranges of the study species and their ancestors, the set of pairwise connections between those areas, and the corresponding set of dispersal functions that specify the probability of migration across each connection through time.

Parameterization of the Biogeographic Model. Following Moore *et al.* (21), the biogeographic model for *Lupinus* comprises 5 areas (Fig. 4 and Fig. S1): eastern Asia (EAS), Europe (EUR), eastern North America (ENA), western North America (WNA), and South America (SAM). There are 5 corresponding connections between these areas: the Bering land bridge (BLB) links WNA to EAS, the North Atlantic land bridge (NALB) links ENA to EUR, the Mesoamerican connection links ENA and WNA to SAM, and two intracontinental connections link ENA to WNA within North America and EAS to EUR within Eurasia. Dispersal functions specifying the probability of migration across each connection are detailed by Moore and Donoghue (6) (Fig. S1).

Analyses. Biogeographic analyses of *Lupinus* were performed using the program AReA (22). The most likely dispersal scenarios at all internal nodes of the study phylogenies were estimated using the maximum-likelihood estimates for lineage dispersal and extinction rates and the branch-specific geographic-range transition probability matrices that were approximated via Monte Carlo simulation using 10^5 simulations/matrix (20–21).

Results. The maximum-likelihood estimate for *Lupinus* ($\ln L = -278.473$) was obtained under equal rates of dispersal and extinction (both parameters 0.1). The inferred biogeographic history indicates that *Lupinus* most likely originated in EUR and ENA and includes 3 episodes of dispersal: 1 from ENA into WNA, 1 from ENA into SAM, and 1 from WNA into SAM (Fig. 4).

IV. Estimating Background Rates of Diversification. *Delineating the training data sets.* Each study group experienced 2 events of interest: single-seeded fruit arose twice in Adoxaceae and biogeographic dispersal into South America occurred 2 times in *Lupinus*. For each event, we defined a corresponding training data set by excluding all species associated with that event (i.e., descendants of the node along which the event was inferred) from the original data matrix. We note that researchers accustomed to evaluating the overall correlation between character states and diversification rates might be tempted to define the training partition by excluding all of the species that exhibit the state of interest. This would result, for example, in specifying a training data set for both of the events in *Lupinus* that excludes all of the species that occur in South America. By contrast, we emphasize that the current method is intended to evaluate the impact of particular historical events on rates of diversification, and therefore views the two episodes of dispersal into South America as two distinct historical events, each requiring definition of a separate training data partition to estimate the relevant background diversification rate. Specifically, the training partition for the Andean dispersal event excludes the species stemming from that event, and that for the Amazonian event excludes the species arising from that event. Similarly, the two training partitions for the two independent origins of single-seeded fruits in the MRCA of *Sinadoxa* and *Viburnum* exclude the set of species associated with those two events. Phylogeny/divergence times from the resulting alignments were subsequently inferred to estimate the background diversification rate (in the section of the tree not associated with the event under consideration).

Phylogeny/divergence time estimation. We approximated the posterior probability density of divergence times for each of the 4 training data partitions as described previously for the entire sequence alignment. These analyses used partition-specific substitution models and assumed that substitution rates evolved under the UCLN relaxed molecular clock model and specified a stochastic Yule prior on divergence times. As before, 4 independent MCMC analyses were run for each training data set, and convergence, mixing, and sampling intensity of each chain were evaluated as described above. For each training data set, we combined the post-burn-in samples from the 4 independent analyses using LogCombiner v.1.4.7 (5) and then summarized the collective marginal probability density for the speciation-rate parameter using Tracer v.1.4 (5).

Fossil calibrations and constraints. Analyses of all 4 training partitions used the same set of calibration priors, except for the complement of *Viburnum* in Adoxaceae (i.e., Adoxoideae), as the two constrained nodes were not included in the reduced taxon sample. Accordingly, we specified a calibration prior on the root of Adoxoideae on the basis of the posterior density of divergence times inferred for this node from analysis of the entire data set. It is conceivable that effects associated with taxon sampling might cause estimates based on the reduced training data sets to differ from those based on the entire data sets. However, estimates derived from the reduced and entire data sets for Adoxaceae and *Lupinus* were virtually identical. To further explore the possibility of taxon-sampling artifacts, we performed a series of analyses of the training data partitions that imposed prior distributions on nodes on the basis of the corresponding posterior distributions inferred from analyses of the entire data set. Again, the topologies and divergence times from the constrained and unconstrained analyses were virtually identical.

Results. The mean and HPD of the marginal posterior densities for the speciation-rate parameter estimated from each of the training data partitions are as follows: $\text{Adoxaceae}_{(Viburnum)} = 0.169$, 95% HPD, [0.054, 0.303]; $\text{Adoxaceae}_{(Sinadoxa)} = 0.120$, 95% HPD, [0.073, 0.170]; $\text{Lupinus}_{(\text{Andean})} = 0.367$, 95% HPD, [0.079, 0.762]; $\text{Lupinus}_{(\text{Amazon})} = 0.681$, 95% HPD, [0.319, 1.119].

V. Estimating the Correlation Between Events and Rates of Diversification. ***Cross-validation predictive density estimation.*** We used our cross-validation predictive density approach to investigate the influence of each historical event on rates of diversification in Adoxaceae and *Lupinus*. For a given event, the predictive diversity distribution was generated by means of Monte Carlo simulation. This merely involves repeatedly sampling pairs of event times and diversification-rate estimates from their respective marginal posterior densities, calculating the expected diversity, and storing each result to a list. The resulting array of expected species diversities is then summarized as the predictive diversity density, from which we subsequently ascertained whether the realized species diversity was significantly higher or lower than predicted (using Eq. **4**).

Analyses. The cross-validation predictive density analyses for each event were performed using tRate. The predictive diversity distribution was estimated by Monte Carlo simulation using 10^7 samples from the marginal posterior densities of t_k and $\lambda_{(k)}$ estimated in steps II and IV, respectively.

Results. The origin of single-seeded fruits in Adoxaceae was inferred to be correlated with significantly elevated species diversity in *Viburnum* ($P = 0.005$), but is evidently associated with a significant decrease in diversification rate in *Sinadoxa* ($P = 0.999$; Fig. 3). Biogeographic dispersal of *Lupinus* into South America appears to be correlated with a significant increase in rates of diversification in the Andean lineage ($P \ll 0.001$), whereas the dispersal of he Amazonian lineage into South America appears apparently had little impact on rates of diversification ($P = 0.47$; Fig. 4).

Testing Fruit Type as a Key Innovation of Adoxaceae Using the Method of Ree (2005). ***Stochastic-mapping based key innovation text.*** For the sake of comparison, we also evaluated the correlation between single-seeded fruits and rates of diversification in Adoxaceae, using the key innovation test of Ree (23). This method works as follows: given a tree with estimated divergence times and observed states of a binary trait across its tips (with states i and j), a character history is generated by means of stochastic mapping. This procedure essentially uses the *Mk*1 model of character evolution to paint segments of the study phylogeny according to the inferred state. The diversification rate of lineages possessing either state is calculated as the number of branching events that occurred while in a given state divided by the time spent in that state: e.g., $\lambda_i = n_i/t_i$. The difference between the two state-specific rates is given by $\Delta = (\lambda_i - \lambda_j)$. To accommodate for uncertainty in the estimate of character evolution, this process is repeated many times, and the average of the resulting Δ values provides the test statistic, $\bar{\Delta}$. The value of the test statistic is compared to a null distribution that is generated by Monte Carlo simulation as follows: the topology of the study tree is held constant, while divergence times are sampled from a stochastic Yule branching process. The $\bar{\Delta}$ statistic is calculated for each simulated tree, and the values from the set of simulated trees collectively compose the null distribution against which the observed value is compared to determine its cumulative probability.

Analyses. We tested the hypothesis that single-seeded fruits are a key innovation of Adoxaceae, using the program keymap v.0.1 (24) [which implements the approach of Ree (ref. 23)]. Inferences were conditioned on the summary chronogram for Adoxaceae (Fig. 2). The observed value of the $\bar{\Delta}$ statistic was estimated as the average of 10^4 stochastic mappings. The null distribution was approximated by generating a sample of 10^5 simulated chronograms. For each simulated tree, the value of the statistic was estimated as the average of 10^4 stochastic mappings.

Results. The key innovation test indicated that the single-seeded fruit form is not correlated with significantly elevated rates of diversification in Adoxaceae ($P = 0.296$: Fig. S2). We note that this test assumes that traits evolved under the *Mk*1 model of character evolution, which is evidently inappropriate for this trait (Table S3).

Combining results from independent events to test for an overall impact on diversification rate. We may wish to test for an overall correlation between a set of 'replicated' historical events of a similar type and diversification rate. For example, we may wish to ask whether the evolution of single-seeded fruits exhibits an overal correlation with rates of diversification in Adoxaceae. This can be achieved by simply combining the posterior predictive P-values for each of the independent events using conventional omnibus statistics. One such statictic, Fisher's combined probability test (25) (FCPT) evaluates whether the accumulation of information from independent tests of a similar null hypotheses collectively reject the global null hypothesis. The test estimates the composite probability that a set of j independent P-values will have a product equal to or smaller than that of the observed set. The FCPT statistic is calculated as:

$$FCPT = -2\sum_{i=1}^{j} \ln[P_i], \qquad \textbf{[S1]}$$

where the test statistic is distributed as χ^2 random variable with $2j$ degrees of freedom (25). Similar to the results obtained using the key-innovation test of Ree (23), combination of the two predictive probabilities under the FCPT indicates that the overall correlation between fruit form and diversification rates in Adoxaceae is non-significant under the two-tailed test (i.e., at the conventional $\alpha = 0.025$; $P = 0.034$).

1. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
2. Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
3. Swofford DL (2000) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0. Sinauer Associates, Sunderland, MA.
4. Drummond AJ, Ho SYW, Phillis MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88.
5. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
6. Moore BR, Donoghue MJ (2007) Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *Am Nat* 170:S28–S55.

7. Baskin JM, *et al*. (2006) Evolutionary considerations of the presence of both morphophysiological and physiological seed dormancy in the highly advanced euasterids II order *Dipsacales*. *Seed Sci Res* 16:233–242.
8. Muller J (1981) Fossil pollen records of extant angiosperms. *Bot Rev* 47:1–142.
9. Gruas-Cavagnetto C (1978) Etude palynologique de l'Eocene du Bassin Anglo-Parisien. *MŽmoires de la SociŽtŽ GŽologique de France, Nouv Ser* 56, *Mem* 131:1–164.
10. Hughes C, Eastwood R (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proc Natl Acad Sci USA* 103:10334–10339.
11. Lavin M, Herendeen PS, Wojciechowski MF (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Syst Biol* 54:575–594.
12. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53:673–684.
13. Pagel M, Meade A (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am Nat* 167:808–825.
14. Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of ancestral states. *Syst Biol* 50:351–366.
15. Nielsen R (2002) Mapping mutations on phylogenies. *Syst Biol* 51:729–739.
16. Huelsenbeck JP, Nielsen R, Bollback, JP (2003) Stochastic mapping of morphological characters. *Syst Biol* 52:131–158.
17. Bollback JP (2006) SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* 7:88.
18. Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst Biol* 50:913–925.
19. Pagel M, Meade A (2007) BayesTraits v1.0. Available at www.evolution.rdg.ac.uk.
20. Ree RH, Moore BR, Donoghue MJ (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.
21. Moore BR, Smith SA, Donoghue MJ (2008) A likelihood model for incorporating fossil data in the inference of historical biogeography. *Evolution*, in press.
22. Smith SA (2006) AReA: Area Reconstruction Analysis v. 2.1. Available at http://blackrim.org/programs/area.html.
23. Ree RH (2005) Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* 59:257–265.
24. Smith SA (2006) Keymap: key innovation testing using stochastic mapping v. 0.1. Available at http://blackrim.org/programs/keymap.html.
25. Fisher RA (1932) *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh), 4th Ed.
26. Donoghue MJ, Baldwin BG, Li J, Winkworth RC (2004) *Viburnum* phylogeny based on the chloroplast *trnK* intron and nuclear ribosomal ITS DNA sequences. *Syst Bot* 29:188–198.
27. Winkworth RC, Donoghue MJ (2005) *Viburnum* phylogeny based on combined molecular data: implications for taxonomy and biogeography. *Am J Bot* 92:653–666.

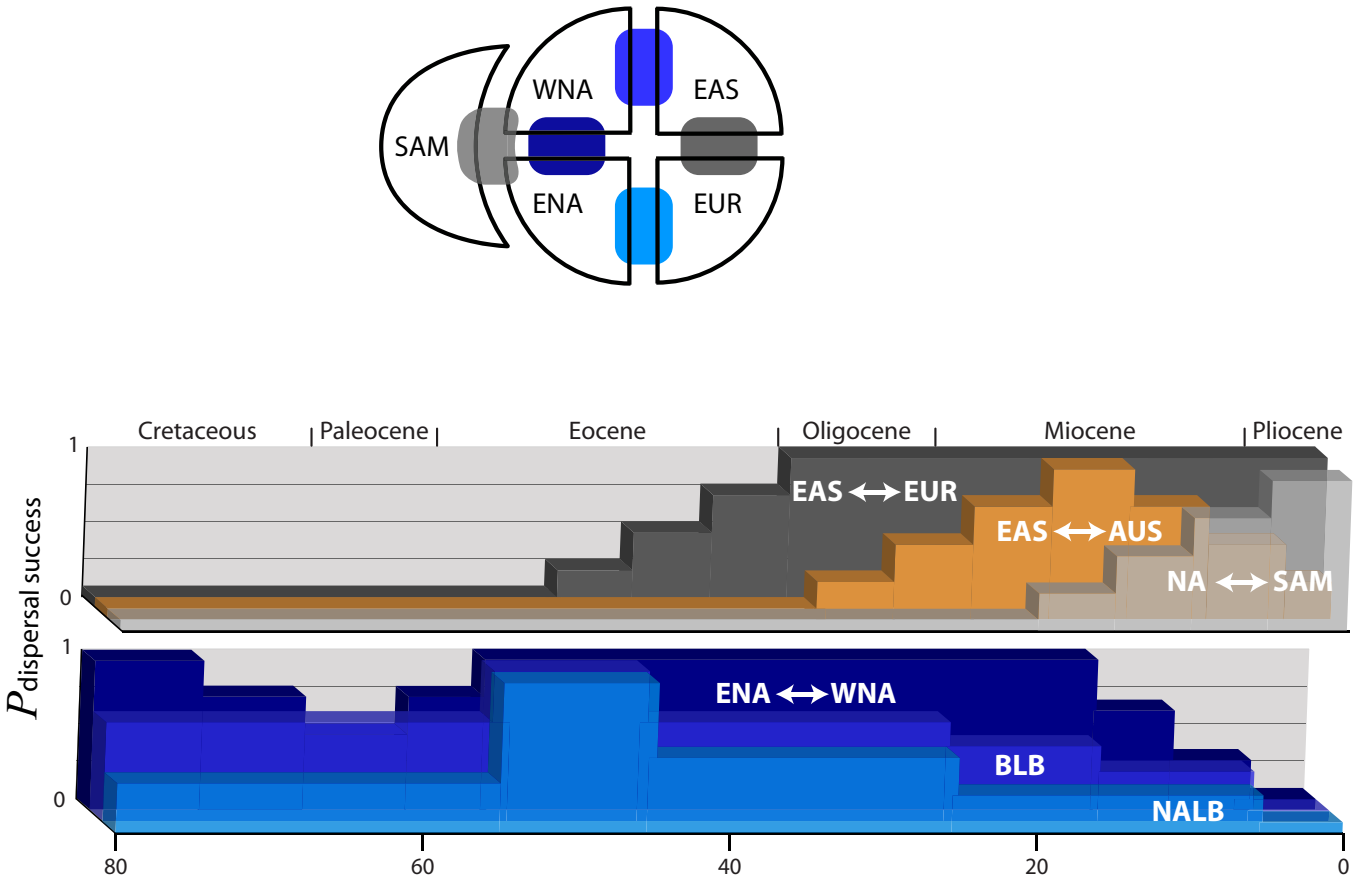

Fig. S1. The biogeographic model used to infer episodes of dispersal in *Lupinus*. The model comprises 6 areas, with 6 corresponding connections between those areas. The probability of successful dispersal across each connection is plotted through time and is symmetrical with respect to direction. Essentially, this biogeographic model specifies the probability that a given lineage will successfully move from one area to another by conditioning the intrinsic rates of lineage dispersal. Abbreviations: EAS, eastern Asia; EUR, Europe; ENA, eastern North America; WNA, western North America; SAM, South America; NA, North America (i.e., WNA + ENA); NH, Northern Hemisphere (i.e., EUR + EAS + NA); BLB, Bering land bridge; NALB, North Atlantic land bridge.

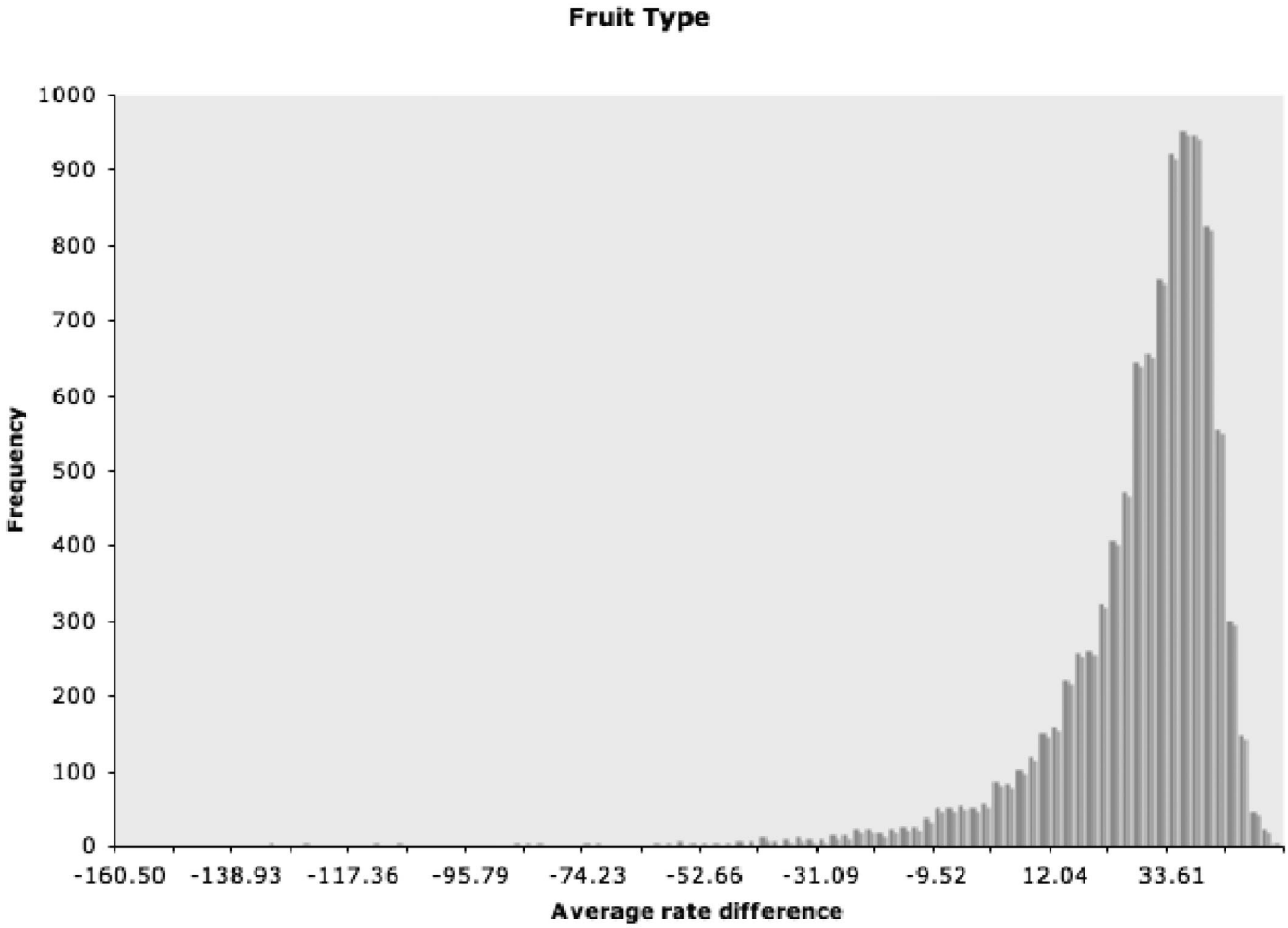

Fig. S2. Testing the hypothesis that fruit type is a key innovation of Adoxaceae. The histogram depicts the null distribution of the test statistic, $\bar{\Delta}$, and the vertical line indicates the observed value of the test statistic, with the corresponding posterior-predictive probability. Results of this application of the key innovation test of Ree (23) indicate that fruit type is not significantly correlated with diversification rate.

Table S1. Properties of the gene regions used to estimate phylogenies for Adoxaceae and *Lupinus*

Taxa	Partition	Source*	*N* taxa	*N* sites	Model selection†	Substitution-rate variation		Substitution-rate autocorrelation¶
						LRT‡	PP§	
Adoxaceae	Combined		63	4263	GTR I Γ	≫0.001	0.364	−2.99*E*-2
	ITS	*i*, *ii*	63	575	GTR I Γ	≫0.001	—	—
	matK	*i*, *ii*	63	1146	TVM Γ	≫0.001	—	—
	trnS-trnG	*iii*	63	655	F81 Γ	≫0.001	—	—
	WAXY 1	*ii*, *iii*	51	994	HKY Γ	≫0.001	—	—
	WAXY 2	*ii*, *iii*	45	894	HKY Γ	≫0.001	—	—
Lupinus	Combined		83	1959	TrN Γ	≫0.001	0.429	−5.02*E*-3
	ITS	*iv*	83	1140	TrN Γ	≫0.001	—	—
	LEGCYCI	*iv*	83	819	TIM Γ	≫0.001	—	—

*(*i*) Donoghue *et al*. 2004 (26), (*ii*) Moore and Donoghue 2007 (6), (*iii*) Winkworth and Donoghue 2005 (27), (*iv*) Hughes and Eastwood 2006 (10).
†Based on ΔAIC values calculated for the set of molecular substitution models evaluated in ModelTest (2).
‡Based on likelihood-ratio test of log-likelihood scores obtained with and without enforcing substitution rate constancy.
§Based on marginal posterior probability estimate of the coefficient of variation in substitution rate estimated with BEAST.
¶Based on marginal posterior probability estimate of the covariance of substitution rate variation estimated with BEAST.

Table S2. Temporal constraints used to estimate divergence times for Adoxaceae and *Lupinus*

Taxa	Node*	Prior distribution	Upper bound	Lower bound	$x/\sigma^{\dagger}$
Adoxaceae	1	Lognormal	86.0	46.0	1.1/1.1
	2	Lognormal	84.0	45.0	1.1/1.1
Lupinus					
	1	Normal	21.0	16.0	18.5/1.0

*Node numbers correspond to those depicted in Figures 2 and 4.
†The mean and standard deviation of the prior distribution.

Table S3. Bayesian inference of the evolution of fruit morphology in Adoxaceae

			Pr(model)§				
ln(*L*)*	Pr(q_{01})† x/σ†	Pr(q_{10})‡ x/σ†	*0Z*	*01*	*00*	*10*	*Z0*
−6.124	0.014/0.009	0.0004/0.011	0.94	0.00	0.06	0.00	0.00

*Marginal log likelihood.

†The mean and standard deviation of the marginal posterior probability distribution for the instantaneous rate of forward change integrated over all models.

‡The mean and standard deviation of the marginal posterior probability distribution for the instantaneous rate of backward change integrated over all models.

§Posterior probability of models of trait evolution sampled by the RJMCMC: model abbreviations are described in the text.

Other Supporting Information Files

Table S4
Table S5
Dataset S1
Dataset S2