Supporting Information

Tsai and Manos 10.1073/pnas.1006225107

SI Materials and Methods

***Epifagus virginiana* Collection and Genetic Analyses.** Plants from the same locality were collected at least 1 m apart, ensuring sampling of different individuals. DNA extractions were performed on fresh, frozen, or silica gel dried corm or stem material following a standard CTAB protocol (1) or using DNeasy Plant Mini Kits (Qiagen). PCR products were produced by Taq DNA polymerase or Phusion High-Fidelity DNA polymerase (New England BioLabs). PCR products were cleaned using an enzymatic protocol (2) with (2 units) Exonuclease I and (0.1 unit) Shrimp Alkaline Phosphatase or (0.1 unit) Antarctic Phosphatase (New England BioLabs). Cycle sequencing was performed following standard procedures using BigDye Terminator v3.1 (Applied Biosystems). Dye-labeled cpDNA amplicons were analyzed on an ABI 3700 or Applied Biosystems 3730xL DNA Analyzer. Sequences were edited and aligned using Sequencher v4.8 (Gene Codes).

The cpDNA primers were specific to *E. virginiana* and designed on the basis of its published cpDNA genome (3) (Table S5). The recovered 41 unique cpDNA haplotypes were analyzed under maximum-parsimony, maximum-likelihood, and Bayesian frameworks using PAUP* v. 4.0b10 (4), GARLI v. 0.951 (5), and Mr. Bayes v. 3.1.2 (6), respectively. A statistical parsimony haplotype network was constructed using TCS v1.21 (7) (Fig. 1*B*). The haplotype network was consistent with the best trees constructed using any of the above approaches.

The microsatellites were developed following Zane et al.'s FIASCO method (8) (Table S5). Microsatellite products for regions MS10, MS34, and MS43 were produced with fluorescently labeled forward primers for genotyping. PCRs for the remaining regions were run following a three-primer method with a forward, a reverse, and a genotyping primer. The forward primer was augmented with a 17- to 19-bp tail of the same sequence as the genotyping primer. The genotyping primer was fluorescently labeled and the same genotyping primer could be used for multiple regions, saving on some primer costs.

Microsatellite PCR products were pooled into sets of two to three loci and run on an Applied Biosystems 3730xL DNA Analyzer with a ROX 500 size standard (Applied Biosystems). Genotypes were scored using GeneMarker v.1.80 (SoftGenetics). Because alleles were not equally spaced (e.g., every 2 bp), a stepwise, repeat-based mutation rate between alleles was not assumed in analyses. Instead, alleles of a locus were ranked by their size and assigned an index value.

The locus MS92 was more complex than the others with several-sized fragments (1–4) recovered for each sample; for instance, a sample might have bands at 345, 351, and 362 bp. *E. virginiana* is a diploid, and very few heterozygotes were recovered at the other loci (four heterozygous loci in three individuals identified, average $F_{IT} = 0.998$), so this pattern appears to be the case of single alleles resulting in multiple bands. Each set of bands was scored as a single allele with three possible alleles in total; all samples were assumed to be homozygous at this locus.

BAPS v.5.2 (9) was used to cluster the microsatellite data. We tried a range of max cluster numbers (max 100) and found 20–22 distinct clusters were supported with the data. Final runs were performed with 50 as the maximum number of clusters possible. Relationships among clusters were visualized via a neighbor-joining phenogram based on Kullback–Leibler (K-L) distances calculated in BAPS.

Phylogeographic visualizations of cpDNA haplotypes and microsatellite cluster assignments on maps (Fig. 1 *B* and *C* and Fig. S1) were created using PhyloGeoViz v. 2.4.4 (Y.-H. E. Tsai, http://phylogeoviz.org) and plotted in Google Earth Pro v.5.2.1.1329 (Google Inc.). Base layers of state boundaries were acquired from GPS Visualizer (A. Schneider, http://www.gpsvisualizer.com/kml_overlay) and modified using Adobe Photoshop and Illustrator CS4 (Adobe Systems). The geographic range map for *Fagus grandifolia* (10) was assumed to be the same for the parasite. The ice margin at the last glacial maximum was plotted from Dyke et al. (11). Colors were chosen on the basis of the relationships between haplotypes or clusters as seen in the network or phenogram.

Allelic richness after rarefaction and *F* statistics were calculated using the R packages vegan v.1.15–4 (J. Oksanen et al., http://CRAN.R-project.org/package=vegan) and hierfstat v.0.04–4 (12).

The black cpDNA haplotype and black microsatellite cluster were found only in the disjunct Mexican population and not included in any further analyses.

***F. grandifolia* Fossil Pollen Dataset.** The threshold of 2% used in calculating F.age and defining regions 13, 9, and 6 is consistent with values adopted in studies of European *Fagus* (13), and it produced colonization ages consistent with prior interpretations (14). Ten of the 1,261 grid cells within the present-day distribution of *F. grandifolia* had <2% *F. grandifolia* pollen at all time slices. For those cells, we estimated the arrival time (F.age) by averaging values in the adjacent 8 cells. All layers were masked to include only points within the present-day distribution; all other cell values were set to zero. To facilitate comparisons among pollen density layers, the time-slice and F.avgP layers were visualized using a common 10-quantile color ramp on the basis of their combined pollen frequency distributions. The F.varP and F.age layers were plotted on a 10-quantile color ramp on the basis of their separate distributions.

Monmonier Analysis Description and Parameters. This method works by first breaking the landscape into tessellations centered on the sampled localities. Then starting on the edge with the highest global pairwise genetic distance [i.e., Reynolds' distance (15) averaged across cpDNA and microsatellite loci] between adjoining localities, a barrier is identified that follows the highest pairwise distance at each edge intersection. Paths in both directions are recorded. This process is repeated by beginning with the next highest global pairwise distance and so forth to find the top one to five genetic barriers. Barriers were computed with the R package Adegenet (16) with nrun = 1–5, scanthres = 0, and threshold = NULL. To minimize barriers found at the edges of the range due to concavity issues, nine virtual points were added at (W 90.38, N 45.77), (W 90.55, N 44.11), (W 90.38, N 42.34), (W 89.05, N 40.82), (W 89.44, N 39.52), (W 91.01, N 38.34), (W 92.61, N 37.40), (W 91.83, N 40.96), and (W 93.69, N 39.41). The R package spatstat v.1.19–2 (17) was used to bin and average barrier data across all of the subsamples. An R script that carries out the cross-validation analysis is available from the first author.

IMa (19) Regional Definitions, Parameters, and Priors. In addition to the regional definitions based on host fossil pollen data (regions 13, 9, and 6 in Fig. 1*A*), two other ways of dividing the landscape were tried to assess the robustness of the estimated demographic parameters to locality assignment to a region. The exact placement of the division between the Northeast (region 9) and the Midwest (region 6) was expected to have little impact on results, because of the widespread sharing of blue alleles and the relatively stable allelic frequencies near the boundary (Fig. 1 *B* and

C). Hence, we focused our additional runs on the boundary between the South (region 13) and the Midwest (region 6), an area prone to error due to few fossil pollen sites. In this case, the exact location of the boundary could be important, resulting in changes in the number of alleles shared between regions. First, because diffuse host populations have been predicted in the area adjacent to the last glacial maximum (19–21), parasite populations could have been present much farther north of region 13. To model this scenario, we included all areas south of the ice margin into the southern region and shrank region 6, thus forming regions 13^{+} and 6^{-} (Fig. S3*A*). Second, populations in the lower Midwest (e.g., Indiana), contain many distinct alleles that are absent from adjacent areas (e.g., the orange cpDNA haplotype, the brown microsatellite cluster; Fig. 1 *B* and *C*). This area corresponds with a biogeographic region, the prairie peninsula, an area with outlying prairie communities and a general absence of trees (22). Interestingly, this area also corresponds with a suggested refuge for the host (19) that was confirmed using a similar Monmonier methodology to that used above (21). Because of the unique parasite alleles found there, and the possible importance of this biogeographic boundary, we created region PP bounded by prairie to the east and the north (Fig. S3*B*). Resulting demographic parameters for both additional regional definitions are shown in Table S2. The migration rates and divergence times between regions have the same relative rankings regardless of regional definition, so we believe our results are robust to the exact boundaries dividing the landscape.

Migration rates, population sizes, and time since divergence were estimated between all regional pairs (excluding the Mexican population) of the parasite, using the program IMa (18). The program was run in "M" mode with the following set of parameters: maximum population sizes ($4N\mu$) = 10–35, maximum migration rates = 10–40, maximum divergence time = 10, 30 chains with a geometric heat mode, heating parameters g1 = 0.95 and g2 = 0.8, and burn-in = 2.5 million steps. All analyses were run three times with different seed values for at least 1.5 million steps (and up to 15 million steps) following burn-in and appeared to have converged according to trendline plots.

E. virginiana has a high rate of selfing ($F_{IT} = 0.998$) that conflicts with IMa's assumption of random mating. Under this violation, we expect estimates of effective population sizes to be reduced by one-half (23) and that divergence times will be shorter due to a faster coalescent rate (24). However, because the selfing rate is consistent among regions ($F_{IT\text{-South}} = 1.00$, $F_{IT\text{-Northeast}} = 0.994$, $F_{IT\text{-Midwest}} = 0.998$), relative comparisons of parameters among regions are not problematic.

To assess the probability that the migration rate from one region was larger than that from another region, migration values were randomly drawn from each marginal distribution. If the value from the first distribution was greater than that from the second, a value of 1 was assigned. If smaller, a value of 0 was assigned. This was repeated 1,000 times. The average value of the 0 or the 1 assignments was equivalent to the probability that the first migration parameter was larger than the second. This method of assessing significance was repeated for each pair of migration parameters, divergence times, and effective population sizes.

Spatial Linear Regression Models: Data Transformation. Because linear models rely on assumptions of relationship linearity and homoscedasticity, several data transformations were tried to minimize assumption violations. Specifically, all data layers were normalized by performing Box–Cox transformations (25), centering the data, and scaling the data by their standard deviations using the qAnalyst v.0.5.1 R package (www.quantide.com). Analyses were performed on both the original and the transformed datasets. Because results were similar, only analyses performed on the original datasets are reported.

In 17 instances, multiple parasite localities fell within the same grid cell. In these cases, the data were pooled to form a single population per grid cell (eight pooled populations in total).

1. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15.
2. Werle E, Schneider C, Renner M, Völker M, Fiehn W (1994) Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res* 22: 4354–4355.
3. Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci USA* 89: 10648–10652.
4. Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)* (Sinauer, Sunderland, MA), p 4.0b10.
5. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis (Univ of Texas, Austin).
6. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
7. Clement M, Posada D, Crandall KA (2000) TCS: A computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659.
8. Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: A review. *Mol Ecol* 11:1–16.
9. Corander J, Siren J, Arjas E (2008) Bayesian spatial modeling of genetic population structure. *Comput Stat* 23:111–129.
10. Tubbs CH, Houston DR (1990) American Beech. *Silvics of North America: 2. Hardwoods. Agriculture Handbook 654*, eds Burns RM, Honkala BH (US Department of Agriculture, Forest Service, Washington, DC), Vol 2, pp 325–332.
11. Dyke AS, Moore A, Robertson L (2003) *Deglaciation of North America,* Geological Survey of Canada (Natural Resources Canada, Ottawa, ON, Canada), p 2.
12. Goudet J (2005) Hierfstat, a package for r to compute and test hierarchical *F*-statistics. *Mol Ecol Notes* 5:184–186.
13. Magri D, et al. (2006) A new scenario for the quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *New Phytol* 171: 199–221.
14. Davis MB (1983) Quaternary history of deciduous forests of eastern North-America and Europe. *Ann Mo Bot Gard* 70:550–563.
15. Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* 105:767–779.
16. Jombart T (2008) adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.
17. Baddeley A, Turner R (2005) spatstat: An R package for analyzing spatial point patterns. *J Stat Softw* 12:1–42.
18. 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.
19. McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology* 86:2088–2098.
20. Bennett KD (1985) The spread of *Fagus grandifolia* across eastern North America during the last 18000 years. *J Biogeogr* 12:147–164.
21. Morris AB, Graham CH, Soltis DE, Soltis PS (2010) Reassessment of phylogeographical structure in an eastern North American tree using Monmonier's algorithm and ecological niche modelling. *J Biogeogr* 37:1657–1667.
22. Transeau EN (1935) The prairie peninsula. *Ecology* 16:423–437.
23. Pollak E (1987) On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics* 117:353–360.
24. Nordborg M, Donnelly P (1997) The coalescent process with selfing. *Genetics* 146: 1185–1195.
25. Box GEP, Cox DR (1964) An analysis of transformations. *J R Stat Soc B* 26:211–252.

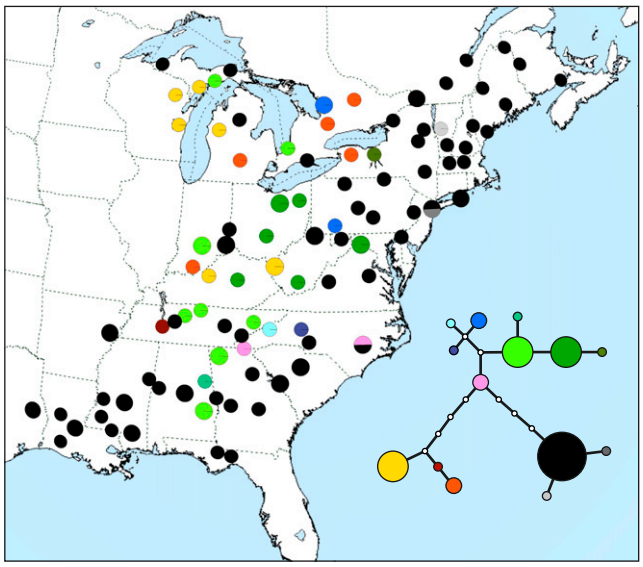

Fig. S1. Genetic structure of *Fagus grandifolia*. Distribution of cpDNA haplotypes and haplotype network (*Inset*) are shown. Redrawn from McLachlan et al. (20). Map was drawn in Google Earth (copyright 2010 Google).

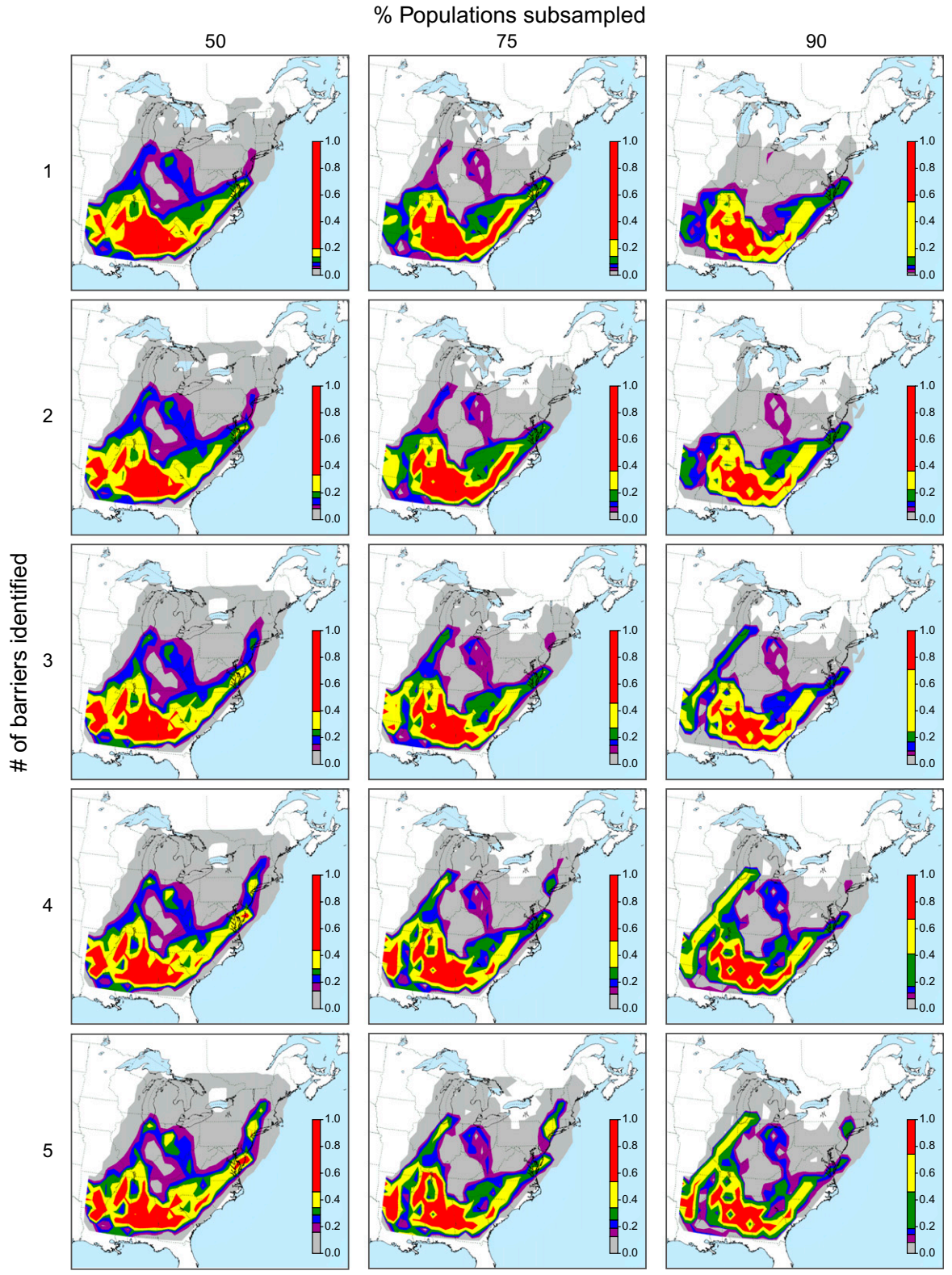

Fig. S2. Robustness of Monmonier analyses of *Epifagus virginiana*. Shown are areas of genetic differentiation found by Monmonier path analyses on cpDNA and microsatellite datasets. Colors correspond to the average number of genetic breaks found at that location over 1,000 cross-validation subsamples. Percentages of populations subsampled vary across columns (50, 75, and 90%), and the number of barriers identified per subsample varies by row (rows 1–5). Color assignments were based on the 50th–90th percentiles: red, 90+%; yellow, 80–89%; green, 70–79%; blue, 60–69%; purple, 50–59%. Maps were drawn in Google Earth (copyright 2010 Google).

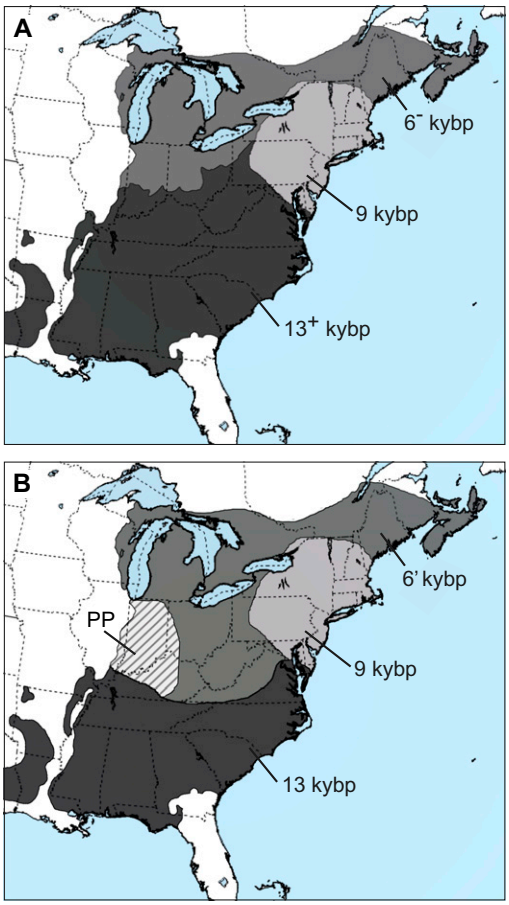

Fig. S3. Alternate regional definitions of *Epifagus virginiana* used in IMa analyses. (*A*) Regions defined by the maximum extent of ice at the last glacial maximum. The boundary between the South and the Midwest has been moved further north to the last glacial maximum. (*B*) In addition to regions defined by host fossil pollen, the area of the prairie peninsula is elevated to regional status. See *SI Materials and Methods* for discussion and justification of these regional boundaries. Maps were drawn in Google Earth (copyright 2010 Google).

Table S1. Comparisons of migration rates among regions of *Epifagus virginiana*

	9→13	6→13	13→9	6→9	13→6	9→6
9→13		0.53	0.10	0.03	0.86	0.01
6→13	0.46		0.06	0.00	0.89	0.00
13→9	0.90	0.94		0.47	0.98	0.28
6→9	0.97	1.00	0.56		1.00	0.18
13→6	0.15	0.11	0.02	0.00		0.00
9→6	0.99	1.00	0.72	0.81	1.00	

Each value is the probability that the migration rate of the row heading is greater than that of the column. Headings refer to migration rates from a region into another forward in time. Regions used in comparisons are defined in Fig. 1*A*. Probabilities were estimated by comparing random samples from the posterior distributions of estimates shown in Table 1.

Table S2. Population demographic parameters of *Epifagus virginiana* with alternate regional definitions

r1	r2	$\theta_1(4N_1\mu)$	$\theta_2(4N_2\mu)$	$\theta_A(4N_A\mu)$	$m_{2\to1}(m_{2\to1}/\mu)$	$m_{1\to2}(m_{1\to2}/\mu)$	$m_{\text{eff }2\to1}(2N_1m_{2\to1})$	$m_{\text{eff }1\to2}(2N_2m_{1\to2})$	$t(t\mu)$
A									
13^+	9	1.68	0.26	8.11	0.10	4.01	0.08	0.51	0.09
		(0.56–2.95)	(0.08–0.60)	(4.98–26.88)	(0.01–3.40)	(0.65–13.28)			(0.04–0.72)
13^+	6^-	3.91	9.05	6.21	0.29	0.10	0.57	0.44	0.96
		(2.95–5.03)	(6.92–11.60)	(3.72–21.46)	(0.10–0.76)	(0.01–0.41)			(0.56–2.80)
9	6^-	0.42	208.52	6.43	4.57	10.12	0.95	1054.84	0.40
		(0.08–0.58)	(111.09–207.48)	(5.39–89.53)	(2.41–13.49)	(8.27–13.01)			(0.18–1.06)
B									
13	9	0.73	0.25	8.63	0.61	0.01	0.22	0.00	0.06
		(0.33–1.66)	(0.08–0.59)	(5.23–27.86)	(0.17–5.83)	(0.22–12.89)			(0.02–0.34)
13	6′	2.74	10.28	5.59	0.44	0.02	0.61	0.12	0.64
		(1.80–3.68)	(8.11–22.61)	(3.43–22.54)	(0.08–1.10)	(0.01–0.37)			(0.24–3.02)
13	PP	1.13	0.56	7.24	1.01	9.35	0.57	2.60	0.10
		(0.60–2.06)	(0.11–1.00)	(3.90–26.53)	(0.26–7.79)	(2.60–14.63)			(0.04–0.54)
9	6′	0.42	212.84	6.57	5.63	9.31	1.18	990.50	0.44
		(0.08–0.58)	(106.15–211.78)	(5.15–75.42)	(2.68–10.96)	(7.31–11.92)			(0.20–0.98)
9	PP	0.25	0.33	7.90	1.13	1.34	0.14	0.22	0.04
		(0.08–0.59)	(0.11–0.77)	(4.59–19.30)	(0.16–9.53)	(0.26–8.53)			(0.00–0.18)
6′	PP	10.16	1.66	3.50	0.05	0.47	0.27	0.39	0.16
		(8.45–73.05)	(0.55–2.54)	(2.76–61.37)	(0.01–0.29)	(0.10–1.66)			(0.12–6.70)

Regions (r1 and r2) are defined in Fig. S3 *A* and *B* and in *SI Materials and Methods*. θ_1, θ_2, and θ_A refer to the scaled effective population sizes of r1, r2, and the ancestral population, respectively. $m_{2\to1}$ and $m_{1\to2}$ are the scaled migration rates forward in time from r2 to r1 and vice versa. $m_{\text{eff}2\to1}$ and $m_{\text{eff}1\to2}$ are the effective migration rates (taking into account population size). t is the scaled divergence time between the regions. All values (except m_{eff}) are scaled by the unknown per gene per generation mutation rate, μ. m, migration rate per gene per generation; n, effective population size; t, time in generations. The 95% credible interval is given below each estimate. Because m_{eff} is a point estimate, no credible interval is given.

Table S3. Comparisons of divergence times among regions of *Epifagus virginiana*

	13–9	13–6	9–6
13–9		0.01	0.04
13–6	0.99		0.84
9–6	0.96	0.16	

Each value is the probability that the divergence time of the row heading is greater than that of the column. Regions are defined in Fig. 1*A*. Probabilities were estimated by comparing random samples from the posterior distributions of estimates shown in Table 1.

Table S4. Univariate host pollen density models predicting genetic distances of *Epifagus virginiana*

Time slice	Intercept	Coefficient	*Z*-value	ρ	AIC	ΔAIC	wAIC
F_{00}	0.04281	−0.00189	0.002	0.949	−5355.3	5.1	0.04
$F_{0.5}$	***0.04669***	***−0.00109***	***0.000***	***0.945***	***−5360.4***	***0.0***	***0.52***
F_{01}	**0.04536**	**−0.00083**	**0.000**	**0.946**	**−5358.6**	**1.7**	**0.22**
F_{02}	**0.04518**	**−0.00073**	**0.001**	**0.946**	**−5357.8**	**2.6**	**0.14**
F_{03}	0.04173	−0.00062	0.003	0.950	−5354.9	5.4	0.03
F_{04}	0.03783	−0.00048	0.019	0.954	−5351.7	8.7	0.01
F_{05}	0.03715	−0.00044	0.024	0.955	−5351.2	9.1	0.01
F_{06}	0.03628	−0.00047	0.029	0.956	−5350.9	9.5	0.00
F_{07}	0.03434	−0.00044	0.090	0.958	−5349.0	11.4	0.00
F_{08}	0.03324	−0.00060	0.204	0.959	−5347.7	12.6	0.00
F_{09}	0.03048	0.00001	0.993	0.961	−5346.1	14.3	0.00
F_{10}	0.03339	0.00207	0.041	0.956	−5350.3	10.1	0.00
F_{11}	0.03402	0.00103	0.037	0.955	−5350.5	9.9	0.00
F_{12}	0.03479	0.00034	0.031	0.955	−5350.8	9.6	0.00
F_{13}	0.03246	0.00015	0.183	0.958	−5347.9	12.5	0.00
F_{14}	0.03417	0.00055	0.047	0.955	−5350.1	10.3	0.00
F_{15}	0.03489	0.00170	0.012	0.954	−5352.5	7.9	0.01
F_{16}	0.03258	0.00337	0.064	0.958	−5349.6	10.8	0.00
F_{17}	0.03135	0.00659	0.188	0.960	−5347.9	12.5	0.00
F_{18}	0.03027	−0.00300	0.804	0.962	−5346.2	14.2	0.00
F_{19}	0.03288	0.03297	0.074	0.957	−5349.3	11.1	0.00
F_{20}	0.03067	0.00379	0.743	0.961	−5346.2	14.1	0.00
F_{21}	0.03066	0.00844	0.813	0.961	−5346.2	14.2	0.00

Coefficients and their corresponding *Z*-values refer to host data layer coefficients, whereas ρ is the spatial lag coefficient. All values of ρ were significant ($P < 0.01$), and spatial lag models were significantly superior to standard models with no spatial components. The remaining residuals in any of the models were not spatially correlated as indicated by Lagrange multipliers. AIC, ΔAIC, and weighted AIC values are reported following Wagenmakers and Farell (1). The three models with >0.10 in probability are in boldface type; the best model, $F_{0.5}$, is in boldface type and italics.

1. Wagenmakers E-J, Farrell S (2004) AIC model selection using Akaike weights. *Psychon Bull Rev* 11:192–196.

Table S5. Primers used to amplify loci of *Epifagus virginiana*

Locus	Type	Sequence
cpDNA		
clpP	Forward	AATGGTTTGCCTGTCCTTTG
	Reverse	ACGTTTAGCATTCCCTCACG
rbcL+atpB	Forward	GACTGAAAATCCTAGTGCCATCA
	Reverse	ACTAAACCGCCATCTTTCCA
Microsatellites		
MS10	Forward	GGTTGGAGAGGAAAAAGGAAA
	Reverse	TGTGTGGAGAGGTTGTGTTGA
MS34	Forward	TGTATTTGCACTGACGGATTC
	Reverse	CGCTCGGTGAATGAGAAAA
MS43	Forward	GTCAAAAATCAGTCCGAGCA
	Reverse	GAATCCATAACACAAAGATGTTGC
MS105	Forward	CAGGAAACAGCTATGACTAGCTTCCCCTCCAATTGCT
	Reverse	AGACTGCAATGTCCCCACAC
	Genotyping	M13 Rev
MS63	Forward	CAGGAAACAGCTATGACGATTTCCCATTGTGGTGCAT
	Reverse	GACCTGCTTGCTTGCATAAAA
	Genotyping	M13 Rev
MS76	Forward	CAGGAAACAGCTATGACTGGGCCAACTAACTAGGGTAA
	Reverse	TTCTGGAAATGAAAGGGAGAAG
	Genotyping	M13 Rev
MS92	Forward	CACGACGTTGTAAAACGACGCTGTTGTCAGGCACTCTG
	Reverse	TCCCCCTCTCACTCTCACTC
	Genotyping	M13 Tail
MS130A	Forward	CACGACGTTGTAAAACGACCAGAGAAAAGGATAAACGACATATCA
	Reverse	CCAAAGGAGACATAAGGGGTAG
	Genotyping	M13 Tail
MS135	Forward	CACGACGTTGTAAAACGACGAGAAACAAGGGAAGTCAAGGA
	Reverse	GTGGACCTGGAGTCTCTGCT
	Genotyping	M13 Tail
M13 Rev		CAGGAAACAGCTATGAC
M13 Tail		CACGACGTTGTAAAACGAC

Sequences are given for the forward, reverse, and (if applicable) genotyping primers for each locus. The genotyping primers were used in three-primer amplifications, where the forward primer contained a tail that matched the genotyping primer.

Other Supporting Information Files

Dataset S1 (XLS)