

Supporting Information

Konrad Lohse¹, Nicholas H. Barton², George Melika³, Graham N. Stone¹

¹Institute of Evolutionary Biology

University of Edinburgh

Kings Buildings

Edinburgh EH9 3JT, UK

² Institute of Science and Technology

Am Campus 1

A-3400 Klosterneuburg

Austria

³ Pest Diagnostic Laboratory

Plant Protection & Soil Conservation Directorate of County Vas

Ambrozy setany 2, 9762 Tanakajd

Hungary

Table 1: Sampling and rearing information of individuals used for sequencing.

species	code	host	country	locality	col. date	oak host
<i>Caenacis lauta</i>	Clau32	<i>Cynips korsakovi</i>	Iran	Azerbaijan, Ardabil For.	10-04	<i>Q. macranthera</i>
<i>Cecidostiba fungosa</i>	Cfun88	<i>Andricus lucidus</i>	Iran	Lorestan, Piran Shahr	10-04	<i>Q. infectoria</i>
	Cfun70	<i>Callirhytis glandium</i>	Hungary	Szentkut	05-02	<i>Q. cerris</i>
	Cfun139	<i>Andricus quercustozae</i>	Spain	Avila, Puerto de Villatoro	03-02	<i>Q. pyrenaica</i>
<i>Cecidostiba semifascia</i>	Csem40	<i>Biorhiza pallida</i>	Iran	Kordestan, Marivan	10-04	<i>Q. infectoria</i>
	Csem60	<i>Biorhiza pallida</i>	Hungary	Godollo	07-01	<i>Q. robur</i>
	Csem50	<i>Biorhiza pallida</i>	Spain	Madrid, El Escorial	02-06	<i>Q. pyrenaica</i>
<i>Hobbya stenonota</i>	Hste48	<i>Pseudoneuroterus macropterus</i>	Iran	Kemanshah, Javanrod	2004	<i>Q. brantii</i>
	Hste60	<i>Aphelonyx cericola</i>	Hungary	Godollo	05-02	<i>Q. cerris</i>
	Hste55	<i>Andricus kollari</i>	Spain	Madrid, El, Escorial	02-06	<i>Q. pyrenaica</i>
<i>Mesopolobus amaenus</i>	Mama50	<i>Dryocosmus caspiensis</i>	Iran	Mazandaran	04-04	<i>Q. castaneifolii</i>
	Mama51	<i>Andricus glossulariae</i>	Hungary	Vitnyed	10-04	<i>Q. infectoria</i>
	Mama55	<i>Andricus burgundus</i>	Spain	Caldes de Malavella	???	<i>Q. suber</i>

Table 2: To test the robustness of the model selection to differences in the number of sampled loci between species, analyses for *C. fungosa* and *M. amaenus* were repeated for 2 subsets of the data including only those loci available for *C. semifascia* (1) and *H. stenonota* (2). The model with the highest $\ln L$ is shown in bold, the simplest model retained in likelihood ratio tests is indicated by *. Note that the set of supported models agrees between analyses on the two subsampled and the full (All) datasets in *M. amaenus*.

Model	k	<i>C. fungosa</i>			<i>M. amaenus</i>		
		All	1	2	All	1	2
panmixia	1	-122.82*	-75.82	-63.82*	-90.09	-52.85	-54.43
polytomy	2	-122.59	-73.26*	-63.78	-87.30	-50.47	-53.09
two-pop.	3	-120.77	-73.19	-63.41	-79.58*	-48.28*	-49.17*
full model	4	-120.01	-71.60	-61.95	-79.49	-47.39*	-48.84
C & W topologies	3	polytomy	polytomy	polytomy	polytomy	polytomy	polytomy

Figure 1: $\Delta \ln L$ plots for ancestral N_e for four oak gall parasitoid species. The maximum likelihood estimate for *C. fungosa* (black) is much larger than that of any other species (*C. semifascia*= blue, *H. stenonota*=green, *M. amaenus*=red). In each species, estimates under all divergence models with substantial support (see Table 2) are shown (Full model = thick dashed lines, two-pop.. = thin dashed lines and polytomy = solid lines). The horizontal line delimits the region of 95 % confidence.

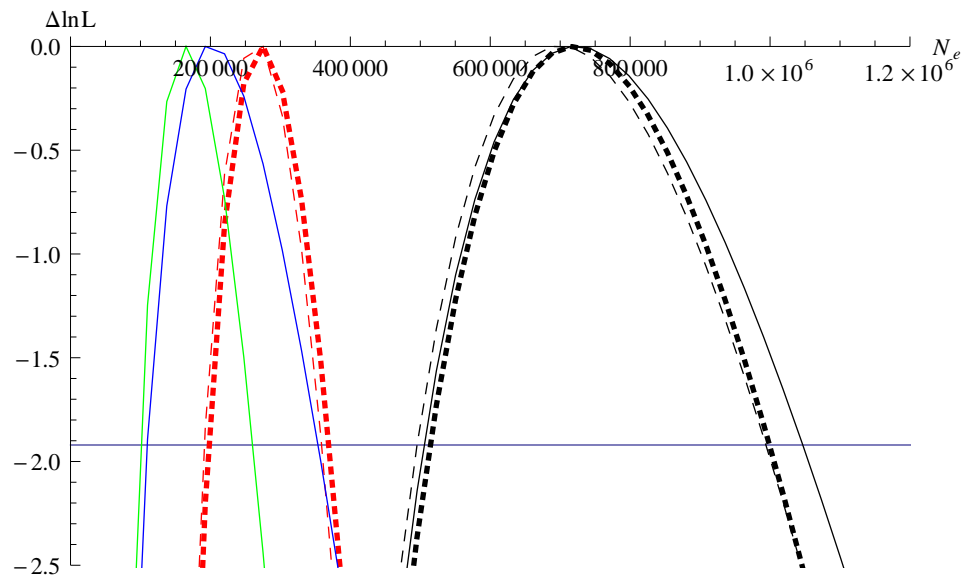


Figure 2: $\Delta \ln L$ plots for *C. fungosa* (top row) and *M. amaenus* (bottom row) for population divergence estimated using either all loci (Full) or only those available for *C. semifasica* (subset 1) or *H. stenonota* (subset 2). Full model = thick dashed lines, two-pop. = thin dashed lines and polytomy = solid lines. The horizontal line delimits the region of 95 % confidence. Although $\ln L$ trajectories are wider when subsets of the data are used, the inference of a significantly older history for *M. amaenus* compared to the other 3 species is robust to differences in the number of loci.

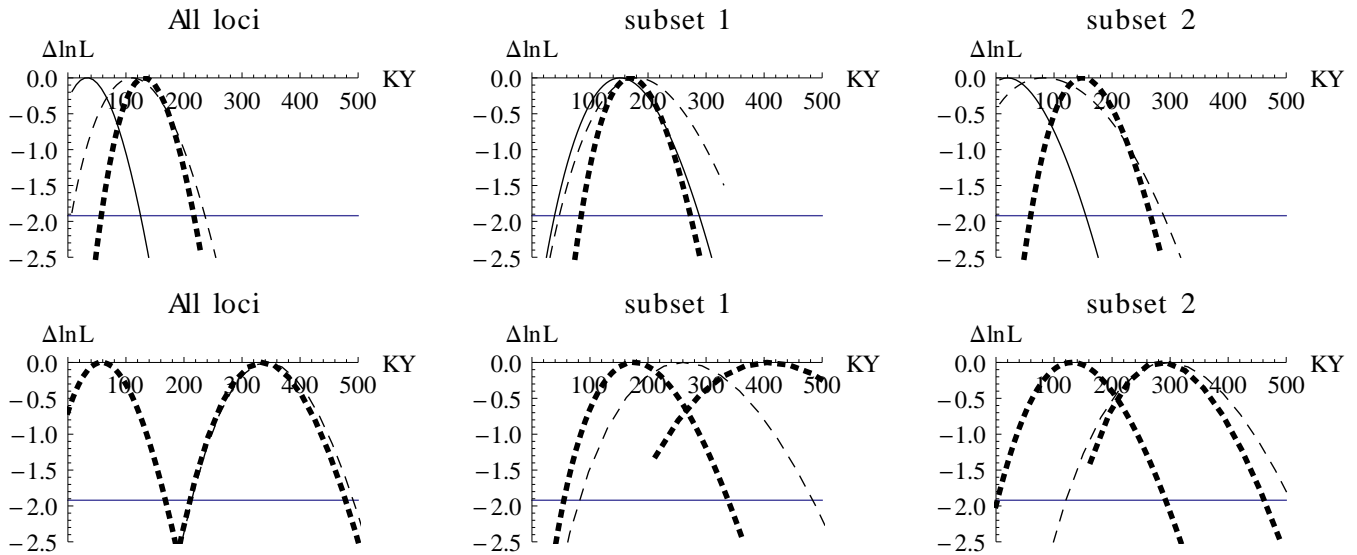


Figure 3: The power to distinguish between alternative models of population divergence in the face of gene flow $M = 2N_e m$. We tested the model comparison the LRT on datasets of 18 loci simulated under an 'Out of the East' divergence history. The following parameters were fixed $\theta = 1.5, T_1 = 0.4, T_2 = 1.1$. Each point shows the proportion of replicats (out of 100) for which a particular model was retained using LRT. Points were joined for ease of comparison: full model = thick dashed, 2 pop. = thin dashed, polytomy = solid lines and panmixia = dotted lines

