A likelihood-based comparison of populations histories in a parasitoid guild

Konrad Lohse[1], Nicholas H. Barton[2], George Melika[3], Graham N. Stone[1]

[1]Institute 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

Running head: Comparing divergence scenarios in a parasitoid guild

Keywords: Population divergence, maximum likelihood, comparative phylogeography, community assembly

Proofs to be sent to:

Konrad Lohse

Institute of Evolutionary Biology

University of Edinburgh

Kings Buildings

Edinburgh EH9 3JT, UK

Abstract

Little is known about the stability of trophic relationships in complex natural communities over evolutionary timescales. Here, we use sequence data from 18 nuclear loci to reconstruct and compare the intraspecific histories of major Pleistocene refugial populations in the Middle East, the Balkans and Iberia in a guild of four Chalcid parasitoids (*Cecidostiba fungosa*, *C. semifascia*, *Hobbya stenonota* and *Mesopolobus amaenus*) all attacking Cynipid oak galls. We develop a likelihood method to numerically estimate models of divergence between three populations from multilocus data. We investigate the power of this framework on simulated data, and — using triplet alignments of intronic loci – quantify the support for all possible divergence relationships between refugial populations in the four paraistoids. Although an East to West order of population divergence has highest support in all but one species, we cannot rule out alternative population tree topologies. Comparing the estimated times of population splits between species, we find that one species, *M. amaenus*, has a significantly older history than the rest of the guild and must have arrived in central Europe at least one glacial cycle prior to other guild members. This suggests that although all four species may share a common origin in the East, they expanded westwards into Europe at different times.

The past two decades have seen a proliferation of studies that use genetic data to draw inferences about the spatial history of species. Population genetic and phylogeographic studies have revealed that regional faunas and floras often share characteristic historical patterns (Avise, 1987). For example, the genetic signatures of past range contractions into southern refugia during glacial maxima followed by expansion out of them into northern areas during warm period have been found in many temperate species (Hewitt, 2000; Schmitt, 2007). Likewise, the same unglaciated areas have acted as refugia for many species and, in Europe, genetic diversity within those southern refugia often shows a decline from east to west, suggesting an earlier, longitudinal spread in that direction (e. g. Koch *et al.*, 2006; Atkinson *et al.*, 2007; Duvaux *et al.*, 2011).

This historical perspective, which seeks to understand how species distributions changed over evolutionary timescales, has been largely absent from the field of community ecology (Hickerson *et al.*, 2010), which instead views regional community composition in terms of the life histories of component species. It therefore remains unclear how trophic links within regional communities have been affected by the drastic range shifts associated with Pleistocene climate cycles. Although phylogenetic studies have demonstrated co-divergence of parasitoids and their associated hosts at the species and deeper levels (Lopez-Vaamonde *et al.*, 2001), few attempts have been made to systematically compare intraspecific histories within communities (but see DeChaine & Martin, 2006; Smith *et al.*, 2011; Dolman & Joseph, 2012). While there are striking examples of specialist associations with tightly linked histories such as highly specialized parasitic or symbiotic interactions (e. g. Hoberg & Brooks, 2008; Espíndola & Alvarez, 2011), the great majority of species share diffuse trophic links with many species rather than strong associations with few.

Oak gallwasps and their associated parasitoid chalcid wasp enemies are a case in point, and provide an excellent model for reconstructing community assembly from genetic data (Stone *et al.*, 2012). Like many insect herbivores (leaf miners, seed feeders etc), oak gall wasps support a diverse guild of chalcid parasitoids (over 100 species in Europe), which although obligate parasitoids of oak galls consists mainly of generalists

that attack a wide range of host galls (Askew, 1961a; Bailey *et al.*, 2009). One hypothesis for the ubiquity of generalism in this and similar temperate parasitoid guilds is that because of the glaciation-associated shifts in climate, species interactions have been repeatedly uncoupled, which limits the potential for co-evolution between hosts and parasitoid and instead selects for minimal host specificty (Stone *et al.*, 2012). If this was the case, we expect to find evidence for incongruent histories within parasitoid guilds.

For classical phylogeography, which has in the past focused overwhelmingly on describing patterns in mitochondrial sequence data, finding concordance across co-distributed species (Avise, 1987) has provided perhaps the best justification for interpreting these patterns in a qualitative way in the first place. However, if we want to actually test how concordant spatial histories are between species, we need a statistical, model-based framework (Edwards & Beerli, 2000; Nichols, 2001; Hickerson *et al.*, 2010; Lim & Sheldon, 2011).

Recently, we have investigated the temporal congruence of Pleistocene histories in the oak gall community by analysing a dataset of mitochondrial DNA sequences from 31 species under a hierarchical model of multispecies divergence between neighbouring pairs of refugia (Stone *et al.*, 2012). This study found that, with few exceptions, divergence between refugia occurred earlier in gallwasp hosts than in their parasitoids, supporting the idea that gallwasps escaped their enemies as they expanded westwards. However, the variance of the coalescent severely limits the information contained in a single locus (Wakeley, 2009). Thus, while Stone *et al.* (2012) were able to infer the number and age of multispecies divergence events across each guild, there was little power to reconstruct the history of any particular species. Furthermore, the analysis was limited to pairs of neighbouring populations, rather than considering multiple refugia jointly, and so did not examine the order of divergence (i. e. the population tree topology). Sampling multiple, independent loci provides the crucial replication to resolve intraspecific histories (Felsenstein, 2006). For example, Jennings & Edwards (2005) and Lohse *et al.* (2010) used likelihood (Yang, 2002) and Bayesian (Rannala & Yang, 2003) methods to estimate divergence times and effective sizes of ancestral populations from nuclear

loci sampled from just a single individual per population. For the oak gall parasitoid *Cecidostiba fungosa*, this model-based analysis supported an eastern Asian origin of Balkan and Iberian refuge populations with divergence from a common ancestral population at most one glacial cycle ago (Lohse *et al.*, 2010). While such minimal triplet samples are of course uninformative about the parameters of current populations, they do contain information about the historical relationships of these populations and are amenable to exact likelihood analysis. In other words, the likelihood of a particular model can be maximised directly from the mutational patterns observed across arbitrary numbers of unlinked loci without loss of information (Yang, 2002; Lohse *et al.*, 2011a).

Here, we extend the likelihood framework of Yang (2002) for triplet samples to investigate all possible population tree topologies and nested models within those topologies. We then apply this method to nuclear sequence data sampled from three refugial populations (the Middle East, the Balkans and Iberia) in four species of chalcid parasitoids of oak galls to compare their longitudinal histories. These include *Cecidostiba fungosa*, previously analysed by Lohse *et al.* (2010), and three other species; *C. semifascia*, *Hobbya stenonota* and *Mesopolobus amaenus*, all Pteromalid chalcids that exclusively attack oak galls (Askew, 1961b). We use likelihhoods to quantify the relative support for all possible divergence scenarios in each species and address three questions; i) Can we infer the order in which refugial populations diverged and – specifically — do all sampled members of the guild share the same population topology and hence a common origin? ii) Are population splitting times compatible with simultaneous divergence of the guild or can we rule out such synchrony? Using simluations we also asked how the power to distinguish between models depends on the timescale of divergence and the number loci and ask how robust these inference are to the presence of post-divergence gene flow.

Methods

Samples and sequencing

The sampling strategy followed Lohse *et al.* (2010). For each species, a single haploid male individual from each of three major Western Palearctic refugia in the Middle East (East) the Balkans (Center) and Iberia (West) was sequenced for a panel of 18 exon priming, intron crossing loci. These markers had previously been developed (Lohse *et al.*, 2011b) and analysed (Lohse *et al.*, 2010) for *C. fungosa* and the outgroup *Caenacis lauta* (GenBank accession nos HM208872-HM209026). East-Center-West triplets for 14 of these loci had been sequenced for *M. amaenus* as part of the marker development (GenBank accession nos HQ596410-HQ596457). Analogous datasets were generated for three individuals of two additional pteromalid species: *Cecidostiba semifascia* and *Hobbya stenonota* (Supporting Information, Table S1). Primers and PCR conditions are described in detail in Lohse *et al.* (2011b). PCR products were sequenced in both directions on an ABI Sanger platform using BigDye chemistry at the NERC GenePool facility, Edinburgh. Complementary reads were aligned using Sequencer v.4.8 and checked by eye. For each locus, ingroup and outgroup sequences were aligned in Muscle (Edgar, 2004). *C. lauta* was used as an outgroup for all four species (Table 1).

Custom made bio-python scripts (available from the authors upon request) were used to compute summary statistics (Watterson's θ), polarize alignments with respect to the outgroup and remove invariant sites and indels. The polymorphism information within each locus can be summarised by counting the six possible types of polarized mutations. Denoting the state of a given SNP as either ancestral (0) or derived (1) these can be written as (1 1 0), (1 0 1), (0 1 1), (1 0 0), (0 1 0) and (0 0 1), where entries in the list corresponds to the three sampling locations i. e. (West, Central, East). Assuming an infinite sites mutation model, each type of mutation corresponds to a unique branch in the genealogy (Patterson *et al.*, 2006). In particular, the

first three types are shared derived (i. e. parsimony informative) mutations which define a unique topology and so observing more than one type of these topologically informative mutations at a locus is incompatible with the assumption of infinite sites and no recombinations. We used this criterion to test for recombination in each alignment by testing for the presence of more than one type of shared derived mutation . This is analogous to the four-gamete test but only requires a minimum of three ingroup samples and therefore has greater power to detection to detect recombination. In total, only four alignments (out of a total of 53 across all four species) showed evidence for recombination and were trimmed to the longest fragment compatible with the assumption of no recombination and infinite sites. All trimmed, outgroup rooted alignmnens are available from Dryad (XXX).

Although the principal aim of our analysis was to compare the relative divergence of refugial populations between species rather than to obtain absolute values, we also applied a molecular clock. Following Lohse *et al.* (2010), a mutation rate (per site and generation) was calibrated using an estimate for the synonymous mutation rate in the closely related pteromalid wasp genus *Nasonia* of 1.375×10^{-8} per year (Oliveira *et al.*, 2008). To apply this to our data (all four species), this rate was multiplied by the ratio of average per site divergence (between *C. fungosa* and *C. lauta*) at synonomous coding sites and divergence across all sites (and loci). Although rate calibrations are notoriously error-prone (Pulquério & Nicholls, 2007), this calibration should at least give an order of magnitude timing of events. We initially tried to account for mutational heterogeneity between loci using the relative divergence between *C. fungosa* and *C. lauta* at each locus. However, given that this did not improve likelihoods and yielded qualitatively similar results (not shown), we assumed the simpler model of a constant (per site) mutation rate across loci in all subsequent analyses. The fact that accounting for mutational hetereogeneity did not improve model fit is perhaps unsurprising given that over very the recent timescales the stochastric variance of the coalescent and the mutational process are expected to outweight any differences in mutation rates between loci which are likely to be subtle in

comparison.

Likelihood computation and model selection

We assume a model of divergence between three populations (labeled A, B, C), such that populations B and C split from each other at some recent time T_1 whereas their shared ancestral population split from population A at a previous time T_1+T_2. Following Yang (2002), the effective size of the ancestral population of all populations is denoted N_0 while the size of the population ancestral to B and C is $N_1 = \frac{N_0}{\alpha}$. Note that because only one gene copy was sampled per population and the model assumes no gene flow between populations, we have no information about the current effective sizes (N_A, N_B, N_C). Divergence times are scaled by twice the effective size of the common ancestral population N_0, e. g. $t_1 = T_1 \times 2N_0 \times g$, where t_1 is the absolute divergence time between B and C and g is the generation time (both in years). All four species are known to have two generations per year (i. e. $g = 0.5$).

We used the recursion derived in Lohse *et al.* (2011a) to obtain an expression for the generating function (GF) of branch lengths under this model (see Appendix 1 and the *Mathematica* given in as Supporting Information). The GF allows calculation of the likelihood of model parameters given the mutational configuration (i. e. the counts of the types of mutations observed at a locus). Assuming that loci are unlinked, the joint likelihood of model parameters for a multilocus dataset is simply the product of likelihoods of individual loci (Hey & Nielsen, 2004).

Note that unlike the model of Yang (2002), our likelihood calculation assumes that genealogies are polarized using an outgroup sequence. All else being equal, this should increase power, but relies on the assumption of an infinite sites mutation model. For a given order of divergence, the full divergence model can be simplified in three ways; by setting either time interval T_1 or T_2 or both to zero. The resulting nested models include a two population divergence model (where populations B and C are joined) ($T_1 = 0$), a

single polytomous split between all three populations ($T_2 = 0$) and – in the simplest case – a single panmictic population ($T_1 = T_2 = 0$) (see Fig. 1). Given that there are three possible orders in which populations can split from each other (i. e. population tree topologies), we have eight models in total. To quantify the relative support for each model in each species, we numerically maximised the joint log likelihood (lnL) across loci using the *FindMaximum* function in *Mathematica* (Wolfram Research, 2010). We used likelihood ratio tests (LRT) to compare each model against all simpler, nested alternatives. Significance was assessed assuming that $2lnL$ follows a χ^2 distribution. The most complex model that provided a significantly better fit than all simpler models nested within it, it was accepted as the most parsimonious model.

Simulations

In order to ascertain how much power there is to distinguish between histories, we tested the model selection scheme on simulated data. Triplet datasets for three different sampling schemes (10, 18 and 100 loci of equal length and mutation rate) were simulated in *ms* (Hudson, 2002). Our aim was to include both the minimum and maximum number of loci available per species in the present study but also consider the gain in power that can be expected from increasing the number of loci by an oder of magnitude, which can be easily achieved using short-read sequencing technology. The power analysis was motivated by the parameters estimates obtained for the four parasitoids and focused on two Pleistocene timescales: Recent divergence was simulated by fixing the time of the oldest split $T_0 + T_1$ to 0.5. Assuming $\theta_0 = \theta_1 = 1.5$ (which for ease of comparison was fixed in all simulations) and nuclear mutation rate calibrations for insects, this correspond roughly to divergence one glacial cycle ago as inferred for *C. fungosa* and *H. stenonota* (see Results). More ancient divergence three glacial cycles ago (as inferred for *M. amaenus*) was simulated by fixing $T_0 + T_1 = 1.5$. In both cases, we kept the time of the oldest split $T_0 + T_1$ constant but varied the more recent divergence time T_1 from 0 to its maximum value. The two extremes for T_1 correspond to the

two-population and polytmony model respectively. We simulated 100 replicat datasets for each parameter combination and sampling scheme and recorded the most parsimonious model as detemined by the LRT for each dataset. Power can be measured simply as the proportion of replicats for which the true model is inferred correctly.

Results

In addition to the 18 and 14 outgroup rooted alignments available for *C. fungosa* and *M. amaenus* respectively, 10 and 11 loci amplified sucessfully in *C. semifascia* and *H. stenonota* (Table 1) (GenBank accession nos XXX). Mean per site diversity across loci as measured by θ_W was considerably higher in *C. fungosa* and *M. amaenus* than in *C. semifascia* and *H. stenonota* ($\theta_W = 0.0160$ and 0.0123 vs. 0.0050 and 0.0076 respectively). However, this difference was only significant for *C. semifascia* (Wilcoxon signed rank test, $p = 0.041$). Both *C. semifascia* and *H. stenonota* also contained a smaller proportion of topologically resolved genealogies (i.e. with parsimony informative sites) compared to the other two species (Table 1).

Model selection

In all four species, models that assume divergence of either the central or western population from a common ancestor as the oldest split (i. e. a non-eastern topology) had no support. In all cases, the maximum likelihood estimate (MLE) for T_2, i. e. the interval between population splitting events was 0 for both topologies. In other words, when fitting these two alternative orders of population splitting, the full model collapsed to a polytomy model. In contrast, under an "Out of the East" topology the MLE for T_2 was non-zero in all species except *H. stenonota* (Table 2).

In both *M. amaenus* and *C. fungosa*, the full "Out of the East" model (i. e. assuming an older divergence of the eastern population from a common ancestral population followed by divergence between central and

western refugia, Fig. 1a) had highest lnL (Table 2). In contrast, simpler models (polytomous or a two-population scenario with central and western populations joined (see fig. 1b and c)) had the highest lnL in *H. stenonota* and *C. semifascia* respectively. In both species, the MLEs for the full model were identical to those under simpler alternatives. However, in all species except *H. stenonota*, the models with the highest lnL were rejected in favour of simpler alternatives using the LRT. In *M. amaenus* the two-population model was retained as the most parsimonious model, whereas in *C. fungosa*, panmixia could not be rejected. While for *H. stenonota*, panmixia could be rejected, this was not possible for *C. semifacia*.

Comparing divergence parameters between species

To assess the evidence for simultaneous divergence between species, we compared MLEs for population divergence times under both the model retained in the LRT (Table 3 and Fig. 3) and all models that provided an improvement in lnL (regardless of whether this was significant). Two conclusions emerge from this: Firstly, estimates for the time of the oldest divergence event generally agree between supported models in each species. Figure 3 shows that this parameter has essentially identical lnL curves under the full and the two-population model in *M. amaenus* and very similar trajectories in *C. fungosa*. In contrast, the polytomy model in *C. fungosa* (and to a lesser extent *C. semifascia*) was associated with a markedly more recent population divergence than that estimated under the two-population model in this species (although the 95 % confidence intervals of these different estimates overlap considerably). Secondly, the divergence of the common ancestral population occurred almost simultaneously in *C. fungosa* and *H. stenonota*. Applying the *Nasonia* calibration, these divergence events fall roughly in the previous Eemian interglacial (131 KYA and 125 KYA for *C. fungosa* and *H. stenonota* respectively). Although, the MLE of the oldest divergence time in *C. semifascia* was more recent than that (59 KY), 95 % C. I. for all three species overlap broadly. In contrast, *M. ameanus* diverged much earlier (343 KY) with 95% C. I. not overlapping those of any other

species regardless of whether the full or a two population model is assumed (Table 3, Fig. 3).

Simulations and sensitivity analysis

Our simulations clearly show that for a large and biologically relevant parameter range the power to distinguish between divergence scenarios is limited. As one might expect power depends both on the number of loci and the depth of population divergence (Fig. 4). When divergence is recent ($T_1 + T_2 = 05$), the most extreme null model of a panmictic population can be rejected less than 50 % of the time, regardless of whether 10 or 18 loci are sampled. However, panmixia is almost always rejected (>95 %) for older divergence histories (i. e. $T_1 + T_2 = 1.5$) . However, even then, it is virtually impossible to correctly identify the (true) full divergence model with 18 loci or less. Instead, LRT almost always favours either one of the two simpler, nested model (polytomy or a 2-population scenario). Which of these two alternatives is supported depends on the relative timing of the more recent split, T_1. If the split is recent ($T_1 < 0.7$), there is strong support for the two population model, if divergence is old, the polytomy model wins out (Fig. 4B). Importantly, the simulation results mirror our inferences on the real data. For example, if we assume that the history inferred under the full model for *M. amaenus* was correct, figure 4B confirms that there is little power to reject the two-population model in favour of the (true) full model. in contrast, panmixia and a polytomous split are comparatively easy to reject, which is excatly what we observe for *M. amaenus*.

A disproportionate number of loci failed to amplify in *C. semifascia* and *H. stenonota*. Given that simpler models generally had higher support in these species compared to *C. fungosa* and *M. amaenus*, an obvious question is how robust our inferences are to the variation in the number of loci. To test for this, we repeated all analyses for *C. fungosa* and *M. amaenus* on two subsets of the data, in each case subsampling only those loci which amplified in either *C. semifascia* or *H. stenonota* (1 and 2 in Supporting Information Table S2). Note that using the same loci rather than just equal numbers in each species also controls for any bias in

amplification success (e. g. longer and hence more informative loci failed to amplify disproportionately in *C. semifascia* or *H. stenonota*). In both species we found that in almost all cases the same models were supported regardless of whether all (18 and 16 respectively) loci or only a subset were used in the analysis (Supporting Information Table S2). Specifically, the ranking of models according to lnL was the same in the subsampled and full analyses in all cases. Likewise, estimates of divergence times and ancestral N_e were comparable to those obtained from the full data in both species (Supporting Information Fig. S2). This confirms that our main results are robust to the differences in sampling effort between species.

Discussion

Our results highlight that even with multiple (10-18) independent loci it is surprisingly difficult to distinguish between simple alternative divergence histories. This is despite the fact that unlike methods that rely on summaries of the data (summary statistics or genetrees), our likelihood calculation uses all available information. As our simulations show, the historical signal contained in sequence data is inherently limited if histories are young. Importantly, the intraspecific histories considered here are recent both on the timescale of mutations and coalescence. In other words, most loci only contained a few variable sites and many were topologically unresolved and a considerable fraction only coalesce in the common ancestral population (Table 1). The same will be true for the Pleistocene histories of any species with large N_e. Despite this, there is no shortage of phylogeographic studies that claim to find signatures of much more complex histories than those we were able to investigate here. However, as has been pointed out before (Nichols, 2001; Knowles, 2002; Hey & Machado, 2003; Beaumont *et al.*, 2010; Barton *et al.*, 2010), few of these provide statistical tests for the historical scenarios they try to infer. While recent histories are hard (or indeed impossible) to resolve using the replication that has been possible using Sanger sequencing, our tests on simulated data show that hundreds of loci. This is encouraging, given the ease with .

Despite the limited ability to distinguish between models, our results demonstrate that key parameters, the time of the oldest split and the effective size of the common ancestral population, are robust to model uncertainty. Firstly, although we cannot rule out alterantive divergence histories under which either central or western populations diverged first for *C. fungosa*, *C. semifascia* and *H. stenonota* (particularly if the internode interval T_2 is short), our finding of improved likelihood under an "Out of the East" model is most compatible with a shared eastern origin of the entire guild, albeit a recent one in most cases. Support for an eastern origin has previously been found for several other parasitoid species (Hayward & Stone, 2006; Nicholls *et al.*, 2010) and their gallwasp hosts (Rokas *et al.*, 2003; Stone *et al.*, 2007; Challis *et al.*, 2007). Secondly, our comparison of relative divergence times across species shows that *M. amaenus* split into distinct refugial populations long before any of the other three species did and so we can rule out a strictly synchronous history in this parasitoid guild. This is in contrast to a recent meta-analysis based on a single locus (mitochondrial DNA) which found no evidecne for different divergence times between eastern and central refugial populations across 15 parasitoid species (Stone *et al.*, 2012). Notably however, *M. amaenus*, the outlier species in the present analysis, was not included in the Stone *et al.* (2012) study. It is worth pointing out that while our comparison between species does not rely on absolute molecular clock calibrations, it does assume that the genome wide mutation rate is comparable between these four species. Although the inferred difference in divergence time between *M. amaenus* and the other 3 species could in theory also be explained by a 2.5-3 fold lower mutation rate in *M. amaenus*, we believe that this is highly unlikely given that all species have the same generation time and are closely related.

Inferring intraspecific divergence histories comes with several challenges (Knowles, 2002; Hey & Machado, 2003). First, the order of divergence (i. e. the population tree topology) is generally not known *a priori*, but is rather one of the parameters to be inferred. Second, it is unclear to what extent a "population tree" is a useful description of population history in the first place. More realistic models of population relationships

may include secondary gene flow (Hey & Nielsen, 2004) or admixture between populations (Hellenthal *et al.*, 2008) or view individuals living in a spatial continuum with no discrete structure at all (Wright, 1943; Barton *et al.*, 2010). However, with few exceptions (Hey & Nielsen, 2004), we lack quantitative methods to estimate parameters under such more complex scenarios or compare them to simpler alternatives. Furthermore, an exhaustive search of model space quickly becomes unfeasible for more parameter-rich models. For example, there are thousands of ways to simplify a divergence and migration model for three populations (Hey, 2010). The advantage of our likelihood method and an analogous Bayesian scheme recently developed by Yang (2010) in the context of species delimitation is that — rather than assuming a known history of divergence — they quantify the support for a set of alternative scenarios. In fact, for a minimal sampling scheme of a single haploid individual per population, evaluating all possible topologies and nested models within them is equivalent to testing all possible assignments of individuals to populations. Thus our method does not even rely on defining population limits *a priori* and so could be used to detect cryptic population structure or reproductive barriers. In practice, maximising the information contained in a single sample per population also minimizes the bias against rare and/or poorly sampled species. The potential importance of rare species when comparing population histories within communities is illustrated by our finding of a different history for *M. amaenus*. Because only a single rearing from the Middle East was available for this species, we were unable to include it in the Stone *et al.* (2012) analysis.

Lohse *et al.* (2010) previously analysed the *C. fungosa* data using the method of Yang (2002), which was originally designed to estimate species splits given a known topology. As expected, this study found almost identical parameter estimates as those obtained here under the full model (which has the highest lnL, Table 2). However, what our previous analysis was unable to reveal was that simpler models may also fit the data. *C. fungosa* stands out from the other parasitoid species analysed here in three key aspects. Firstly, it has the greatest model uncertainty despite the fact that the largest number of loci was available in this species.

Secondly, the effective size of its common ancestral population (N_0) is around 2.5 fold larger than estimates for the three other parasitoid species regardless of the model (Supporting Information Fig. S1). This is also reflected by the fact that *C. fungosa* has the highest per site diversity (θ_W) across loci (Table 1) despite its recent population divergence time. It is tempting to speculate that the larger ancestral N_e is a consequence of the greater abundance and host range of *C. fungosa*, which has been recorded in over twice as many different gall types than any of the other species (Askew, 1961b; Bailey *et al.*, 2009). However, this assumes that its lifehistory has remained unchanged at least over the last glacial cycle. While positive correlations between census size and nuclear diversity have been found across insects generally (for a recent review see Frankham, 2011), correlations of N_e and lifehistory traits remain to be explored within communities. However, this of course requires comparisons across larger sets of taxa. Finally, under the full model, estimates of T_2, the time between population divergence events and the effective population size N_1 during this interval, both converge to zero in *C. fungosa* (both in the present study and the Lohse *et al.* (2010) analysis). Lohse *et al.* (2010) showed that even when increasing the number of individuals sampled per population, these two parameters remain highly confounded. This may suggest that an important aspect of the history of *C. fungosa* is not captured by simple divergence models. For example, a strong bottleneck accompanying divergence between central and western refugia would be compatible with low and uncertain estimates of these parameters and gene flow following divergenc could have the same effect. We perfomed additional simulations to investigate how robust our inferences are to such model misspecification. Specifically we asked, given the timing of divergence inferred for *M. amaenus* (under the full model), what level of post divergence gene flow is required to erode the signal for a two population model? In other words, is it possible that some of the species inferred to have diverged more recently, actually co-diverged with *M. amaenus* but experienced gene flow following divergence? To roughly match the parameters inferred for *M. amaenus* we fixed $T_1 + T_2 = 1.5$ and $T_1 = 0.26$ (see vertical line in Fig. 4B) and simulated replicate datasets (of 18 loci)

with increasing amounts of symmetric migration between all populations (varying $M = 4Nm$, the number of migrants per generation, from 0-2). In agreement with a previous simulation study (Eckert & Carstens, 2008), our robustness analysis revealed that migration does indeed erode phylogenetic signal (Supporting Information Figure S3). Although rather high levels of postdivergence geneflow ($M > 0.5$) are required for there to be an appreciable chance of erroneously inferring a polytomous split or panmixia, we can of course not exclude the possibility of postdivergence gene flow without modelling it explicity.

In general, there is much scope for increasing the realism of model based inference and analogous expressions for the likelihood of triplet genealogies under more complex models including population size changes, migration and admixture can be derived (Lohse *et al.*, 2011a). However, because of the inherent stochasticity of the coalescent, much larger volumes of data are required to distinguish those more realistic models from simpler alternatives in practice. Whole genomes which can now be sequenced cost-effectively even in non-model organisms offer maximum replication across loci and should make it possible to accurately estimate recent divergence and pick up signatures of secondary gene flow (Green *et al.*, 2010). Likelihood analysis and model selection based on it provides an efficient way to extract information from such genomic datasets in the gallwasp community and other systems.

Acknowledgments

We thank Majide Tavakoli, Juli Pujade-Villar and Pablo-Fuentes Utrilla for contributing specimens. Mike Hickerson and three anonymous reviewers gave helpful comments on earlier versions of the manuscript. This work was supported by funding from the UK Natural Environment Research Council to KL (NE/I020288/1) and GNS (NE/H000038/1, NE/E014453/1, NER/B/504406/1, NER/B/S2003/00856).

References

Askew, R.R. (1961a). On the biology of the inhabitants of oak galls of Cynipidae (H ymenoptera) in Britain. *Transactions of the Society for British Entomology*, 14, 237âĂŞ258.

Askew, R.R. (1961b). Some biological notes on the pteromalid (Hym. Chalcidoidea) genera *Caenacis* Förster, *Cecidostiba* Thomson and *Hobbya* Delucchi, with descriptions of two new species. *Entomophaga*, 6, 58–67.

Atkinson, R., Rokas, A. & Stone, G.N. (2007). Longitudinal patterns in species richness and genetic diversity in european oaks and oak gallwasps. In S. Weiss, editor, *Phylogeography in southern European refugia*, pages 127–154. Springer, Dordrecht, The Netherlands.

Avise, J. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18, 489–522.

Bailey, R., Schönrogge, K., Cook, J.M., Melika, G., Csóka, G., Thúroczy, C. & Stone, G.N. (2009). Host niches and defensive extended phenotypes structure parasitoid wasp communities. *PLoS Biology*, 7(8), e1000179.

Barton, N.H., Kelleher, J. & Etheridge, A.M. (2010). A new model for extinction and recolonisation in two dimensions: Quantifying phylogeography. *Evolution*, 64(9), 2701–2715.

Beaumont, M.A., Nielsen, R., Robert, C., Hey, J., Gaggiotti, O., Knowles, L., Estoup, A., Panchal, M.and Corander, J., Hickerson, M., Sisson, S.A., Fagundes, N., Chiki, L., Beerli, P., Vitalis, R., Cornuet, J.M., Huellsenbeck, J., Foll, M., Yang, Z., Rousset, F., Balding, D. & Excoffier, L. (2010). In defence of model-based inference in phylogeography. *Molecular Ecology*, 19, 436–446.

Challis, R.J., Mutun, S., Nieves-Aldrey, J.L., Preuss, S., Rokas, A., Aebi, A., Sadeghi, E., Tavakoli, M. &

Stone, G.N. (2007). Longitudinal range expansion and cryptic eastern species in the western palaearctic oak gallwasp *Andricus coriarius*. *Molecular Ecology*, 16(10), 2003–2014.

DeChaine, E.G. & Martin, A.P. (2006). Using coalescent simulation to test the impact of Quaternary climate cycles on divergence in an alpine plant-insect association. *Evolution*, 60(5), 1004–1013.

Dolman, G. & Joseph, L. (2012). A species assemblage approach to comparative phylogeography of birds in southern australia. *Ecology and Evolution*, 2(2), 354–369.

Duvaux, L., Belkhir, K., Boulesteix, M. & Boursot, P. (2011). Isolation and gene flow: inferring the speciation history of european house mice. *Molecular Ecology*, 20(24), 5248–5264.

Eckert, A.J. & Carstens, B.C. (2008). Does gene flow destroy phylogenetic signal? the performance of three methods for estimating species phylogenies in the presence of gene flow. *Molecular Phylogenetics and Evolution*, 49(3), 832–842.

Edgar, R. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32(5), 1792–1797.

Edwards, S.V. & Beerli, P. (2000). Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, 54, 1839–1854.

Espíndola, A. & Alvarez, N. (2011). Comparative phylogeography in a specific and obligate pollination antagonism. *Plos One*, 6(12), e28662.

Felsenstein, J. (2006). Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? *Molecular Biology and Evolution*, 23(3), 691–700.

Frankham, R. (2011). How closely does genetic diversity in finite populations conform to predictions of neutral theory? large deficits in regions of low recombination. *Heredity*, 32(5), 1792–1797.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.Y., Hansen, N.F., Durand, E.Y., Malaspinas, A.S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prufer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Hober, B., Hoffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D. & Paabo, S. (2010). A draft sequence of the Neanderthal genome. *Science*, 328(5979), 710–722.

Hayward, A. & Stone, G.N. (2006). Comparative phylogeography across two trophic levels: the oak gall wasp *Andricus kollari* and its chalcid parasitoid *Megastigmus stigmatizans*. *Molecular Ecology*, 15(2), 479–489.

Hellenthal, G., A., A. & Falush, D. (2008). Inferring human colonisation history using a copying model. *PLoS Genetics*, 4(5), e1000078.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.

Hey, J. (2010). Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, 27, 905–920.

Hey, J. & Machado, C.A. (2003). The study of structured populations - new hope for a difficult and divided science. *Nature Reviews Genetics*, 4(7), 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(2), 747–760.

411 Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler,
412 L., Victoriano, P.F. & Yoder, A.D. (2010). Phylogeography's past, present, and future: 10 years after
413 Avise 2000. *Molecular Phylogenetics and Evolution*, 54(1), 291–301.

414 Hoberg, E. & Brooks, D. (2008). A macroevolutionary mosaic: host-switching, geographical colonization
415 and diversification in complex host-parasite systems. *J. Biog*, 35, 1533–1550.

416 Hudson, R.R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioin-*
417 *formatics*, 18, 337–338.

418 Jennings, W.B. & Edwards, S.V. (2005). Speciational history of Australian grass finches (*Poephila*) inferred
419 from thirty gene trees. *Evolution*, 59(9), 2033–2047.

420 Knowles, L.L. (2002). Statistical phylogeography. *Molecular Ecology*, 11, 2623–2635.

421 Koch, M.A., Kiefer, C. & Ehrlich, D. (2006). Three times out of Asia Minor: the phylogeography of *Arabis*
422 *alpina* l. (Brassicaceae). *Molecular Ecology*, 15, 825–839.

423 Lim, H.C. & Sheldon, F.H. (2011). Multilocus analysis of the evolutionary dynamics of rainforest bird
424 populations in Southeast Asia. *Molecular Ecology*, 20, 3414–3438.

425 Lohse, K., Harrison, R.J. & Barton, N.H. (2011a). A general method for calculating likelihoods under the
426 coalescent process. *Genetics*, 58(189), 977–987.

427 Lohse, K., Sharanowski, B., Nicholls, J.A., Blaxter, M. & Stone, G.N. (2011b). Developing EPIC markers
428 for chalcidoid hymenoptera from EST and genomic data. *Molecular Ecology Resources*, 3(11), 521–529.

429 Lohse, K., Sharanowski, B. & Stone, G.N. (2010). Quantifying the population history of the oak gall
430 parasitoid *C. fungosa*. *Evolution*, 58(4), 439–442.

Lopez-Vaamonde, C., Rasplus, Y.J., Weiblen, G. & Cook, J.M. (2001). Molecular phylogenies of fig wasps: partial co-cladogenesis of pollinators and parasites. *Molecular Phylogenetics and Evolution*, 21, 55–71.

Nicholls, J.A., Preuss, S., Hayward, A., Melika, G., Csóka, G., Nieves-Aldrey, J.L., Askew, R.R., Tavakoli, M., Schönrogge, K. & Stone, G.N. (2010). Concordant phylogeography and cryptic speciation in two western Palaearctic oak gall parasitoid species complexes. *Molecular Ecology*, 19, 592–609.

Nichols, R. (2001). Gene trees and species trees are not the same. *Trends in Ecology & Evolution*, 16(7), 358–364.

Oliveira, D.C.S.G., Raychoudhury, R., Lavrov, D.V. & Werren, J.H. (2008). Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Molecular Biology and Evolution*, 25(10), 2167–2180.

Patterson, N., Richter, D.J., Gnerre, S., Lander, E.S. & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, 441(7097), 1103–1108.

Pulquério, M. & Nicholls, R.A. (2007). Dates from the molecular clock: how wrong can we be? *Trends in Ecology & Evolution*, 22(4).

Rannala, B. & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4), 1645–1656.

Rokas, A., Atkinson, R.J., Webster, L., Csóka, G. & Stone, G.N. (2003). Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-mediterranean oak gallwasp *Andricus quercustozae*. *Molecular Ecology*, 12(8), 2153–2174.

Schmitt, T. (2007). Molecular biogeography of europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, 4(11), doi:10.1186/1742–9994–4–11.

452 Smith, C., Tank, S., Godsoe, W., Levenick, J., Strand, E., Esque, T. & Pellmyr, O. (2011). Comparative
453 phylogeography of a coevolved community: concerted population expansions in joshua trees and four
454 yucca moths. *PloS One*, 6(10), e25628.

455 Stone, G.N., Challis, R.J., Atkinson, R.J., Csóka, G., Hayward, A., Melika, G., Mutun, S., Preuss, S., Rokas,
456 A., Sadeghi, E. & Schönrogge, K. (2007). The phylogeographical clade trade: tracing the impact of
457 human-mediated dispersal on the colonization of northern Europe by the oak gallwasp *Andricus kollari*.
458 *Molecular Ecology*, 16, 2768–2781.

459 Stone, G.N., Lohse, K., Nicholls, J.A., Fuentes-Utrilla, P., Sinclair, F., Schönrogge, K., Csóka, G., Melika,
460 G., Nieves-Aldrey, J.L., Pujade-Villar, J., Tavakoli, M., Askew, R.R. & Hickerson, M.J. (2012). Recon-
461 structing community assembly in time and space reveals enemy escape in a western palaearctic insect
462 community. *Current Biology*, *in press*.

463 Takahata, N., Satta, Y. & Klein, J. (1995). Divergence time and population size in the lineage leading to
464 modern humans. *Theoretical Population Biology*, 48, 198–221.

465 Wakeley, J. (2009). *Coalescent theory*. Roberts & Company Publishers, Greenwood Village, Colorado.

466 Wolfram Research, I. (2010). *Mathematica, Version 8.0*. Wolfram Research, Inc., Champaign, Illinois.

467 Wright, S. (1943). Isolation by distance. *Genetics*, 28(2), 114–138.

468 Yang, Z. (2002). Likelihood and Bayes estimation of ancestral population sizes in hominoids using data
469 from multiple loci. *Genetics*, 162(4), 1811–1823.

470 Yang, Z. (2010). A likelihood ratio test of speciation with gene flow using genomic data. *Genome Biology*
471 *and Evolution*, 2, 200–211.

Appendix

Assuming the full divergence model described above (Methods) and a sample of three sequences a, b and c (the labelling corresponds to the sampled population), we can write down an expression for the generating function (GF) of the vector of branch lengths $\underline{t} = (t_a, t_b, t_c, t_{ab}, t_{ac}, t_{bc})$. Using the recursion of Lohse *et al.* (2011a, eq. 5 and 12) it is simplest to initially assume a sligthlyu different model where population divergence times are exponentially distributed with rates Λ_1 and Λ_2. The GF under this model is defined as $\psi[a/b/c] = E[e^{-\underline{t}.\underline{\omega}}]$ where $\underline{\omega} = (\omega_a, \omega_b, \omega_c, \omega_{ab}, \omega_{ac}, \omega_{bc})$ is a vector of dummy variables corresponding to the branch lengths $\underline{t}$ and is given by the following set of equations:

$$\begin{aligned}
\psi[a/b/c] &= \frac{1}{\Lambda_1 + \omega_a + \omega_b + \omega_c}\Lambda_1\psi[a/b,c] \\
\psi[a/b,c] &= \frac{1}{\alpha\beta + \Lambda_2 + \omega_a + \omega_b + \omega_c}\left(\Lambda_2\psi[a,b,c] + \alpha\psi[a/\{b,c\}]\right) \\
\psi[a,b,c] &= \frac{1}{3\beta + \omega_a + \omega_b + \omega_c}\left(\frac{1}{\beta + \omega_a + \omega_{ab}} + \frac{1}{\beta + \omega_b + \omega_{ac}} + \frac{1}{\beta + \omega_c + \omega_{bc}}\right) \\
\psi[a/\{b,c\}] &= \frac{\Lambda_2}{(\Lambda_2 + \omega_a + \omega_{bc})\,(1 + \omega_a + \omega_{bc})}
\end{aligned} \tag{1}$$

β is an inheritance scalar (1 for diploids and $4/3$ for haplodiploids as in the analysis above) and $\alpha = \frac{N_0}{N_1}$. This has solution:

$$\psi[a/b/c] = \frac{\Lambda_1\Lambda_2\left(\frac{2\beta+\omega_b+\omega_c+\omega_{ab}+\omega_{ac}}{(\beta+\omega_c+\omega_{ab})(\beta+\omega_b+\omega_{ac})} + \frac{3\alpha\beta+\Lambda_2+(1+\alpha)\omega_a+\alpha\omega_b+\alpha\omega_c+\omega_{bc}}{(\beta+\omega_a+\omega_{bc})(\Lambda_2+\omega_a+\omega_{bc})}\right)}{(3\beta + \omega_a + \omega_b + \omega_c)\,(\Lambda_1 + \omega_a + \omega_b + \omega_c)\,(\alpha\beta + \Lambda_2 + \omega_a + \omega_b + \omega_c)} \tag{2}$$

We denote the GF for the case of interest, i. e. divergence at discrete times T_1 and $T_1 + T_2$ as $P[\underline{\omega}]$. Because $\psi[a/b/c] = \int \Lambda_1\Lambda_2 P[\underline{\omega}]e^{-\underline{\Lambda}.\underline{T}}d\underline{T}$, $P[\underline{\omega}]$ is given by dividing (2) by Λ_1 and Λ_2 and inverting with respect to Λ_1 and Λ_2. The expression can be obtained using the *InverseLaplaceTransform* function in *Mathematica* but is cumbersome (see Supporting Information, nb.file). However, a drastic simplification is

achieved if we condition on a particular topology of the genealogy by taking the limit with respect to those ω that are incompatible with that topology (see Lohse *et al.*, 2011a). A further simplification arises from the symmetries in branch lengths. For a given topology, $P[\underline{\omega}]$ only depends on the interval between successive coalescence events. For example, for topology $\{\{b,c\},a\}$, $t_b = t_c = t_3$, $t_{bc} = t_2$ and $t_a = t_3 + t_2$. in other words, t_2 and t_3 are the time intervals during which there are two and three lineages respectively. Defining the corresponding dummy variables ω_2 and ω_3, the GF for a genealogy congruent with the order of population divergence is:

$$P[\omega_2,\omega_3|G_{bc},T_1,T_2,\alpha] = \lim_{\substack{\omega_{ab}\to\infty \\ \omega_{ac}\to\infty}} P[\underline{\omega}| = \frac{e^{-\omega_2 T_1}\left(\frac{e^{-\omega_2 T_2}(-3\alpha\beta-\alpha\omega_3)}{-\alpha\beta+\omega_2-\omega_3} + \frac{e^{(-\alpha\beta-\omega_3)T_2}(2\alpha\beta+\omega_2-\omega_3+\alpha\omega_3)}{-\alpha\beta+\omega_2-\omega_3}\right)}{(\beta+\omega_2)(3\beta+\omega_3)} \tag{3}$$

where G_{bc} is a shorthand notation for a congruent topology $\{\{b,c\},a\}$.

Similarly, the GF for an incongruent (either with branch t_{ab} or t_{ac}) genealogy is:

$$P[\omega_2,\omega_3|G_{ac},T_1,T_2,\alpha] = \lim_{\substack{\omega_{ab}\to\infty \\ \omega_{bc}\to\infty}} P[\underline{\omega}| = \frac{e^{-\omega_3 T_1-(\alpha\beta+\omega_3)T_2}}{(\beta+\omega_2)(3\beta+\omega_3)} \tag{4}$$

Note that if we set all ω to zero (and assume $\beta = 1$), 2 goes to 1 and 3 and 4 above reduce to the well-known result of Takahata *et al.* (1995) for topological probabilities, i. e. $1 - \frac{2}{3}e^{-\alpha T_2}$ and $\frac{1}{3}e^{-\alpha T_2}$ for congruent and incongruent genealogies respectively.

Assuming that mutations in interval t_2 and t_3 are Poisson distributed with rates $2\theta/2$ and $3\theta/2$ respectively, where the per locus mutation rate is $\theta/2 = 2N_0\mu$, the joint probability of observing k_2 and k_3 mutations can be obtained by taking successive derivatives of (3) and (4) with respect to ω_2 and ω_3 (eq. 1 Lohse *et al.*, 2011a):

$$p(k_2, k_3|G_i, T_1, T_2, \alpha) = (-1)^{k_2+k3} \frac{\theta^{k_2}(3\theta/2)^{k_3}}{k_2!k_3!} \left(\frac{\partial^{k_2+k_3} P[\omega_2, \omega_3|G_i, T_1, T_2, \alpha]}{\partial {\omega_2}^{k_2} {\omega_3}^{k_3}} \right)_{\substack{\omega_2=\theta \\ \omega_3=3\theta/2}} \quad (5)$$

For a known triplet topology G_i, there are only four possible branches and the corresponding mutations can be classed into three types, those on the internal branch, k_i, those on the two shorter external branches k_{eS} and those on the longer external branch k_{eL}. Their joint probability $p(k_i, k_{eS}, k_{eL})$ can be found from (5) by summing over all possible ways these can be partitioned amongst the two coalescent intervals (Lohse *et al.*, 2011a, Supporting Information):

$$p(k_i, k_{e1}, k_{e2}|G_i, T_1, T_2, \alpha) = \sum_{j=0}^{k_{e2}} \binom{k_{e1} + k_{e2} - j}{k_{e2} - j} \frac{1}{3}^{k_{e2}-j} \frac{2}{3}^{k_{e1}} \binom{k_i + j}{j} \frac{1}{2}^{k_i+j} p(k_i + j, k_{e1} + k_{e2} - j|G_i, T_1, T_2, \alpha) \quad (6)$$

where the last term corresponds to (5).

Loci with no topologically informative mutations (i. e. $k_i = 0$) constitute a separate class G_0. Finding the probability of mutational configurations for this case involves summing over the contributions from the three topology classes. Analogous to 6, these are weighted by the binomial probabilities of distributing the keS mutations onto the two shorter external branches (with k_eS1 and k_eS2 mutations on each).

$$p(k_a, k_b, k_c|G_0, T_1, T_2, \alpha) = \sum_i \frac{1}{2} \binom{k_{eS1} + k_{eS2}}{k_{eS1}} p(0, k_{eS}, k_{eL}|G_i, T_1, T_2, \alpha) \quad (7)$$

Table 1: Length (excluding indels) of the alignment with the outgroup, number of polymorphic sites (S) and topologically informative mutations (those on the internal branches, k_i) in triplet for 18 nuclear loci. The topology of the triplet genealogy at each locus is denoted according to which sample is basal (east = E, center = C, west = W, no topologically informative sites = 0) and given in brackets. The bottom row gives the mean θ_W per site across loci. *indicates alignments that were trimmed to exclude likely recombinant portions.

	C. fungosa			*C. semifascia*			*H. stenonota*			*M. amaenus*		
Locus	length	**S**	k_i	length	**S**	top	length	**S**	top	length	**S**	top
AntSesB	606	2	1 (E)							563	3	
nAcRbeta	748	0	0				234	0	0			
RACK	560	3	0	561	1	0				738*	6	2 (E)
ran	499	2	0	472	1	0	476	2	0	447	3	1 (E)
RpL10ab	955	3	1 (E)							966	9	1 (E)
RpL13a	446*	14	4 (E)	776	5	1 (C)						
RpL15	618	2	0							608	6	3 (E)
RpL27	501	14	6 (E)				508	2	0	518	2	2 (E)
RpL37a	220	0	0	220	0	0	220	2	0	218	0	0
RpL37	866	20	1 (W)	666	0	0	679	3	0	370*	9	2 (W)
RpL39	463	0	0				467	2	1 (C)	545	5	1 (E)
RpS15	739	28	7 (C)									
RpS18	813	6	1 (E)	768	2	2 (E)						
RpS23	268	6	3 (E)	268	0	0	267	2	0	268	1	1 (E)
RpS4	754	1	0	250*	5	1 (W)	705	3	1 (C)	531*	4	1 (C)
RpS8	422	5	1 (E)	470	1	0	468	4	1 (E)	452	1	0
sansfille	446	2	1 (C)	433	1	0				434	2	0
Tctp	493	3	0	465	2	0	477	3	1 (C)	389*	6	1 (E)
Mean θ_W		0.0160			0.0050			0.0076			0.0123	

Table 2: lnL and of all models nested within the full divergence model of three populations with topology (E, (C, W)) (Fig. 1a) for four parasitoid species. The 2nd column gives the number of model parameters (k). The model with the highest lnL in each species is shown in bold, the simplest model retained in likelihood ratio tests of nested models is indicated by *. Models with alternative order of population divergence had no support.

Model	k	*C. fungosa*	*C. semifascia*	*H. stenonota*	*M. amaenus*
panmixia	1	-122.82*	-44.97*	-49.15	-86.92
polytomy	2	-122.59	-44.67	**-46.71***	-84.98
2 pop.	3	-120.77	**-44.34**	-46.71	-79.01*
full model	4	**-120.01**	-44.34	-46.71	**-78.90**
C & W topologies	3	polytomy	polytomy	polytomy	polytomy

Table 3: Maximum likelihood estimates of scaled divergence times and ancestral population sizes θ for the model retained in the LRT and all models with a higher lnL (see Table 2) for four parasitoid species. For ease of comparison between models, the time of the oldest population split is given in each case and —for the full model only— the time inbetween population splits T_2. Corresponding absolute values of N_e and τ are shown in brackets.

Model	θ_1 (N_0)	θ_2 (N_1)	T_2 (τ_2)	oldest T (τ)
C. fungosa				
panmixia	5.70 (7.84×10^5)			
polytomy	5.25 (7.23×10^5)			0.046 (33 KY)
two-pop.	5.09 (7.00×10^5)	2.76 (3.79×10^5)		0.158 (111 KY)
full model	5.26 (7.19×10^5)	$\rightarrow 0$	$\rightarrow 0$	0.182 (131 KY)
C. semifascia				
panmixia	1.87 (2.57×10^5)			
polytomy	1.46 (2.01×10^5)			0.177 (35.6 KY)
two-pop.	1.35 (1.85×10^5)	2.71 (3.73×10^5)		0.322 (59.7 KY)
H. stenonota				
polytomy	1.20 (1.65×10^5)			0.755 (125 KY)
M. amaenus				
two-pop.	1.58 (2.17×10^5)	3.45 (4.57×10^5)		1.58 (343 KY)
Full	1.67 (2.30×10^5)	2.79 (3.21×10^5)	1.20 (277 KY)	1.46 (335 KY)

Figure 1: The full divergence model between three populations with a population tree topology (E,(W, C)) (a) can be further simplified by setting either interval T_1 or T_2 or both to zero resulting in three nested models; (b) divergence between two populations (with C and W merged into a single population), (c) a polytomous split of the common ancestral population and (d) a single panmictic population.

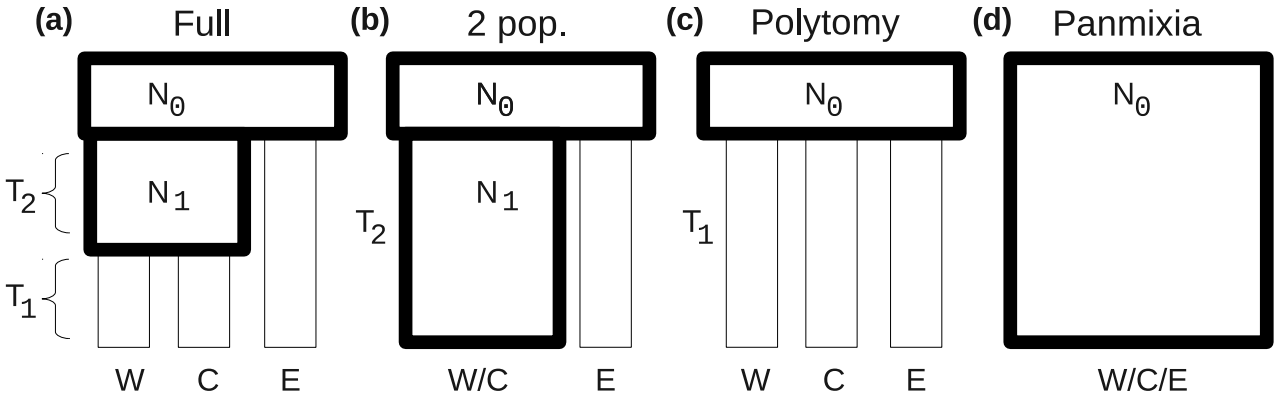

Figure 2: Assuming infinite site mutations and an outgroup, each polymorphic site can be placed onto a unique branch in the underlying genealogy unambiguously. For example, there are 6 polymorphic sites in *RpS18* in *C. fungosa*. These can be classed into 3 types according to the genealogical branch they fall on (0 denotes the ancestral, 1 the derived state relative to the outgroup *C. lauta*). In *RpS18* a single shared derived mutation, i. e. parsimony informative site (white dot), defines the topology (E,(C,W)).

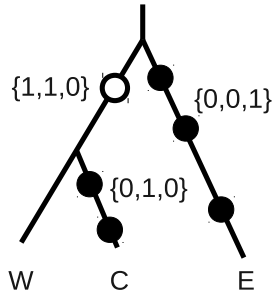

Figure 3: ΔlnL plots for divergence times (in KY) between refugial populations for four oak gall parasitoid species. In each species, plots for the divergence times under the most parsimonious model as determined by LRT and all models with a higher lnL are shown. Full model = thick dashed lines, two-pop. = thin dashed lines and polytomy = solid lines. The horizontal line delimits the region of 95 % confidence. Note that there are two curves for the full model one for each divergence time (T_1 and $T_2 + T_2$). However, because in *C. fungosa* the MLE for T_2 converges to zero, the lnL curves are near identical and appear as one.

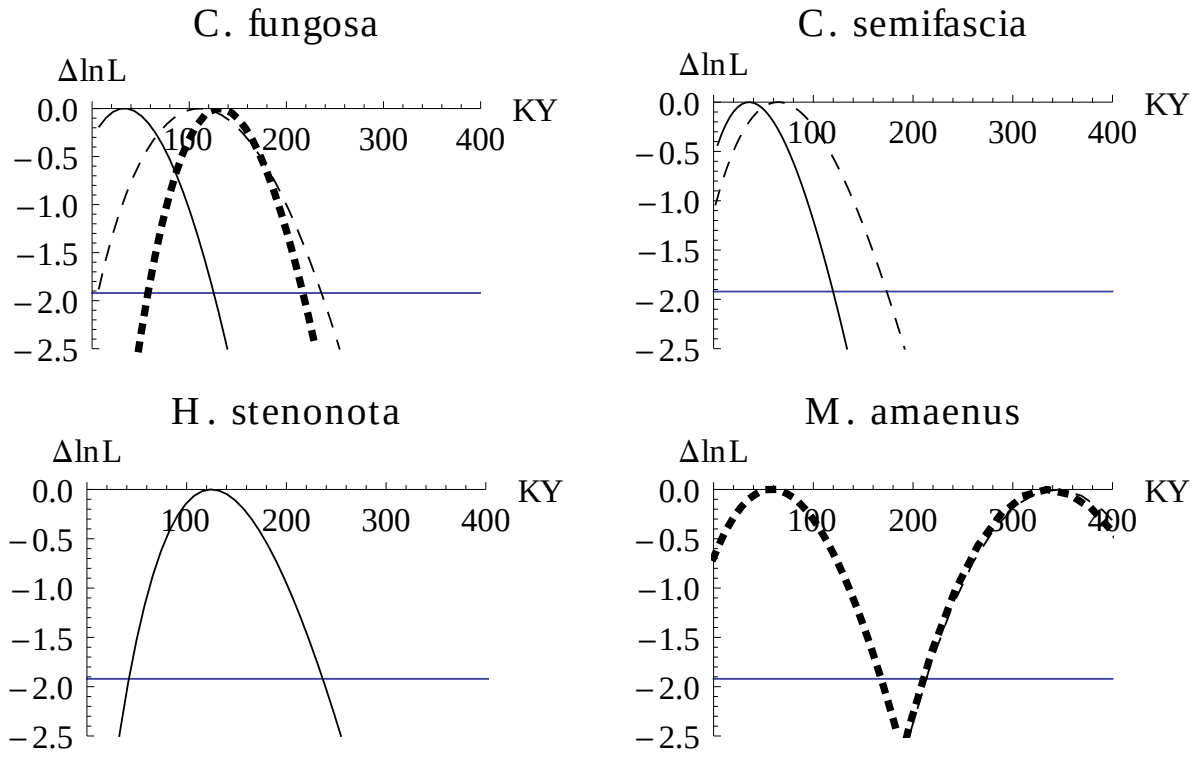

Figure 4: The power to distinguish between alternative models of population divergence plotted against T_1 the time of the more recent split. Each point shows the proportion of replicates (out of 100) for which a particular model was retained using LRT. Points were joined for ease of comparison with the same labelling as in Fig. 3, i. e. full model = thick dashed, two-pop. = thin dashed, polytomy = solid lines and panmixia = dotted lines. Panels in the top row (A–C) correspond to old $T_1 + T_2 = 1.5$, those in the bottom row (D–F) to recent $T_1 + T_2 = 0.5$ divergence histories. Power was determined from simulated datsets for varying numbers of loci: 10 (A, D), 18 (B, E) and 100 (C, F). The MLE estimnate for T_1 inferred for *M. amaenus* under the full model is shown in B) as a vertical line