Lifecycle closure, lineage sorting, and hybridization revealed in a phylogenetic analysis of European oak gallwasps (Hymenoptera: Cynipidae: Cynipini) using mitochondrial sequence data

Antonis Rokas,[a,b,*] George Melika,[c] Yoshihisa Abe,[d] Jose-Luis Nieves-Aldrey,[e] James M. Cook,[f] and Graham N. Stone[a]

[a] *Institute of Cell, Animal and Population Biology, Ashworth Laboratories, West Mains Road, King's Buildings, University of Edinburgh, Edinburgh EH9 3JT, UK*
[b] *Howard Hughes Medical Institute and Laboratory of Molecular Biology, R.M. Bock Labs, University of Wisconsin-Madison, 1525 Linden Drive, Madison, WI 53706-1596, USA*
[c] *Systematic Parasitoid Laboratory, Central Plant Protection and Soil Conservation Service, Kelcz-Adelffy St. 6, Köszeg 9730, Hungary*
[d] *Kyoto Prefectural University, Faculty of Agriculture, Lab of Applied Entomology, Kyoto 606-8522, Japan*
[e] *Museo Nacional de Ciencias Naturales, Departamento de Biodiversidad y Biología Evolutiva, José Gutiérrez Abascal, Madrid 28006, Spain*
[f] *Department of Biological Sciences, Imperial College, Silwood Park, Ascot, Berkshire SL5 7PY, UK*

Received 2 January 2002; received in revised form 3 August 2002

Abstract

Oak gallwasps are cyclically parthenogenetic insects that induce a wide diversity of highly complex species- and generation-specific galls on oaks and other Fagaceae. Phylogenetic relationships within oak gallwasps remain to be established, while sexual and parthenogenetic generations of many species remain unpaired. Previous work on oak gallwasps has revealed substantial intra-specific variation, particularly between regions known to represent discrete Pleistocene glacial refuges. Here we use statistical phylogenetic inference methods on sequence data for a fragment of the mitochondrial cytochrome *b* gene to reconstruct the relationships among 62 oak gallwasp species. For 16 of these we also include 23 additional cytochrome *b* haplotype sequences from different Pleistocene refuge areas to test the effect of intra-specific variation on inter-specific phylogeny reconstruction. The reconstructed phylogenies show good intra-generic resolution and identify several conserved clades, but fail to reconstruct either very recent or very ancient divergences. Nine of the 16 species represented by multiple haplotypes are not monophyletic. The apparent discordance between the recovered gene tree and the current taxonomic classification can be explained through: (a) collapsing of some species currently known only from either a sexual or a parthenogenetic generation into a single cyclically parthenogenetic entity; (b) sorting of ancestral polymorphism in diverging lineages, and (c) horizontal transfer of haplotypes, perhaps due to hybridization within glacial refuges. Our conclusions emphasise the need for careful intra-specific sampling when reconstructing phylogenies for radiations of closely related species and imply that for certain taxonomic groups full phylogenetic resolution (using molecular markers) may not be attainable.
© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Gallwasp; Cynipidae; Cynipini; Lineage sorting; Lifecycle; Likelihood; Bayesian inference; Cytochrome *b*; Hybridization; Insect systematics

1. Introduction

Gallwasps (Hymenoptera: Cynipidae) are members of the Cynipoidea, a major lineage of predominantly parasitoid wasps within the Hymenoptera (Ronquist, 1999). Gallwasps are obligate parasites of plants, and either induce their own galls in plant tissues or develop as inquilines within the galls induced by other gallwasps (Askew, 1984; Stone et al., 2002). Each species induces a highly characteristic and often complex gall, leading to an incredible diversity of gall structures within the

* Corresponding author. Fax: 1-608-262-9343.
E-mail address: arokas@facstaff.wisc.edu (A. Rokas).

whole group (Stone and Cook, 1998). Cynipidae contains about 1360 described species divided into six tribes (Liljeblad and Ronquist, 1998; Nieves-Aldrey, 2001). Biogeographic and fossil evidence suggest that the Cynipidae originated at least as long ago as the mid Cretaceous (83 mya) (Ronquist, 1999).

Oak gallwasps (tribe Cynipini) comprise the most species-rich tribe of gallwasps, with around 1000 known species in 40 genera world-wide, predominantly in the Northern Hemisphere (Nieves-Aldrey, 2001; Ronquist, 1999). The major lineages of oak gallwasps are thought to have diverged in central America (Kinsey, 1936) alongside their oak hosts (Manos et al., 1999), and the greatest richness of oak gallwasps is found in the Nearctic with an estimated 700 species in 29 genera (Weld, 1957, 1959, 1960). The Palaearctic fauna is less species-rich (ca. 163 species in 10 genera) (Askew, 1984; Nieves-Aldrey, 1987; Nieves-Aldrey, 2001; Stone et al., 2002), and the majority of European oak gallwasps are contained in just four genera—*Andricus* (116 species), *Plagiotrochus* (14), *Neuroterus* (13), and *Cynips* (7) (Askew, 1984; Nieves-Aldrey, 2001; Stone et al., 2002).

Despite recent advances in Cynipini phylogenetics (Cook et al., 2002; Liljeblad and Ronquist, 1998; Rokas, 2001; Rokas et al., 2002; Ronquist, 1994; Stone and Cook, 1998) little is known about relationships within the tribe. The Cynipini are cyclically parthenogenetic, with a lifecycle involving obligate alternation between a spring sexual generation and a summer/autumn parthenogenetic generation (Stone et al., 2002). The two generations of all oak galling cynipids differ substantially in size, and in some cases in morphological traits used in the past to define genera. Typically the sexual generation wasps are much smaller (mass 1–2 mg) and shorter lived (a week or less) than the parthenogenetic generation wasps (mass ca. 10 mg, lifespan up to 1 month). Early gall taxonomy was also based extensively on gall structure, and the enormous differences in gall traits between the two generations of almost all oak cynipids hampered the linking of generations. As a consequence, historically the two generations of a single species have been allocated to different genera causing great taxonomic confusion. The problem is particularly severe in the Nearctic (e.g., Drown and Brown, 1998; Lyon, 1996; Melika and Abrahamson, 2000), whereas the taxonomy of Western Palaearctic genera and species is less problematic (e.g., Melika et al., 2000; Nieves-Aldrey, 2001).

Previously (Rokas et al., 2002) we assessed the utility of eight DNA sequence markers in reconstructing phylogenetic relationships at various levels of divergence within the Cynipidae. Two mitochondrial loci (cytochrome *b* and cytochrome oxidase I) were shown to be fast-evolving and potentially useful for lower-level phylogenetics of gallwasp taxa. We have also shown that, for a number of species, regions of Southern Europe and Asia Minor corresponding to discrete Pleistocene glacial refuges possess refuge-specific haplotypes (Rokas, 2001; Rokas et al., 2001; Stone et al., 2001). In addition, it has been demonstrated experimentally that closely related species can form viable hybrids (Folliot, 1964), and it is known that horizontal transfer of haplotypes by introgression between closely related species occurs in a range of insects (e.g., Powell, 1983; Sota et al., 2001; Thelwell et al., 2000). Taken together, these facts raise the possibility that refuge-specific haplotypes might be shared within a closely related group of hybridizing taxa, such that members of several species within the same refuge could be more similar in haplotype sequence than conspecifics from different refuges.

Here we use the cytochrome *b* sequence data from 62 species to analyze the impact of incorporating intra-specific haplotype variation in resolving relationships among oak gallwasp species. In particular, we assess monophyly of 16 geographically widespread species by including additional haplotypes, resulting in an extended dataset of 85 sequences. The haplotypes used are derived from genetically divergent Western Palaearctic refuges (Atkinson, 2000; Rokas, 2001) for sets of *Andricus* species shown in previous analyses to be close relatives (Cook et al., 2002; Stone and Cook, 1998), in an attempt to detect lineage sorting events or intra-refugial hybridization. Furthermore, we include sequences for 28 *Andricus* species currently known only from a single generation. Combination of one species of each type within a single clade with high bootstrap support would provide strong indirect evidence for their representing two halves of a single cyclically parthenogenetic lifecycle.

2. Materials and methods

2.1. Study species

The dataset analysed in this study is composed of 62 oak gallwasp species. Fifty-seven of these are from 9 genera of European gallwasps, representing more than a third of the total number of species known to exist in Europe. From these species, 28 are currently known only from a single generation—either sexual (7 species) or parthenogenetic (21 species). The other five species are Japanese (*Andricus kashiwaphilus*, *Andricus mukaigawae*, *Andricus symbioticus*, and *Trichagalma serratae*) or American (*Andricus spectabilis*). The dataset also includes 23 additional haplotypes from 16 species (10 *Andricus* species, 4 *Cynips* species, 1 *Neuroterus* species and 1 *Biorhiza* species) sampled from distinct Pleistocene ice age refuges in Europe and Turkey. Species nomenclature follows Melika et al. (2000). No outgroup was used since the other gallwasp tribes are, according to the molecular marker employed in this study, only very distantly related to Cynipini (Rokas et al., 2002).

2.2. Molecular methods

DNA was extracted by using the DNeasy Tissue kit (QIAGEN cat. 69504). A 433 bp fragment of the mitochondrial cytochrome *b* gene was amplified by PCR using the CB1/CB2 primer combination of Jermiin and Crozier (1994) (alias CB-J-10933 and CB-N-11367 of Simon et al., 1994) and previously described conditions (Rokas et al., 2001; Stone and Cook, 1998). Sequencing was carried out using Perkin–Elmer BigDye Terminator chemistry on an ABI 377 sequencer. Both strands were sequenced to minimise PCR artefacts, ambiguities and base-calling errors. Chromatogram output was checked by eye. All sequences are deposited in GenBank (AJ131065–AJ131070, AJ228448–AJ228461, AJ228463–AJ228472, AJ228474–AJ228479, AJ228481, AF339625, AF395138, AF539551–AF539591, AF539795).

Four gallwasp species (*Andricus aries*, *Andricus theophrastii*, *Andricus tomentosus*, and *Neuroterus lanuginosus*) failed direct sequencing, due to the production of multiple cytochrome *b*-like PCR products. To resolve the problem, a strategy of cloning was adopted. PCR products were cloned in a blunt-ended vector following the manufacturer's instructions (Zero Blunt TOPO Cloning kit, Invitrogen, cat. K4500-01). For each of the four species, up to 10 colonies were selected, cultured and purified (QIAprep spin miniprep kit, QIAGEN cat. 27104) for subsequent sequencing.

2.3. Phylogenetic analysis

Maximum likelihood, *ML*. Phylogenies were estimated by maximum likelihood (ML) analysis of the sequence data using PAUP* versions 4.0b8-10 (Swofford, 2002). The best-fit ML model for each dataset was identified using likelihood ratio tests (Huelsenbeck and Rannala, 1997) as implemented in Modeltest 3.06 (Posada and Crandall, 1998). Parameters allowed to vary in model-fitting were base composition, substitution rates and rate heterogeneity among sites. The best-fit model incorporated unequal base frequencies, one rate for transitions and four different rates for transversions as well as among site variation (proportion of invariable sites = 0.29, shape parameter α of the γ distribution = 0.37). The parameter values suggested by Modeltest were used to search for the ML topology, using a heuristic search with TBR swapping and 10 random-taxon-addition replications. Bootstrap values were generated using the same parameter values and the nearest-neighbour-interchange (NNI) search algorithm on 500 pseudoreplicate datasets.

Bayesian inference. Phylogenies were also estimated using the program MRBAYES, version 2.0 (Huelsenbeck and Ronquist, 2001). In MRBAYES, sampling of trees from the posterior probability distribution is achieved by implementation of the Metropolis-coupled Markov chain Monte Carlo algorithm—$(MC)^3$ for short. The $(MC)^3$ algorithm allows running of multiple Markov chains. A run with four chains was performed for 2,000,000 generations, under a general time-reversible model (all six types of substitution occurring at different rates) with parameter value estimation for base frequencies, substitution matrix values and rate heterogeneity; rate heterogeneity was estimated both by using a γ distribution for the variable sites as well as by assuming a certain portion of sites to be invariable. The burn-in time was 100,000 generations.

The alignment and topologies reported in this study are available electronically from TreeBASE (http://www.herbaria.harvard.edu/treebase/, TreeBASE Study Accession No. S793).

3. Results

3.1. Multiple cytochrome b-like fragments in four gallwasp species

Sequencing of multiple clones for cytochrome *b* from the four gallwasp species (*A. aries*, *A. theophrastii*, *A. tomentosus*, and *N. lanuginosus*) that failed direct sequencing revealed multiple cytochrome *b*-like fragments. Out of the seven *A. aries* clones, one had the correct length (433 bp) and reading frame (no stop codons/indels), whereas the other six clones had a 1 bp indel (434 bp) and differed by 3.7% (uncorrected *p* distance) from the first clone. In *A. theophrastii*, five clones had the correct length and reading frame, whereas the other five although of correct length had a stop codon. The two clone groups differed by 10.2%. In *A. tomentosus*, five distinct sequences were found. Only one clone had the correct length and a full open reading frame. Five clones had the correct length but translation revealed the presence of a stop codon. One clone was 2 bp longer (435 bp), whereas the last two clones were shorter (375 and 376 bp, respectively). The lowest divergence observed between these five sequences was 2.6%. In *N. lanuginosus*, all four clones recovered a single sequence of correct length and reading frame. For the phylogenetic analyses presented in this study, only the haplotype with the correct length and reading frame from each species was used.

3.2. Phylogenetic relationships among European oak gallwasps and the effect of intra-specific variation on phylogenetic patterns

Phylogenetic reconstructions using maximum likelihood and Bayesian inference, and their respective bootstrap/posterior probability branch support values, are shown in Figs. 1 and 2, respectively. The most striking result is that 10 out of the 16 species for which
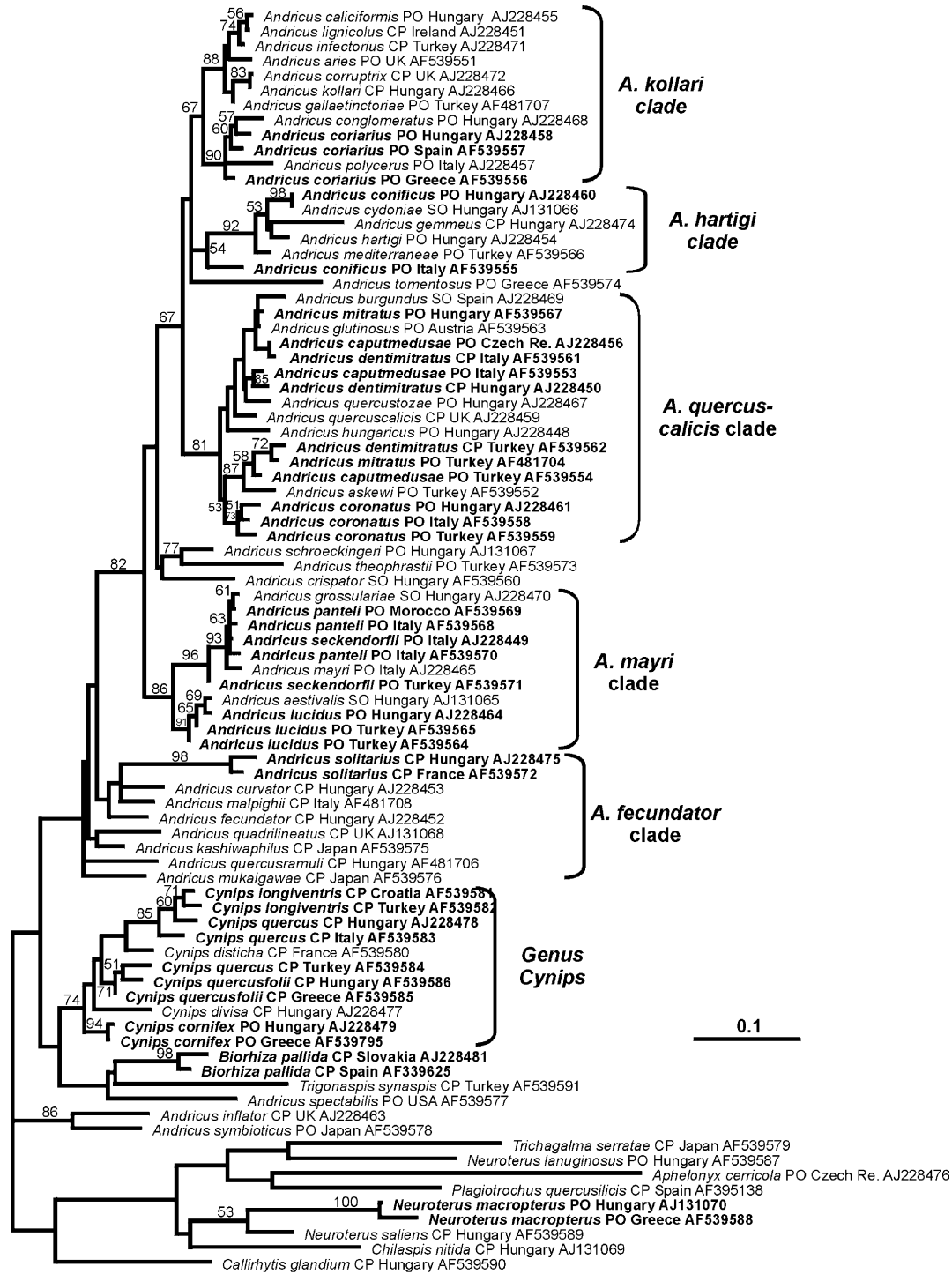
Fig. 1. Phylogram representing the ML topology found using a heuristic search algorithm. The root of the tree is arbitrary and has been chosen for best visualisation. Bootstrap values from 500 pseudoreplicates are superimposed. Branches without bootstrap support values are supported by less than 50% of replicates. Species for which multiple haplotypes have been included are shown in boldface. After the species name, the generation(s) from which the species is known (SO, sexual only; PO, parthenogenetic only; CP, cyclical parthenogenetic), the country of origin and the GenBank accession number are listed.

more than one haplotype has been included are not monophyletic, thus generating incongruence between the cytochrome *b* gene tree and the current taxonomy. To provide further support to our findings regarding non-monophyly of these species, we calculated the posterior probability of each of the 16 species being monophyletic using the dataset generated from the Bayesian analysis (Table 1). Assuming a cut-off *p*-value
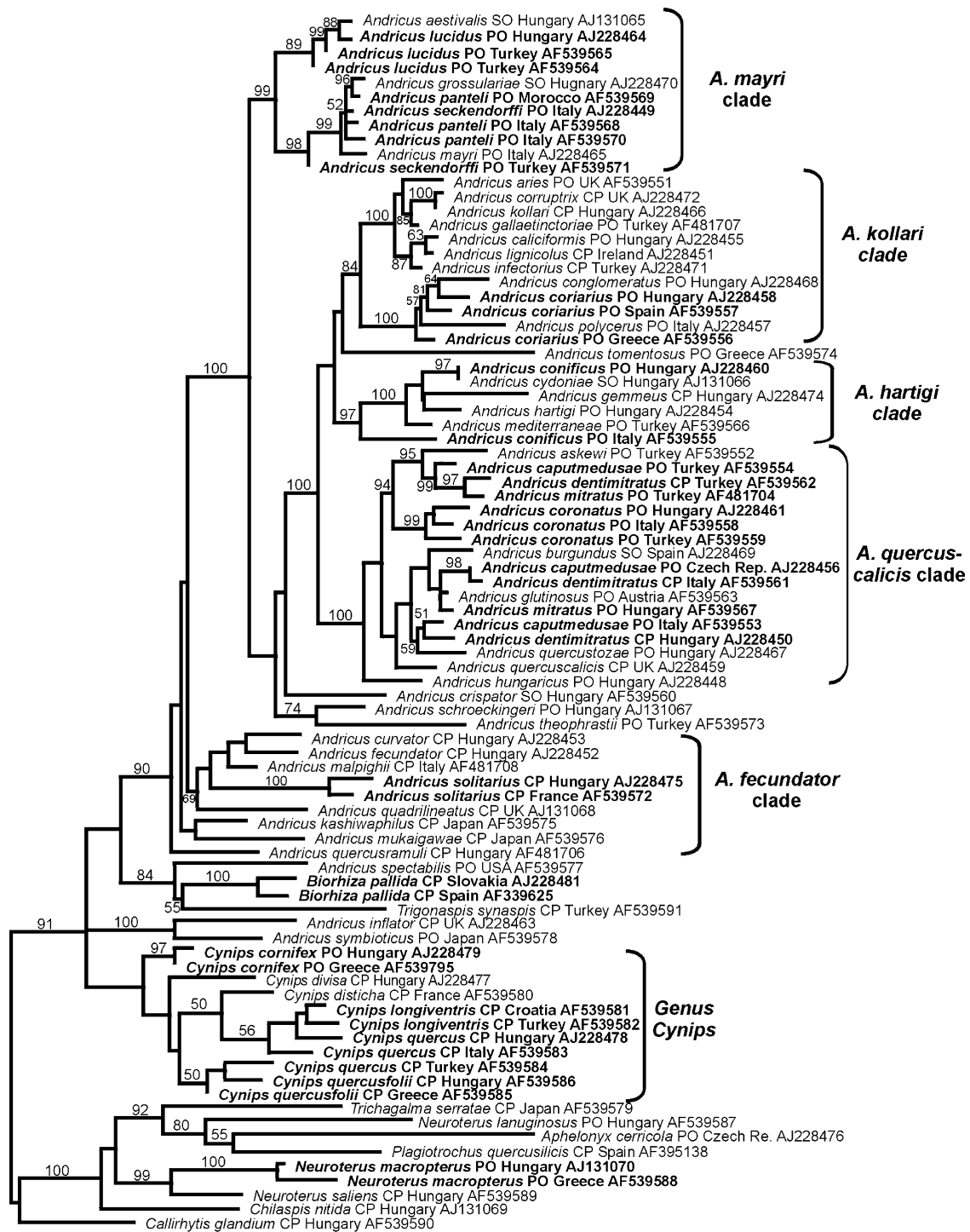
Fig. 2. Fifty percent majority-rule consensus phylogram from the Bayesian analysis. The tree is mid-point rooted. Branches without posterior probability values are supported by less than 50% of the sampled trees. Species labels and information contained there-in same as in Fig. 1.

of 0.01 (1%), monophyly is rejected for 9 of the 16 species. The only apparently non-monophyletic species for which monophyly cannot be rejected on the basis of posterior probabilities is *Cynips quercusfolii* (Table 1, Figs. 1 and 2).

Non-monophyly is particularly common within *Andricus*, where only two of the 10 species contributing

multiple haplotypes (*A. coronatus* and *A. solitarius*) had monophyletic haplotype lineages (Figs. 1and 2, Table 1). Haplotypes from each non-monophyletic species are always grouped within the same clade (sensu Cook et al., 2002; Stone and Cook, 1998) in the *Andricus* lineage (for example, all the *Andricus caputmedusae* haplotypes are placed within the *Andricus quercuscalicis*

Table 1
Posterior probability values for the monophyly of each of the 16 species of this study for which more than one haplotype was included in the analysis

Species	Posterior probability for species monophyly (%)	Grouping status according to Figs. 1 and 2
Andricus caputmedusae	0	Non-monophyletic
Andricus conificus	0	Non-monophyletic
Andricus coriarius	0	Non-monophyletic
Andricus coronatus	99.5	Monophyletic
Andricus dentimitratus	0	Non-monophyletic
Andricus lucidus	0.1	Non-monophyletic
Andricus mitratus	0	Non-monophyletic
Andricus panteli	0.3	Non-monophyletic
Andricus seckendorfii	0	Non-monophyletic
Andricus solitarius	100	Monophyletic
Biorhiza pallida	100	Monophyletic
Cynips cornifex	96.6	Monophyletic
Cynips longiventris	48.5	Monophyletic
Cynips quercus	0	Non-monophyletic
Cynips quercusfolii	14.8	Non-monophyletic
Neuroterus macropterus	100	Monophyletic

clade; Figs. 1 and 2). Non-monophyly is less pronounced outside *Andricus*, where 5 out of 6 species contributing multiple haplotypes are monophyletic. The only non-monophyletic species is *C. quercus* (Table 1).

In summary, two general patterns of incongruence between the cyt*b* tree and current taxonomic classification are present in Figs. 1 and 2:

(a) Non-monophyly. Haplotypes for several gallwasp species lie within a single clade that also includes haplotypes of one or more additional species. Examples include the distribution of haplotypes for *Andricus lucidus* and *Andricus panteli* in the *A. mayri* clade, and *Andricus coriarius* in the *A. kollari* clade. Possible causes of such patterns are considered in detail below.

(b) Groupings of haplotypes for closely related species from the same glacial refuge. The single clear example of this phenomenon is seen within the main *Andricus* clade for *A. caputmedusae*, *A. dentimitratus*, and *A. mitratus*. Haplotypes from these three species sampled in Turkey form a clade separated by high branch support values from haplotypes for the same three species sampled in Italy/Hungary. This pattern is exactly as expected for hybridization among species within shared, genetically discrete refuges.

4. Discussion

The phylogenies reconstructed from both methods show good resolution of lineages at intermediate taxonomic depths, and identify several robustly supported clades, but fail to reconstruct very recent or very old divergences. Most species for which multiple haplotypes were available are shown to be non-monophyletic. There are three main causes of discordance between gene and species phylogenies—assuming that estimation of the former is correct: (a) gene duplication, (b) lineage sorting of ancestral polymorphisms, and (c) horizontal transfer (Maddison, 1997). Our data suggest that all three processes may have occurred/are occurring in oak gallwasps.

4.1. Gene duplication: cytochrome b pseudogenes

Cloning of cytochrome *b* for the four oak gallwasp species that failed direct sequencing revealed multiple cytochrome *b*-like sequences. Only one sequence from each species was found to be of the correct length and reading frame, suggesting that the other sequences found are probably nuclear mitochondrial pseudogenes (numts), as indicated by their mutational degeneration (presence of stop codons and/or indels) (Bensasson et al., 2001). Alternative explanations for the existence of additional mitochondrial-like sequences include heteroplasmy and intra-mitochondrial duplications (Bensasson et al., 2001; Mirol et al., 2000); reasons why both of these alternatives are rather unlikely explanations in studies like this one have been discussed in detail by Mirol et al. (2000). However, even sequences with correct length and reading frame are not necessarily amplifications of the 'true' mitochondrial copy; evidence from other gallwasp species suggests that, occasionally, more than one haplotype of correct length and reading frame exist (Rokas et al., in preparation). This is a potentially serious problem for phylogenetic reconstruction, since the derived numts are essentially paralogues to the mitochondrial loci.

4.2. Reconciling discordance between the cytb tree and current oak gallwasp taxonomy

The striking feature from this analysis is the lack of monophyly for 9 of the 16 species represented by multiple haplotypes (Table 1, Figs. 1 and 2). Here we consider in turn three probable causes for incongruence between the cytochrome *b* tree and the oak gallwasp taxonomy: (a) closure of cyclically parthenogenetic lifecycles, (b) lineage sorting, and (c) horizontal transfer through introgressive hybridization. We then discuss the implications of lineage sorting and hybridization for estimates of genetic divergence among taxa.

(a) *Closure of lifecycle pairs.* Although cyclical parthenogenesis has been established in almost all of the oak gallwasps whose biology has been examined in detail (Stone et al., 2002), many European species are still known only from either a sexual or a parthenogenetic generation. The adults and galls of the two generations of a single species are so distinct morphologically that

they cannot be paired on the basis of structure alone (Ambrus, 1974; Stone et al., 1995; Stone and Cook, 1998). Until recently, the pairing of parthenogenetic and sexual generations has required caging experiments in the field and as a result this process is far from complete (Folliot, 1964; Lund et al., 1998; Wehrmaker, 1998). It is therefore possible that clades in our phylogenies that are paraphyletic because they contain one taxon known only from a sexual generation, and another known only from an parthenogenetic generation may actually be the two alternate generations of a cyclical parthenogenetic species. The phylogenies presented here suggest two potential pairings. The first case is found in the *A. mayri* clade; if *A. aestivalis* (currently known only from its sexual generation) is the sexual generation of *A. lucidus* (currently known only from its parthenogenetic generation), then the apparent non-monophyly of *A. lucidus* haplotypes disappears (the *A. aestivalis* sequence is from Hungary and, as expected, groups with a haplotype from the same refuge). A second case is found in the *A. hartigi* clade, where *A. conificus* (currently known only from its parthenogenetic generation) may be the parthenogenetic generation of *A. cydoniae* (currently known only from its sexual generation), although in this case the non-monophyly of *A. conificus* does not disappear. Similar pairings may exist between *Andricus grossulariae* (sexual) and *A. panteli* (parthenogenetic) in the *A. mayri* clade.

(b) *Lineage sorting of ancestral polymorphism.* Wherever polymorphism exists in a marker prior to speciation, shared ancestral polymorphisms may generate incongruence between a gene tree and the species tree. Previous work on the cytochrome *b* locus in gallwasps has shown extensive intra-specific variation, the coalescence time of which can exceed the time depth reconstructed for separation between species (Rokas et al., 2001; Stone and Cook, 1998; Stone et al., 2001). Sorting of ancestral polymorphism is thus not unexpected in oak gallwasps, and results in non-monophyly of daughter species. Although lineage sorting and hybridization can be difficult or impossible to discriminate in very recent speciation events (Sota et al., 2001), the two can be discriminated by considering spatial patterns in haplotype similarity among species. By definition, the divergence of ancestral polymorphisms must predate the divergence of species (Edwards and Beerli, 2000). If stochastic sorting of ancestral polymorphism is responsible for similarity in haplotypes across species, there is thus no reason to expect similar haplotypes in different species to be sampled from sympatric populations. If, however, haplotype similarity across species is generated by hybridization, then these similarities should be observed in sympatric (and hence potentially interbreeding) populations of those species. A difficulty in distinguishing between these scenarios is that inadequate sampling of intra-specific variation in hybridizing taxa will lead to inference of sorting of ancestral polymor-

phism. Since the latter is probable a priori among closely related species, we take this cause as the more parsimonious explanation for non-monophyly of intra-specific haplotype diversity, and reserve hybridization as the explanation for cases with a clear correlation between sequence similarity and sympatry, as described below.

Incongruities between gene and species trees associated with sorting of ancestral polymorphism depend on the level of ancestral polymorphism, and the coalescence time of those polymorphisms in the phylogeny. Coalescence theory suggests that lineage sorting of ancestral polymorphisms will be faster for mtDNA loci because they have a population size one quarter of that for nuclear genes (Moore, 1995; see Hoelzer, 1997, for exceptions; Palumbi et al., 2001). Nuclear loci with high mutation rates are thus expected to bear a stronger signature of lineage sorting. A 590 base pair fragment of the nuclear gene long-wavelength opsin (LW *Rh*) showed good phylogenetic resolution for 21 species of gallwasps (Cook et al., 2002), but alleles from different phylogeographic refuges have yet to be examined. The high percentage of non-monophyletic haplotypes for a given species of oak gallwasp in the mitochondrial phylogenies presented here suggests that the species phylogeny of oak gallwasps may be difficult to resolve fully with molecular data. Similar predictions may be made for other taxonomic groups (e.g., many insects), that have experienced high speciation rates (e.g., Farrell, 1998; Kambysellis et al., 1995).

(c) *Horizontal transfer of haplotypes through introgressive hybridization.* Horizontal transfer of mitochondrial haplotypes is always a possibility where reproductive barriers between lineages are not complete, and several studies have shown introgression of the nuclear genome of one insect species by the mitochondrial genome of another through hybridization (e.g., Powell, 1983; Sota et al., 2001; Thelwell et al., 2000). Rearing experiments have shown that the closely related oak gallwasp species *Andricus kollari* and *Andricus lignicolus* can mate and produce offspring capable of back-crossing to both parents for up to three generations (Folliot, 1964). These two species are members of the same species group (the *A. kollari* clade) but are not sister taxa, so hybridization among even more closely related species is a possibility under natural conditions. We suggest that our data for *A. caputmedusae*, *Andricus dentimitratus*, and *Andricus mitratus* in the *A. quercuscalicis* clade provide evidence for horizontal exchange of haplotypes by introgressive hybridization, because haplotypes from different species from a given glacial refuge (Turkey) are more similar than members of the same species across refuges. As discussed above, this result is unlikely to arise through sorting of ancestral polymorphism alone. Although haplotypes from other refuges (Hungary and Italy) are included for this set of species, no clustering

according to geography is occurring. While this could be interpreted as lack of support for hybridization between these species, it has been shown that cytochrome *b* sequence data do not distinguish between these two refuges in all species examined so far (Rokas et al., 2001; Stone et al., 2001) (Rokas et al., unpublished data).

Of the three *Andricus* species involved, only one (*A. dentimitratus*) has a known sexual generation. Population genetic analyses strongly infer the presence of an unknown sexual generation in *A. caputmedusae* (Atkinson, 2000), while nothing is known of the lifecycle of *A. mitratus*. The inferred horizontal exchange of haplotypes provides further indirect evidence of sexual reproduction in all three species. All three species probably have sexual generations that develop in small spring bud galls on oaks in the section *Cerris* (Atkinson, 2000; Cook et al., 1998; Cook et al., 2002). Mating in related *Andricus* takes place on the surface of the gall (Stone and Sunnucks, 1993; Stone et al., 2001), and where the galls of two or more closely related species are at high density, it is possible that hybridization could occur under natural conditions.

The discussion above raises that general point that the total divergence between two species, as measured by the average pairwise divergence between lineage tips in the two species, includes three components: (a) allelic divergence between the species that has accumulated after lineage splitting (i.e., genuine phylogenetic signal for establishing species monophyly), (b) allelic divergence within the common ancestral species, and (c) divergence in alleles exchanged through hybridization. The second and third components may introduce significant errors into estimates of intra-specific diversity and of inter-specific genetic divergence—and hence into other parameters (such as past population size and molecular dating of speciation events) estimated from these. For example, if a significant portion of the diversity between the two species is ancestral, ignoring intra-specific variation in each species (as an estimate of that ancestral diversity) may lead to a gross over-estimation of inter-specific divergence (at least for recently diverged species) (see also Edwards, 1997). In contrast, high levels of introgression will lead to over-estimates of genetic diversity within species, and an underestimate of inter-specific divergence. Given that both errors due to lineage sorting and introgression will be more pronounced as the ancestral species becomes more structured (Nei and Takahata, 1993), these effects should be seriously considered in phylogeny reconstruction for widespread species with refuge-specific genetic diversity.

5. Conclusion

A range of interesting and general molecular phylogenetic questions can be addressed with oak gallwasps.

For example, the lineage sorting of ancestral polymorphisms (a possible consequence of the genetic structure of populations due to the Pleistocene glaciation cycles) may have a serious effect on within-genus phylogenetics. Furthermore, the possibility of introgressive hybridization between species and the reality of, as yet, unmatched generations make the task of phylogenetic reconstruction even more difficult.

The phylogenetic history of a group such as oak gallwasps, where divergence times between taxa in a well-sampled data matrix differ by an order of magnitude, is unlikely to be resolved by any single locus. A combined approach, employing several loci that show different and complementary rates (and modes) of evolution, is probably the only way to obtain well-supported phylogenetic hypotheses (Hillis et al., 1996; Rokas and Holland, 2000; Rokas et al., 2002). Lessons learned from single-locus studies will be a most useful guide towards achieving this goals.

References

Ambrus, B., 1974. Cynipida–Gubacsok–Cecidia Cyniparum. Academic Press, Budapest, Hungary.

Askew, R.R., 1984. The biology of gall wasps. In: Ananthakrishnan, T.N. (Ed.), The Biology of Galling Insects. Oxford and IBH Publishing Co., New Delhi, pp. 223–271.

Atkinson, R.J., 2000. The genetic analysis of natural history, reproductive strategy, and population structure in European oak gallwasps (Hymenoptera: Cynipidae). Ph.D. Thesis, University of Oxford, Oxford, p.286.

Bensasson, D., Zhang, D.-X., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends. Ecol. Evol. 16, 314–321.

Cook, J.M., Rokas, A., Pagel, M., Stone, G.N., 2002. Evolutionary shifts between host oak species and host plant organs in *Andricus* gallwasps. Evolution, in press.

Cook, J.M., Stone, G.N., Rowe, A., 1998. Patterns in the evolution of gall structure and life cycles in oak gallwasps (Hymenoptera: Cynipidae). In: Csoka, G., Mattson, W.J., Stone, G.N., Price, P.W. (Eds.), The Biology of Gall-Inducing Arthropods. U.S. Department of Agriculture, Forest Service, North Central Research Station, St. Paul, Minnesota, pp. 261–279.

Drown, D.M., Brown, J.M., 1998. Molecular phylogeny of North American oak-galling Cynipini (Hymenoptera: Cynipidae) supports need for generic division. In: Csoka, G., Mattson, W.J., Stone, G.N., Price, P.W. (Eds.), The Biology of Gall-Inducing Arthropods. U.S. Department of Agriculture, Forest Service, North Central Research Station, St. Paul, Minnesota, pp. 241–246.

Edwards, S.V., 1997. Relevance of microevolutionary processes to higher level molecular systematics. In: Mindell, D.P. (Ed.), Avian Molecular Evolution and Systematics. Academic Press, New York, pp. 251–278.

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

Farrell, B.D., 1998. "Inordinate fondness" explained: Why are there so many beetles? Science 281, 555–559.

Folliot, R., 1964. Contribution a l' etude de la biologie des Cynipides gallicoles (Hymenoptera: Cynipoidea). Annals des Sciences Naturelles, Zoologie (Paris), Series 12 6, 407–564.

Hillis, D.M., Mable, B.K., Moritz, C., 1996. Applications of molecular systematics. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer, Sunderland, MA, pp. 515–543.

Hoelzer, G.A., 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. Evolution 51, 622–626.

Huelsenbeck, J.P., Rannala, B., 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276, 227–232.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Jermiin, L.S., Crozier, R.H., 1994. The cytochrome *b* region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: sequence divergence in Hymenoptera may be associated with nucleotide content. J. Mol. Evol. 38, 282–294.

Kambysellis, M.P., Ho, K.F., Craddock, E.M., Piano, F., Parisi, M., Cohen, J., 1995. Pattern of ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny. Curr. Biol. 5, 1129–1139.

Kinsey, A.C., 1936. The origin of higher categories in *Cynips*. Ind. Univ. Publ., Sci. Ser. 4, 1–334.

Liljeblad, J., Ronquist, F., 1998. A phylogenetic analysis of higher-level gall wasp relationships (Hymenoptera: Cynipidae). Syst. Entomol. 23, 229–252.

Lund, J.N., Ott, J.R., Lyon, R.J., 1998. Heterogony in *Belonocnema treatae* Mayr (Hymenoptera: Cynipidae). Proc. Entomol. Soc. Wash. 100, 755–763.

Lyon, R.J., 1996. New cynipid wasps from the southwestern United States (Hymenoptera: Cynipidae). Pan-Pacific Entomol. 72, 181–192.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Manos, P.S., Doyle, J.J., Nixon, K.C., 1999. Phylogeny, biogeography and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). Mol. Phylog. Evol. 12, 333–349.

Melika, G., Abrahamson, W.G., 2000. Review of the cynipid gall wasps of the genus *Loxaulus* Mayr (Hymenoptera: Cynipidae) with descriptions of new species. Proc. Entomol. Soc. Wash. 102, 198–211.

Melika, G., Csoka, G., Pujade-Villar, J., 2000. Check-list of oak gall wasps of Hungary, with some taxonomic notes (Hymenoptera: Cynipidae, Cynipinae, and Cynipini). Ann. Hist.-Nat. Mus. Natl. Hung. 92, 265–296.

Mirol, P.M., Mascheretti, S., Searle, J.B., 2000. Multiple nuclear pseudogenes of mitochondrial cytochrome *b* in *Ctenomys* (Caviomorpha, Rodentia) with either great similarity to or high divergence from the true mitochondrial sequence. Heredity 84, 538–547.

Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49, 718–726.

Nei, M., Takahata, N., 1993. Effective population size, genetic diversity, and coalescence time in subdivided populations. J. Mol. Evol. 37, 240–244.

Nieves-Aldrey, J.L., 1987. Estado actual del conocimiento de la subfamilia Cynipinae (Hymenoptera, Parasitica, and Cynipidae) en la Peninsula Iberica. Eos 63, 179–195.

Nieves-Aldrey, J.-L., 2001. Hymenoptera: Cynipidae. Museo Nacional de Ciencias Naturales, CSIC, Madrid.

Palumbi, S.R., Cipriano, F., Hare, M.P., 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. Evolution 55, 859–868.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Powell, J.R., 1983. Interspecific cytoplasmic gene flow in the absence of nuclear gene flow: evidence from *Drosophila*. Proc. Natl. Acad. Sci. USA 80, 492–495.

Rokas, A., 2001. Tree-thinking, molecules and gallwasps; analysing evolutionary patterns in European gallwasps using a molecular phylogenetic approach. Ph.D. Thesis, University of Edinburgh, Edinburgh, p. 255.

Rokas, A., Atkinson, R.J., Brown, G.S., West, S.A., Stone, G.N., 2001. Understanding patterns of genetic diversity in the oak gallwasp *Biorhiza pallida*: demographic history or a *Wolbachia* selective sweep? Heredity 87, 294–304.

Rokas, A., Holland, P.W.H., 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol. Evol. 15, 454–459.

Rokas, A., Nylander, J.A.A., Ronquist, F., Stone, G.N., 2002. A maximum likelihood analysis of eight phylogenetic markers in gallwasps (Hymenoptera: Cynipidae); implications for insect phylogenetic studies. Mol. Phylog. Evol. 22, 206–219.

Ronquist, F., 1994. Evolution of parasitism among closely related species: phylogenetic relationships and the origins of inquilism in gall wasps (Hymenoptera: Cynipidae). Evolution 48, 241–266.

Ronquist, F., 1999. Phylogeny, classification and evolution of the Cynipoidea. Zool. Scr. 28, 139–164.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Sota, T., Ishikawa, R., Ujie, M., Kusumoto, F., Vogler, A.P., 2001. Extensive trans-species mitochondrial polymorphisms in the carabid beetles *Carabus* subgenus *Ohomopterus* caused by repeated introgressive hybridisation. Mol. Ecol. 10, 2833–2847.

Stone, G.N., Atkinson, R., Rokas, A., Csóka, G., Nieves-Aldrey, J.-L., 2001. Differential success in northwards range expansion between ecotypes of the marble gallwasp *Andricus kollari*: a tale of two lifecycles. Mol. Ecol. 10, 761–778.

Stone, G.N., Cook, J.M., 1998. The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. Proc. R. Soc. Lond., Ser. B 265, 979–988.

Stone, G.N., Schönrogge, K., Atkinson, R.J., Bellido, D., Pujade-Villar, J., 2002. The population biology of oak gallwasps (Hymenoptera: Cynipidae). Annu. Rev. Entomol. 47, 633–668.

Stone, G.N., Schönrogge, K., Crawley, M.J., Fraser, S., 1995. Geographic and between-generation variation in the parasitoid communities associated with an invading gallwasp, *Andricus quercuscalicis* (Hymenoptera: Cynipidae). Oecologia 104, 207–217.

Stone, G.N., Sunnucks, P., 1993. Genetic consequences of an invasion through a patchy environment—the cynipid gallwasp *Andricus quercuscalicis* (Hymenoptera: Cynipidae). Mol. Ecol. 2, 251–268.

Swofford, D.L., 2002. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer, Sunderland, MA.

Thelwell, N.J., Huisman, R.A., Harbach, R.E., Butlin, R.K., 2000. Evidence for mitochondrial introgression between *Anopheles bwambae* and *Anopheles gambiae*. Insect Mol. Biol. 9, 203–210.

Wehrmaker, A., 1998. On the beginnings of cecidology in 19th century North America: Bassett's discovery of heterogony in oak gall wasps (Hymenoptera: Cynipidae). In: Csóka, G., Mattson, W.J., Stone, G.N., Price, P.W. (Eds.), The Biology of Gall-inducing Arthropods. U.S. Department of Agriculture, Forest Service, North Central Research Station, St. Paul, Minnesota, pp. 106–110.

Weld, L.H., 1957. Cynipid Galls of the Pacific Slope. Privately printed, Ann Arbor, Michigan.

Weld, L.H., 1959. Cynipid Galls of the Eastern United States. Privately printed, Ann Arbor, Michigan.

Weld, L.H., 1960. Cynipid Galls of the Southwest. Privately printed, Ann Arbor, Michigan.