A single origin of gall association in a group of parasitic wasps with disparate morphologies

Alejandro Zaldivar-Riverón a,*, Sergey A. Belokobylskij b,c, Virginia León-Regagnon a, Juan José Martínez d, Rosa Briceno e, Donald L.J. Quicke f,g

b Zoological Institute, Russian Academy of Sciences, Universitetskaya nab. 1, St. Petersburg 199034, Russia
c Museum and Institute of Zoology, PAN, Wilcza 64, Warsaw 00-679, Poland
d División de Entomología, Museo Argentino de Ciencias Naturales “Bernardino Rivadavia”, Av. Angel Gallardo 470, C1405DJR, Buenos Aires, Argentina
e Universidad Centroccidental “Lisandro Alvarado”, Decamato de Agronomía, Depto. de Ciencias Biológicas, Sección de Entomología, Cabudare, Estado Lara, Venezuela
f Division of Biology and Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot, Berkshire, SL5 7PY, UK
g Department of Entomology, Natural History Museum, London SW7 5BD, UK

Received 17 August 2006; revised 10 May 2007; accepted 17 May 2007
Available online 9 June 2007

Abstract

The braconid wasp subfamily Doryctinae mainly comprises idioibiont ectoparasitoids of other insect larvae. In recent years, however, members of a few genera have been discovered to be associated with galls from various unrelated host plant families, with some of these being gall inducers whereas others are suspected as being predators of gallers. Because of their considerable morphological differences, these gall-associated taxa traditionally have been placed in separate tribes or even in other subfamilies. In this study, we investigate the phylogenetic relationships among representatives of a number of different doryctine genera, including five of its seven gall-associated genera using two genetic markers. Here we analyzed the length-variable 28S sequence data based on secondary structure both excising the unalignable regions and recoding them according to indel length. In addition, multiple alignments were carried out with a range of gap-opening and extension parameters. The combined (28S + CO1) phylogenetic hypotheses obtained, both excluding and recoding the unalignable regions, recover a clade comprising the five gall-associated genera, and most of the analyses using multiple alignments also support this relationship. These results support a scenario in which secondary phytophagy evolves from initially attacking primary gall-forming hosts. The relationships recovered are also more congruent with a model that explains the macroevolution of insect plant association in the Doryctinae as reflecting geographic proximity rather than host plant relationships. Further, our phylogenetic hypotheses consistently show that one of the main morphological features employed in the higher level classification of the Doryctinae is actually highly homoplastic.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Braconidae; Doryctinae; Cyclostomes; Character evolution, Molecular phylogeny; Secondary structure alignment; Multiple alignment

1. Introduction

In the Hymenoptera, gall formation, predation of gall-inducers and inquilinism are related phenomena that occur in several lineages, principally in the Chalcidoidea and Cynipoidea, but which have also been discovered recently in a few members of the cyclostome group of the otherwise almost entirely parasitic family Braconidae (see Wharton...
and Hanson, 2005 for a review). Starting from the first report of phytophagy in *Allorhogas Gahan* almost 20 years ago (Macêdo and Monteiro, 1989), cecidogenesis (i.e., gall development) in the Braconidae has subsequently been confirmed only in other two cyclostome genera, *Monitoriella Hedqvist* (Infante et al., 1995) and *Mesosota van Achterberg* (Austin and Dangerfield, 1998). Some other braconid wasps have only ever been reared from galls, though some of these appear to display various forms of gall association rather than being the primary gall-formers themselves. Most of these gall-associated species belong to the large and cosmopolitan subfamily Doryctinae, though this feature is also displayed by species of three genera of the Mesostoidea (sensu Zaldivar-Riverón et al., 2006) (*Mesosota*, *Aspilodemon Fischer*, and *Hydrangeocola* Brèthes; Brèthes, 1927; Austin and Dangerfield, 1998; Oda et al., 2001).

Doryctine wasps are mainly idiobiont (i.e., they arrest the host development at the time of parasitization) ectoparasitoids of xylophagous Coleoptera larvae (Belokobylskij et al., 2004a,b). However, gall association is exhibited by species of seven Neotropical doryctine genera, most of which have been reared from different unrelated host plant families. In *Allorhogas*, several species are known to induce galls in seeds of species of Melastomataceae, Bignoniaceae and at least four legume (Fabaceae) genera, though some other species are thought to be parasitoids of gall inhabitants (Macêdo and Monteiro, 1989; Macêdo et al., 1998; Marsh, 2002; Marsh et al., 2000). Moreover, the type species of *Percnobracon Kieffer*, *Pe. stenopterus Kieffer*, was reared from ceccidomyiid (Diptera) galls on the legume *Prosopis strombulifera* (Lam.) Benth (Kieffer, 1910), and recently this and two other species of the genus have been reared from the Argentinian endemic *Pr. caldenia* Burkart (Martinez, 2006). Species of *Monitoriella*, on the other hand, induce leaf galls on *Philodendron Schott* (Araeaceae) (Infante et al., 1995), whereas some members of *Labania Hedqvist* and *Psenobolus Reinhard* have been observed to develop in galls located on aerial roots and leaves, and on the syconia of *Ficus* L. (Moraceae), respectively (Marsh, 2002; Ramirez and Marsh, 1996). Finally, the only described species of *Donquickea Massh* and *Mononeuron Fischer* were reported to be reared from ceccidomyiid galls of *Mikania* sp. (Asteraceae) and *Eugenia rotundifolia Casar* (Myrtaceae) (Penteado-Dias, 2000) and from a species of *Dugetia* A.St.-Hil (Annonaceae) (Fischer, 1981), respectively.

Since the process of host shift in phytophagous insects involves a number of genetic and ecological factors (e.g., secondary metabolites, competitors, predators; Thompson, 1999), this is normally expected to occur more between closely related host plant species rather than in distant, unrelated families (e.g., Becerra, 1997; Köpf et al., 1998). The different host plant families inhabited by the above doryctine genera would therefore suggest that at least the confirmed ceccidogenic *Allorhogas* and *Monitoriella* are distantly related. This idea has been generally supported by the considerable external morphological differences shown by the different gall-associated doryctine genera. For this reason, these taxa have been placed in separate and in some cases exclusive tribes (e.g., Labanini, Percnobraconini) or even in other subfamilies (e.g., Monitoriellinae: Wharton, 1993; but see Zaldivar-Riverón et al., 2006). It has been argued, however, that the external morphology of doryctine wasps is quite homoplastic (Belokobylskij et al., 2004b), and thus several features employed to differentiate its higher taxa could have actually evolved separately in repeated occasions.

Recently, a simultaneous molecular and morphological phylogenetic analysis among representatives of the cyclostome braconid subfamilies strongly supported the gall-associated *Labania* and *Monitoriella* as being sister taxa (Zaldivar-Riverón et al., 2006). That study, however, did not sample other gall-associated doryctine genera and only included a scarce number of members of this group. Thus, relationships both within the subfamily and among the gall-associated genera were far from conclusive.

Different strategies have been proposed for aligning length-variable rDNA sequence data, though unalignable regions present in rDNA sequence alignments still are commonly excluded and therefore the results obtained are based exclusively on regions that appear to be length-conserved. In the case of secondary structure alignment (Gillespie, 2004), strict criteria of matched base-pairing are used to delimit putatively homologiseable stem regions from ambiguously alignable loops and expansion zones. However, coding approaches also exist that align length-variable regions conditionally based on indel length (e.g., Simmons and Ochoterena, 2000; Lutzoni et al., 2000; Zaldivar-Riverón et al., 2006). Alternatively, computer algorithms are often employed to find best matches between bases in regions of ambiguous alignment, and the incorporated variation may contribute to the phylogenetic analysis performed. These include multiple alignment (e.g., Clustal X, Thompson et al., 1997; MALIGN, Wheeler and Gladstein, 1994) and dynamic-based alignments (e.g., POY, Wheeler, 1996; Gladstein and Wheeler, 1996), and when using them it is increasingly common to use a ‘sensitivity’ approach, where the stability of the results is assessed across a range of user-defined parameters. Here we utilize three different alignment strategies, secondary structure alignment excluding unalignable regions, secondary structure alignment with conditional alignment of indel regions, and multiple alignment using a range of parameters to assess the sensitivity of our results to alignment procedure.

In this study, we report a close relationship among most of the currently known gall-associated doryctine wasps based on a molecular phylogenetic analysis using two gene fragments. We discuss the implications of this finding on the origin of cecidogeny and on the evolution of host plant shifts displayed among these taxa. Moreover, we assess the evolution of one of the key features employed to distinguish among supraspecific taxa within the group, the length of acrosternite (strongly sclerotised basal sternal plate) of
the first metasomal segment, and consider the consequences that our results will have in the higher level classification of the Doryctinae.

2. Materials and methods

2.1. Taxon sampling

DNA sequence data were generated for a total of 51 taxa belonging to 50 different doryctine genera. The taxon sampling comprised representative genera from the Afro-tropical, Australian, Neotropical, Oriental, and Palaearctic regions, which belong to most of the tribes recognized by Belokobylskij (1992). Monophyly of the Doryctinae, previously supported by relatively little convincing morphological evidence (see Quicke et al., 1992), was tested by including members of seven cyclostome subfamilies, with the rhyssaline Rhyssalus Haliday used for rooting all trees: previous studies have consistently recovered the Rhyssalinae as the sister group of all of them (Belshaw et al., 1998, 2000; Dowton et al., 2002; Zaldivar-Rivero et al., 2006). Details of the included material, voucher information, and GenBank/EMBL Accession Numbers are given in Table 1.

2.2. DNA sequence data and alignment

The genetic markers examined included ~650 bp of the second and third domains of the nuclear 28S rDNA gene and 603 bp of the COI mtDNA gene. Genomic DNA was extracted from ethanol preserved and up to 15 years old dry pinned specimens. The DNA extraction and amplification procedures employed were those described in Zaldivar-Rivero et al. (2006). 28S sequences were obtained using the primers designed by Belshaw and Quicke (1997) (fwd: 5'-GGC AAC AAG TAC CGT GAG GG-3') and Mardulyn and Whitfield (1999) (rev: 5'-TAG TTC ACC ATC TTT CGG GTC CC-3'). COI sequences were amplified using the primers designed by Folmer et al. (1994) (LCO 5'-GTT CAA CAA ATC ATA AAG ATA TTG G-3', HCO 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3').

COI sequences were unambiguously aligned based on their translated amino acids; however, alignment of the length-variable 28S gene fragment was more problematic and here we employ three different strategies and compare the results to assess the sensitivity of our findings to choice of alignment method.

2.2.1. Secondary structure alignment

Alignment of the 28S sequences followed the 28S bracoid secondary structure alignment proposed by Gillespie et al. (2005). Length variable regions were identified and initially characterized according to the three categories proposed by Gillespie (2004) [regions of ambiguous alignment (RAAs), regions of slipped-strand compensation (RSCs), and regions of expansion and contraction (RECs)]. RAAs and their adjacent RECs were then combined into single unalignable regions for subsequent analyses using conditional alignment (see below). The secondary structure alignment excluding these ambiguous regions is termed 28SN.

2.2.2. Indel coding and conditional alignment

Regions of the 28S data set whose positional homology could not be confidently assigned were recoded following the approach used by Zaldivar-Rivero et al. (2006). This approach offers an objective criterion for preserving all the nucleotide variation that is not in conflict due to sequence length variation and additionally includes information on indel length by attributing a ‘morphological’ character state to indels of identical length regardless to their base composition. This data set, comprising the secondary structure alignable regions (28SN) together with the unalignable regions of identical length and the ‘morphological’ states referring to indel lengths is referred to as 28SA.

2.2.3. Multiple alignment

Multiple alignments of the 28S data were created with Clustal X (Thompson et al., 1997) with each of the following range of gap opening-gap extension penalties: 1:1, 2:1, 2:2, 4:1, 4:2, 4:4, 6:1, 6:2, 6:4, 6:6, 8:1, 8:2, 8:4, 8:6, 8:8, 10:1, 10:2, 10:4, 10:6, 10:8, 10:10. All of these used a DNA transition: transversion weight ratio of 1:0.1, and the default pairwise alignment parameters.

2.3. Phylogenetic analysis

Bayesian MCMC analyses were performed for the separate and combined data sets (with 28SN and 28SA matrices) using MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003). Two independent analyses were run simultaneously for each data set, each consisting of 2 million generations, sampling trees every 1000 generations, and using four chains and uniform priors on three topologies, with the latter implying non-uniform prior probabilities of clades due to these are dependent on the number of taxa in a clade as well as the number of taxa in the analysis (Pickett and Randle, 2005). The GTR + I + Γ model of sequence evolution was the model selected for the two gene markers examined according to the likelihood ratio test implemented with MrModeltest version 2.2 (Nylander, 2004). There is no reason to assume a priori that the states derived from the 28SA indel characters, whose states are coded as ‘morphological’ (i.e., with numerical coding), could have different frequencies. The Mk + Γ (Markov k; Lewis, 2001) model of evolution was therefore used for the above characters. This model was originally designed to deal with morphological characters and considers equal base frequencies, equal mutation rates and makes the likelihood conditional on characters being variable (Lewis,
Table 1
Provenances and voucher and EMBL/GenBank Accession Numbers of the taxa included

<table>
<thead>
<tr>
<th>Voucher No.</th>
<th>28S</th>
<th>COI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Taxon</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acanthorhogas sp., Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-713</td>
<td>DQ498942</td>
</tr>
<tr>
<td>Acrophasmus exilis Marsh, Costa Rica&lt;sup&gt;d&lt;/sup&gt;</td>
<td>IB-02</td>
<td>DQ498943</td>
</tr>
<tr>
<td>Ateleykus arawak, Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-723</td>
<td>AY935471&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Allorhogas sp., La Pampa, Argentina&lt;sup&gt;dt&lt;/sup&gt;</td>
<td>IB-18</td>
<td>DQ498927</td>
</tr>
<tr>
<td>Allorhogas sp., Cartago, Turrialba, Costa Rica&lt;sup&gt;dt&lt;/sup&gt;</td>
<td>IB-20</td>
<td>—</td>
</tr>
<tr>
<td>Aphelopsia striato Braet &amp; Barbalho, French Guyana (holotype)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-838</td>
<td>DQ498941</td>
</tr>
<tr>
<td>Barbalboa contohpleba Marsh, Costa Rica&lt;sup&gt;tf&lt;/sup&gt;</td>
<td>Jo-723</td>
<td>DQ498935</td>
</tr>
<tr>
<td>Barbalboa licina Marsh, Costa Rica&lt;sup&gt;tf&lt;/sup&gt;</td>
<td>Jo-739</td>
<td>—</td>
</tr>
<tr>
<td>Caenopachys hartigii (Ratzeburg), Corsica, France&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>DQ498941</td>
</tr>
<tr>
<td>Caenophanes sp., Africa&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Jo-781</td>
<td>AY935465&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Curtisella hansoni, Costa Rica&lt;sup&gt;d&lt;/sup&gt;</td>
<td>IB-03</td>
<td>DQ498944</td>
</tr>
<tr>
<td>Doryctes heydenii Reinhard, Palaearctic&lt;sup&gt;d&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Doryctes leucogaster (Nees), Israel&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-876</td>
<td>AY935472&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ecphylus sp., South America&lt;sup&gt;c&lt;/sup&gt;</td>
<td>—</td>
<td>AJ302894&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Euscelinus sarawacus Westwood, Thailand&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-871</td>
<td>DQ498920</td>
</tr>
<tr>
<td>Hansonorum sp. nov. aff. pali, Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-728</td>
<td>DQ498934</td>
</tr>
<tr>
<td>Hexabola suctus Curtis, Ascot, UK&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Jo-650</td>
<td>AY935473&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hemidoryctes carbonarius (Ashmead), Diego Suarez Province, Madagascar&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-592</td>
<td>DQ498921</td>
</tr>
<tr>
<td>Heterospilus prosopidis Viereck, Silwood culture, UK&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>AY935449&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Heterospilus sp., Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>IB-01</td>
<td>DQ498926</td>
</tr>
<tr>
<td>Histeromeroides sp., Rio Napo, Ecuador&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Jo-866</td>
<td>DQ498940</td>
</tr>
<tr>
<td>Hypodoryctes sibiricus Kokujev, no data&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>AJ302895&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ipodoryctes sp., St. Phillippe, Reunion Island&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-719</td>
<td>DQ498939</td>
</tr>
<tr>
<td>Leptorhaconotus sp. (1), Madagascar&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-655</td>
<td>AY935479&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Liobracon sp., Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-714</td>
<td>AY935467&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Masonius sp., Cuyagua, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-578</td>
<td>DQ498933</td>
</tr>
<tr>
<td>Megaloprostus sp., Colombia&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-654</td>
<td>AY935466&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Monitoriella sp., Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-733</td>
<td>AY935457&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Nottospathius sp., Cauagaua, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-596</td>
<td>AY935477&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Odontobracon cf., Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-712</td>
<td>AY935468&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Onitigus bicolor Braet, Barbalho &amp; van Achterberg, French Guyana (holotype)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Jo-834</td>
<td>DQ498937</td>
</tr>
<tr>
<td>Ontsira imperator (Haliday), Honshu Ibaraki, Japan&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-971</td>
<td>DQ498919</td>
</tr>
<tr>
<td>Pedinus sp., Costa Rica&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>DQ498936</td>
</tr>
<tr>
<td>Percnonbracon cf., Hobby Paris, Santa Rosa, La Pampa, Argentina&lt;sup&gt;d&lt;/sup&gt;</td>
<td>IB-19</td>
<td>DQ498930</td>
</tr>
<tr>
<td>Platyspathius sp., Benin&lt;sup&gt;a&lt;/sup&gt;</td>
<td>AL-183</td>
<td>DQ498915</td>
</tr>
<tr>
<td>Psenobolas sp., Camboto, Venezuela&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Jo-666</td>
<td>DQ498929</td>
</tr>
<tr>
<td>Pseudodoryctes sp., Malawi&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Jo-679</td>
<td>DQ498922</td>
</tr>
<tr>
<td>Rhaconotus emarginatus Marsh, Costa Rica&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-720</td>
<td>DQ498914</td>
</tr>
<tr>
<td>Rhacontsira sp., Hoa Binh Province, Vietnam&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-923</td>
<td>DQ498917</td>
</tr>
<tr>
<td>Rinamba opacicollis Cameron, Madagascar&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-588</td>
<td>DQ498923</td>
</tr>
<tr>
<td>Schlettereriella sp., Kibale, Uganda&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-673</td>
<td>AY935478&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Semirhytus sp., Municipality of Lara, Yacambu, Venezuela&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Jo-628</td>
<td>DQ498931</td>
</tr>
<tr>
<td>Semirhytus sp., no data&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Jo-983</td>
<td>—</td>
</tr>
<tr>
<td>Spathiomorpha sp. aff. longipalpis, Japan&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-711</td>
<td>DQ498924</td>
</tr>
<tr>
<td>Spathius agrili Yang, China</td>
<td>—</td>
<td>AY920299&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Spathius sp., Tollara Province, Madagascar&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-591</td>
<td>AY935476&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Spathius (Antespathius) sp., Madagascar&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Jo-910</td>
<td>DQ498918</td>
</tr>
<tr>
<td>Stenocorse sp., Belize&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Jo-750</td>
<td>AY935475&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Syngaster lepidus Brullé, Australia&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>AJ245698&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tarasco sp., Camboto, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-926</td>
<td>DQ498932</td>
</tr>
<tr>
<td>Tarasco sp., Camboto, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-587</td>
<td>—</td>
</tr>
<tr>
<td><strong>Outgroups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cystomastax sp., Cucagua, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-574</td>
<td>AY935445&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hormius sp., Mahajanga Province, Madagascar&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-582</td>
<td>AY935455&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Oncophanes sp., Ascot, UK&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>AY935481&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pambolus sp., Choroni, Venezuela&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Jo-597</td>
<td>AY935458&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
2001). The 28S nucleotide and indel characters and the three COI codon positions were each treated as unlinked partitions. Likelihoods reached their plateau ranges before 100,000 generations in all the analyses; therefore, their first 100 sampled trees were discarded as burn-in. Relationships derived from the postburn-in samples were similar in the two analyses run for each data set and thus they were pooled and used to produce a 50% majority rule consensus tree with posterior probabilities of clades.

The analyses for the combined data sets derived from the 28S multiple alignments used the same search strategies and parameters as above, except for the number of run generations, which was of 1 million. Burn-in in all the analyses was determined to be completed by 100,000 generations. We also carried out maximum parsimony (MP) analyses for the combined and separate data sets that resulted from the 28S multiple alignments with PAUP* version 4.0b10 (Swofford, 1998). Alignments from each parameter combination were analyzed using 10,000 random additions followed by TBR branch-swapping and holding no more than one most parsimonious tree (MPT) for swapping at any one time. The resulting MPTs from each search were then used as starting trees for further TBR searches with unlimited maxtrees. Application of reweighting on these using the maximum value of the reten-

tation index as the reweighting function, followed by reweighting characters back to unity (Quicke et al., 2001) failed to recover any more parsimonious trees.

The matrices of the separate and combined data sets with 28SN and 28SA and their topologies reconstructed can be downloaded from the TreeBase web page (accession no. S1800). The matrices with the multiple alignments analyzed and their resulting topologies can be obtained upon request to AZR.

2.4. Test of alternative hypotheses

A Bayesian approach (Buckley et al., 2002; Reeder, 2003; Brandley et al., 2004) was implemented for hypothesis testing of alternative topologies not present in our majority consensus trees derived from the two simultaneous analyses. In particular, we examined the monophyly of each of the large tribes Doryctinae, Hecabolini and Spathiini (sensu Belokobylskij, 1992). Of these, the gall-associated Allorhogas and Psenobolus were placed in the Hecabolini and Spathiini, respectively. In this approach, a 95% credible set of trees sampled after burn-in was assembled for each simultaneous analysis, considering a

given alternative phylogenetic hypothesis as statistically rejected when this was absent in all of the credible set of trees. The presence of the alternative hypotheses within the 95% posterior intervals was detected by creating the selected topological constraint with MacClade 4.06 (Maddison and Maddison, 2003) and then using this to filter the set of credible trees using the option ‘compatible with constraints’ included in PAUP* version 4.0b10 (Swofford, 1998).

The Shimodaira-Hasegawa (SH; Shimodaira and Hase-
gawa, 1999) test was also performed with PAUP* version 4.0b10 (Swofford, 1998) to test for significant differences between the likelihood of the 28SN + COI Bayesian phylogeny and the likelihood of alternative Bayesian topologies that enforced the above tribes each as monophyletic. The SH tests were carried out using the full optimization sampling with 1000 replicates. The alternative Bayesian topologies were obtained with the constraint option available in MrBayes version 3.1.2, using the same strategies and parameters employed in the unconstrained analyses. Burn-in in these analyses was determined occur after 300,000 generations.

2.5. Analysis of character evolution

The evolution of gall-association and of one morphological feature used to distinguish higher-level taxa in the Dor-
yctinae, the length of acrosternite of the first metasomal tergite, were investigated mapping their selected character states onto the 2SN + COI Bayesian phylogeny employing maximum likelihood character optimizations using Mesquite 1.6 (Maddison and Maddison, 2004). The morphological feature was defined to have three states [(0) 0.2–0.25 times as long as tergite, petiole absent; (1) 0.3–0.5 times as long as tergite, petiole present but short; (2) 0.6–0.85 times as long as tergite, petiole present and long]. The gall association trait was defined as binary [(0) absent/unknown; (1) present]. Ecological information was scored from literature and personal observations, and morphological data was obtained from the sequenced specimens and from literature. The character states scored for the two features examined can be retrieved from the electronic supplementary material. The Markov k-state one parameter model (Mk1; Lewis, 2001) was used for the likelihood character optimizations, which assumes a single rate for transitions between character states. A likelihood ratio test was carried out to discriminate for the best estimate of ancestral state at each node following Pagel (1994). In this
test, a state is considered as the best estimate for a given branch if it differs by $\geq 2.0$ from one or more states with higher negative log-likelihood values, and considers the ancestral reconstruction as ambiguous when the difference between states is $\leq 2.0$.

The full hierarchical Bayesian method for inferring ancestral states (Huelsenbeck and Bollback, 2001) was also performed for the gall association character with MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) using the 28SN + COI data set. This method incorporates uncertainty in the tree, branch lengths, and substitution model parameters, approximating the posterior probability of ancestral states employing Markov Chain Monte Carlo (Huelsenbeck and Bollback, 2001). The Bayesian analysis performed to infer the ancestral state for the above character was run using the same parameters and settings as the original phylogenetic analyses, except by including the gall association character in the matrix analyzed and constraining the node with the five gall associated genera examined (see Section 3).

3. Results

3.1. Phylogenetic reconstructions

Despite several attempts were made to obtain both 28S and COI sequences for all the ethanol preserved and dry pinned specimens examined, COI sequences could not be obtained for three taxa [Psenobolus, Spathius Nees (Antespispathius Belokobylskij), and Ecphysus Foerster]. Moreover, in five genera (Tarasco Marsh, Hypodoryctes Kokujev, Barbalhoa Marsh, Semirimythus Szepligeti, and Allorrhagos) sequences of the two genetic fragments were obtained from different specimens, but in the case of Tarasco sp. and Hypodoryctes sibiricus Kokujev, the sequences were from different individuals of the same species. The 28S data set consisted of 538 unambiguously aligned positions and 15 unalignable regions. Inclusion of the recoded unalignable regions added 263 characters to the data set, of which 249 were from the nucleotide clusters created and 14 from the ‘morphological’ indel characters.

The majority rule consensus trees that resulted from the separate 28SN and COI Bayesian analyses performed are shown in Figs. 1 and 2, respectively. The COI Bayesian phylogeny recovered the highest number of significantly supported clades in comparison to the analyses with the 28SN and 28SA data sets (COI = 21; 28SN = 18; 28SA = 19). The topologies reconstructed from the 28SN and 28SA data sets were mainly similar, with no conflicting significantly supported clades among them. The three separate analyses performed recovered a clade with gall-associated genera. A clade with Heterospiulus Haliday and the gall-associated genera included except Labania was recovered in the 28SN topology (BPP = 0.66), whereas in the 28SA topology the five gall-associated genera were recovered in a single clade together with the above species of Heterospiulus (BPP = 0.89). Moreover, of the four gall-associated genera examined in the COI Bayesian analysis, three of them (Allorrhagos, Labania, and Monitoriella) were significantly recovered as monophyletic (BPP = 1.0). None of the separate analyses recovered the taxa with elongated first metasomal tergite to be monophyletic.

The 28SN + COI Bayesian topology with its posterior probabilities and the posterior probabilities of similar clades recovered in the 28SA + COI Bayesian topology are shown in Fig. 3. The number of significant posterior probabilities of clades considerably increased in the two combined analyses with respect to the separate ones. A total of 22 clades were similar and significantly supported in the two combined analyses, but two clades were only significantly supported in the analysis with 28SN and four clades increased their BPP to significant values in the analysis with 28SA (Fig. 3).

The two simultaneous analyses both show a non-monophyletic Doryctinae, though the clades implicated are not significantly supported and involve the morphologically less derived genera Doryctes Haliday and Ontisra Cameron. In both topologies, a clade with the species of the latter two genera (BPP: 28SN, 28SA = 1.0) appears as the sister group of a clade with two divisions, one containing all the non-doryctine subfamilies except rhysyalines and the other one with the remaining doryctine genera. This major doryctine clade (BPP: 28SN, 28SA = 1.0) shows an apparent geographic structure that is composed of two main clades. One of these contains genera mainly distributed in the Palearctic, Oriental, and Afrotropical regions (BPP: 28SN = 0.62; 28SA = 0.56). The second clade has the Neotropical Megaloprotocus Schulz at the base (only with 28SN; BPP = 0.58), followed by a clade with two species of Spathius + Caenophanes Foerster (BPP: 28SN, 28SA = 1.0), then by a clade with the two genera of the cosmopolitan Holcobraconini, Liobracon Szepligeti and Odontobracon Cameron (BPP: 28SN, 28SA = 1.0), a clade with two endemic Australian genera, Jarra Marsh & Austin and Synagaster Brullé (BPP: 28SN, 28SA = 1.0), and finally by a large, significantly supported clade represented by doryctine genera that are mainly distributed in the Neotropics (BPP: 28SN = 0.69; 28SA = 0.98).

The five gall-associated doryctine genera were recovered together in the two simultaneous analyses in a single clade within the large ‘Neotropical’ clade, and recoding of the unalignable regions increased its BPP, though not to a significant value (BPP: 28SN = 0.49; 28SA = 0.7). Moreover, a clade with all the gall associated genera excluding Percobracon appeared significantly supported in the 28SA + COI topology (BPP = 0.95). The clade with the five gall-associated genera was recovered as the sister group of a clade with the two included species of the widely distributed and polyphagous Heterospiulus, and this relationship increased its BPP to a significant value when the unalignable regions were recoded (BPP: 28SN + COI = 0.69; 28SA + COI = 0.99).

Most of the topologies derived from the MP and Bayesian analyses using the multiple alignments for the 28S with
various gap opening-gap extension penalties recovered a clade with at least four of the gall-associated genera included (Fig. 4). Moreover, all these analyses placed the taxa with elongated first metasomal tergite in separate clades along the tree. Six of the 21 MP analyses with the 28S alone recovered MPTs containing a group with the gall-associated genera except Labania, and 13 analyses reconstructed MPTs that placed them together with the latter and the Australian Jarra and Syngaster. Fourteen of the combined analyses using MP recovered the monophyly of the five gall-associated genera in at least some of their MPTs obtained, whereas 19 of the 50% majority rule consensus trees derived from the simultaneous Bayesian analyses also recovered this relationship.

3.2. Tests of alternative hypotheses

The Bayesian test of alternative phylogenetic hypotheses for the two combined analyses showed that the topologies recovering the tribes Doryctini, Hecabolini and Spathiini (sensu Belokobylskij, 1992) each as monophyletic are statis-
tically rejected \((P < 0.05)\). The SH tests also show that the three Bayesian topologies derived from the 28SN + COI data set with the above tribes forced to be monophyletic are significantly worse explanations of the data than the unconstrained Bayesian topology \((P < 0.0001)\). The Bayesian alternative topologies used for the latter test can be retrieved from the electronic supplementary material.

3.3. Character evolution

Fig. 3 shows the ancestral states of the two examined features that were mapped onto the 28SN + COI Bayesian phylogeny based on maximum likelihood optimizations. Origin of gall association appears significantly supported once in the \textit{Percnobracon + Psenobolus + Allorhogas + Labania + Monitoriella} clade. On the other hand, at least eight origins are shown for a moderately to considerably elongated acrosternite in the first tergite. This feature is shown to have evolved three times within the Rhaconotini clade \([\textit{Platyspathius Viereck, Leptorhaconotus Granger and Spathius (Antespathius)}]\), twice within the remaining members of the ‘Palaeartic-Oriental-Afrotropical’ clade with \textit{Schlettereriella Szépligeti} and \textit{Spathiomorpha Tobias}, once with the species of \textit{Spathius (Spathius Belokobylskij)}, and two times within the ‘Neotropical’ clade, one with the gall-associated \textit{Percnobracon} and \textit{Pse-
nobilis, and the remaining one in the Notiospathius Matthews & Marsh + Masonius Marsh + Tarasco + Hansonomurum Marsh clade.

The full hierarchical Bayesian method shows that the ancestral state being gall associated has a probability of 0.93 for the clade involved.

4. Discussion

Our study shows phylogenetic evidence that suggests a close evolutionary relationship among a group of gall-associated wasps within the mainly parasitic wasp subfamily Doryctinae. This relationship was recovered in our two simultaneous phylogenetic analyses, though the clade involved was only significantly supported in the 28SA + COI Bayesian phylogeny. The result obtained was surprising because the taxa involved display disparate external morphologies and are associated with distantly related plant families.

The exclusive presence of cecidogenic and presumed galler-predating taxa within a same clade agrees with the generally accepted scenario about the origin of sec-
secondary phytophagy in parasitic wasps, in which phytophagy is hypothesized to have arisen via attacking of cone-boring, galler or seed predator insect larvae in living plant tissues, with the subsequent shift to feeding on the associated particularly nutritious plant material (Ronquist, 1995; Quicke, 1997). This hypothesis is also supported by the presence of both gall-formers and predators of gallers within Allorhogas (Maceêdo et al., 1998; Wharton and Hanson, 2005). In addition, Allorhogas and Psenobolus, the latter probably a parasitoid of gallers in figs (P. Hanson, pers. comm.), are recovered as sister taxa in the 28S + COI Bayesian phylogeny (though with low BPP), whereas the cecidogenic Monitoriella and the presumed gall former Labania are significantly supported as sister taxa in both analyses.

The fact that all the known gall-forming doryctines are from the Neotropics where they attack a range of only distantly related plant families is more congruent with a model that explains the macroevolution of insect plant associations as primarily reflecting geographical proximity (Bernays and Chapman, 1994; Dobler et al., 1996), though chemical similarities between the host plants involved (Futuyma and McCafferty, 1990; Becerra and Venable, 1999) cannot be ruled out.

Similar to the separate and combined molecular analyses performed by Zaldivar-Riverón et al. (2006, Figs. 2a, b and 3a) with the same two markers, our phylogenies do not recover the Doryctinae as monophyletic, with the support of the clades involved being weak. In fact, a monophyletic Doryctinae was only recovered with a significant BPP when morphological data were included. Despite this, the relationships recovered here reveal a geographic pattern that is congruent with the sequential break up of Gondwanaland with three defined groups:

Fig. 4. Summary of groups recovered in the separate and combined analyses based on Clustal X alignments of the 28S with different gap opening and gap extension penalties. Black squares are analyses that recovered the five gall-associated genera included as monophyletic in any of the MPTs obtained and in the 50% majority rule consensus tree derived from the Bayesian method. Grey squares are analyses that recovered at least four of the above genera in a single clade. Numbers in the Bayesian analyses refer to posterior probabilities of clades. Numbers in the parsimony analyses are percentages of MPTs that grouped gall-associated genera (*, 100% of MPTs). (a) Excluding Labania; (b) excluding Percnobracon; (c) excluding Psenobolus; (d) excluding Labania and including the two species of Heterospilus; (e) excluding Labania and including the two species of Heterospilus, Syngaster, and Jarra.
one mainly composed of taxa from the Palaeartic, Oriental, and Afrotropical regions, a second one of Australian, and the remaining one of Neotropical taxa, with the latter two being significantly supported as sister groups.

The presence of suites of characters associated with adaptations of parasitism has commonly led to distantly related wasp taxa being classified together (Achterberg, 1988; Quicke, 1997; Quicke and Belshaw, 1999). Our phylogenetic analyses have consistently shown that one of the main morphological features employed to distinguish the Spathiini, one of the largest tribes of the subfamily, has evolved on numerous separate occasions. We suggest that an extended first metasomal tergite in distantly related lineages probably reflects convergence as a result of a shared life history strategy: the attack of concealed host larvae, as in the case of wood- and bark-boring coleopterans through enabling use of a longer ovipositor while maintaining a near vertical ‘drilling’ posture. Finally, the results of this study also support a previous work that suggests the existence of high levels of morphological homoplasy among the different doryctine genera (Belokobylskij et al., 2004b), which can be inferred by the high disparity between the morphological classification of the group and our resulting hypotheses of phylogenies. This high level of morphological homoplasy is also observed by the significant differences found between the likelihood of the 28SN + COI phylogeny and the likelihood of alternative topologies that enforced the larger doryctine tribes Hecabolini, Spathiini, and Doryctini each as monophyletic.

The significantly supported relationships presented in our study should therefore be used to construct an alternative higher classification in the Doryctinae.

Acknowledgments

We thank Kees van Achterberg, Yves Braet, José Clavijo, Brian Fisher, Georg Georgen, Carolina Godoy, Paul Hanson, Nina Laureme, John Noyes, and Mark Shaw for providing material for this study. This work was in part supported by a Natural Environment Research Council Grant NE/C519538/1 to DLJQ (and Andy Purvis), a post-doctoral fellowship given by the Universidad Nacional Autónoma de México (UNAM-DGAPA) to AZR, a Grant given by the Russian Foundation for Basic Research (No. 07-04-00454) to S.A.B., a PhD scholarship given by the Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) to J.J.M., and by two Grants given by FONACIT-Venezuela (S1-2000000479, LAB-2000001593) to R.B.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.05.016.

References


