Research article

The phylogeny of the social wasp subfamily Polistinae: evidence from microsatellite flanking sequences, mitochondrial COI sequence, and morphological characters

Elisabeth Arévalo*1, Yong Zhu2, James M Carpenter3 and Joan E Strassmann4

Address: 1Department of Biology, Providence College, Providence, RI, 02918-0001, USA, 2Department of Epidemiology and Public Health, Yale University, 60 College Street, P.O. Box 208034, New Haven, CT 06520-8034, USA, 3Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192, USA and 4Department of Ecology and Evolutionary Biology, MS 170, Rice University, 6100 Main St., Houston, TX 77005-1892, USA

Email: Elisabeth Arévalo* - earevalo@providence.edu; Yong Zhu - Yong.Zhu@yale.edu; James M Carpenter - carpente@amnh.org; Joan E Strassmann - strassm@rice.edu

* Corresponding author

Abstract

Background: Social wasps in the subfamily Polistinae (Hymenoptera: Vespidae) have been important in studies of the evolution of sociality, kin selection, and within colony conflicts of interest. These studies have generally been conducted within species, because a resolved phylogeny among species is lacking. We used nuclear DNA microsatellite flanking sequences, mitochondrial COI sequence, and morphological characters to generate a phylogeny for the Polistinae (Hymenoptera) using 69 species.

Results: Our phylogeny is largely concordant with previous phylogenies at higher levels, and is more resolved at the species level. Our results support the monophyly of the New World subgenera of Polistini, while the Old World subgenera are a paraphyletic group. All genera for which we had more than one exemplar were supported as monophyletic except Polybia which is not resolved, and may be paraphyletic.

Conclusion: The combination of DNA sequences from flanks of microsatellite repeats with mtCOI sequences and morphological characters proved to be useful characters establishing relationships among the different subgenera and species of the Polistini. This is the first detailed hypothesis for the species of this important group.

Background

A robust phylogenetic tree is of paramount importance as a beginning point for many kinds of evolutionary studies [1-4]. Understanding the evolutionary history of a character can allow us to tease apart factors important for adaptation to a specific environmental condition from those due to history [5,6]. However, obtaining accurate phylogenetic relationships can sometimes be difficult. Historically, morphological traits were used in combination with the principles of cladistics [7,8] to elucidate phylogenetic relationships. Morphological traits require specialists who can detect sometimes minute differences between species, and such expertise requires years of training.

Nucleic acid sequence information, primarily from mitochondrial DNA (mtDNA) has allowed us to make enormous strides towards resolving phylogenies during the last twenty years [9-15]. But these sequence regions have
the disadvantage that they are often not evolving at ideal rates to reveal histories of desired traits. The branching pattern of gene trees and species trees are not necessarily overlapping in time and proceed concurrently. Also, evolution of mtDNA genes might not necessarily correspond to the speciation process [4]. Mitochondrial DNA genes are essentially single linkage groups and might not reveal the actual whole organism phylogeny [16,17].

For nearly as long as sequence data have been available, there has been a heated debate about which kinds of data, molecular or morphological, provide the most accurate result when topologies obtained from different data sets are in conflict. Molecular data are undoubtedly considered by many to be more reliable than morphology (i.e., more likely to produce a "true" tree, e.g. [18]). It is clear that the perception of morphological data as inferior to molecular data also extends to value judgments on the quality of a particular morphological analysis based on how well the results of that analysis conform to the results of molecular analyses (e.g. [19]). Such standards are usually not applied to independent molecular analyses, which are also often in conflict with each other (e.g. [20]). No single particular data set may completely conform with our best estimate based on total evidence. The best estimator for a species phylogeny might therefore not necessarily depend on the sole use of one set of characters (morphological or molecular), but on a combination of different characters that can prove to be a better estimator of species trees [8,13,21-24].

Nuclear markers are independent from mitochondrial markers, and can be independent from each other if they are scattered through the genome. Nuclear genes represent a largely untouched resource for molecular systematics [25]. However, few nuclear genes, besides some rDNA sequences, have been used in phylogenetic studies [26]. The earliest molecular studies in Hymenoptera focused almost entirely on mitochondrial 16S rDNA [27-32]. More recently, the nuclear 28S rDNA gene has been used successfully for analyses of relationships within superfamilies, but with less success at higher taxonomic levels [33-36]. It has also been applied to social wasps [37], but its use is controversial [38]. Microsatellite loci are abundant, single copy, and widely distributed throughout the genome [39-41]. Few studies have used microsatellites to reconstruct phylogenies although they have been widely used for population genetic studies (e.g. [42-46]). Though their sequences are short and the repeat regions themselves are highly variable, sequence information from several microsatellite flanking regions can be combined for robust, informative and easily aligned characters that can help resolve difficult phylogenetic relationships. The single copy flanking sequences are much less variable than are numbers of repeats themselves, making them easier to align.

Here we present a phylogeny for a very important group of social wasps. Social insects of the order Hymenoptera constitute a highly diverged group, with cosmopolitan distributions. Because of the existence of different castes, division of labor, altruism, and a haplodiploid genetic system, these social insects represent a unique system to study kin selection [47-50]. The subfamily Polistinae (paper wasps) is the second largest of six subfamilies of the family Vespidae, and is entirely composed of social wasps. It is made up of four tribes (Polistini, Epiponini, Mischocyttarinii and Ropaliiini), with 943 species ([51], unpublished). This group probably arose in the mid-to late Jurassic. Carpenter [51] has estimated that this group diverged from other Vespidae at the breakup of Gondwana, as much as 140 million years ago. A molecular study, which used microsatellite loci to establish the conservation and polymorphisms across different species of this subfamily, discovered interesting results for divergence times among species [52]. The divergence time for *Polistes bellicosus* and *P. annularis* was estimated to be between 10 to 80 million years. The divergence time for the entire tribe is larger than among the species of the genus *Polistes* and smaller than for the subfamilies of Vespidae (between 80 and 175 million years), and Vespinae and Polistinae probably split in the Mid Jurassic [52]. The tribe Polistini is the only one with a cosmopolitan distribution (Table 1). The genus *Polistes* was formerly subdivided into 12 different subgenera, either distributed in the New World (*Fuscopolistes, Aphanilopterus, Palisotius, Epicenemius, Onerarius*) or Old World (*Gyrostoma, Stenopistes, Nygmopolistes, Megapolistes, Polistella, Sulcopistes* and *Polistes sensu stricto*). More recently, seven of these subgenera have been synonymized [53], with one subgenus for the New World species (*Aphanilopterus*) and three for the Old World species (*Gyrostoma, Polistella and Polistes sensu stricto*). The New World subgenera were synonymized because two were found to be paraplectic, a result not supported by the present study. The other three tribes, on the other hand, are either distributed in the New World (Epiponini, Mischocyttarini) or the Old World (Ropaliiini).

The subfamily Polistinae is characterized by two major behavioral groups, based on their mode of colony founding and mechanism of reproductive dominance [54,55]. Independent founding is characterized by small, simply constructed nests without a protective paper envelope, founded by a single queen or few queens, without the assistance of workers [56,57]. By contrast, swarm founding species have large colonies initiated by swarms consisting of multiple queens and a large number of workers [55].
Table 1: List of Tribe (including subgenera, in the case of the Polistini tribe), species, locality and collector for all the specimens used in the study.

| Tribe/Subgenera         | Species               | Locality of Origin | Donor |
|-------------------------|-----------------------|--------------------|-------|
| Polistini/Fuscopolistes | Polistes bellicosus   | USA                | EA    |
| Polistini/Fuscopolistes | Polistes apachus      | USA                | JMC   |
| Polistini/Fuscopolistes | Polistes aurifer      | USA                | JMC   |
| Polistini/Fuscopolistes | Polistes carolina     | USA                | PS/ES/DCQ |
| Polistini/Fuscopolistes | Polistes dorsalis     | Texas              | EA    |
| Polistini/Fuscopolistes | Polistes metricus     | USA                | EA    |
| Polistini/Fuscopolistes | Polistes fuscatus     | USA                | JMC   |
| Polistini/Aphanilopterus| Polistes annularis    | USA                | JES/DCQ |
| Polistini/Aphanilopterus| Polistes buysoni      | Chile              | JMC   |
| Polistini/Aphanilopterus| Polistes canadensis   | Venezuela          | JES/DCQ |
| Polistini/Aphanilopterus| Polistes lanio        | Argentina          | JMC   |
| Polistini/Aphanilopterus| Polistes cavayptia    | Argentina          | JMC   |
| Polistini/Aphanilopterus| Polistes exclamans    | USA                | EA    |
| Polistini/Aphanilopterus| Polistes similimus    | Bolivia            | JMC   |
| Polistini/Aphanilopterus| Polistes crinitus americanus | Puerto Rico | JMC   |
| Polistini/Aphanilopterus| Polistes versicolor   | Venezuela          | JES/DCQ |
| Polistini/Aphanilopterus| Polistes instabilis   | S. Texas           | JES/DCQ |
| Polistini/Palisotius    | Polistes major        | USA                | JMC   |
| Polistini/Epicnemius    | Polistes cinerascens  | Argentina          | JMC   |
| Polistini/Nympopolistes | Polistes tenereciosus | Malaysia           | JMC   |
| Polistini/Megopolistes  | Polistes rothneyi     | Japan              | JMC   |
| Polistini/Megopolistes  | Polistes jakahamae    | FL                 |       |
| Polistini/Polistella    | Polistes stigmus      | Australia          | JMC   |
| Polistini/Polistella    | Polistes snelleni     | Japan              | JMC   |
| Polistini/Polistella    | Polistes japonicus    | FI                 |       |
| Polistini/P. sensu stricto | Polistes nimpus      | Cavriglia          | EA, FZ |
| Polistini/P. sensu stricto | Polistes chinesis     | FI                 |       |
| Polistini/P. sensu stricto | Polistes dominulus   | Italy              | EA, FZ |
| Polistini/P. sensu stricto | Polistes biglumis bimaculatus | M. Genevre | MCL   |
| Polistini/P. sensu stricto | Polistes gallicus    | Florence           | EA, FZ |
| Polistini/P. sensu stricto | Polistes marginalis  | South Africa       | JMC   |
| Mischocyttarini         | Mischocyttarus alfenii | Venezuela         | JES/DCQ |
| Mischocyttarini         | Mischocyttarus immarginatus | Costa Rica | JMC   |
| Mischocyttarini         | Mischocyttarus melanarius | La Trinidad | UM    |
| Mischocyttarini         | Mischocyttarus mexicanus | USA              | JMC   |
| Mischocyttarini         | Mischocyttarus pallidpectus | Costa Rica | JMC   |
| Mischocyttarini         | Mischocyttarus phthisicus | Puerto Rico | JMC   |
| Mischocyttarini         | Mischocyttarus mastigophorus | Costa Rica | JMC   |
| Ropalidiini             | Belonogaster juncea colonialis |         |       |
| Ropalidiini             | Parapolybia varia     | ST                 |       |
| Ropalidiini             | Ropalida latebalteata | Malaysia          | JMC   |
| Ropalidiini             | Ropalida socialistica | FI                 |       |
| Ropalidiini             | Ropalida fasciata    | FI                 |       |
| Ropalidiini             | Ropalida sp.         | Malaysia          | ST    |
| Epiponini               | Apoica pallens       | Venezuela          | JES/DCQ |
| Epiponini               | Agelaia multipicta   | Venezuela          | JES/DCQ |
| Epiponini               | Angiopolybia pallens | Peru               | CRS   |
| Epiponini               | Parachartergus colobopterus | Venezuela | JES/DCQ |
| Epiponini               | Parachartergus fraternus | Venezuela | JES/DCQ |
| Epiponini               | Protopolybia exigua  | Venezuela          | CRH   |
| Epiponini               | Protonectarina sylveireae | Brazil           | JES/DCQ |
| Epiponini               | Brachygastra augusti | Venezuela          | JES/DCQ |
| Epiponini               | Brachygastra mellifica | South Texas | JES/DCQ |
| Epiponini               | Brachygastra lecheguana | Argentina     | JMC   |
| Epiponini               | Brachygastra bilineolata | Trinidad      | JMC   |
| Epiponini               | Polybia occidentalis | Venezuela          | JES/DCQ |
The independent-founding group is made up of five genera: Polistes, Mischocyttarus, Belonogaster, Parapolybia, and some species of Ropalidia. Colonies are initiated by one or several inseminated egg-layers and no workers. Egg layers go through a solitary phase after insemination (few days to months). Soon after nest initiation, one fertilized female becomes the sole egg-layer. These groups of species are characterized by small colony sizes (rarely more than 50 adult wasps) and the lack of a protective envelope on the nest [56,58,59]. Reproductive dominance in this group is at least partly maintained through aggression by the queen.

The swarm-founding group is made up of 20 genera of the tribe Epiponini, the genus Polybioides, and some Ropalidia species. Colonies are initiated by a large number of workers and a small number of egg-layers. There is no colony phase without workers. Workers choose the site and built the nest while egg-layers wait. Swarm founding wasps characteristically have large colony sizes (usually over 50 adult wasps), and nests are often provided with a protective envelope. Reproductive dominance in this group is obtained through worker aggression on potential queens, and there is a tendency towards morphological castes in some but not most species [50,55].

Many behavioral, ecological and evolutionary questions in this group could be asked in a phylogenetic framework including characteristics of nest structure [5,60,61], reproductive strategies [62-64], and solutions to potential conflicts of interest among colony members [57,65,66].

For the present study, we used a combination of nuclear (three microsatellite loci flanking regions and a coding system generated by the different repeat motifs, with priming sites largely conserved throughout the Polistinae subfamily; [52], adult morphological [2,53,67], larval [68-76], behavioral [77] and mitochondrial (COI) characters to obtain a better understanding of the relationships in this group. A study by Zhu et al. [78] used a subset of the characters and species presented in this study. The use of these characters was to reconstruct the evolution of microsatellite repeats within the Polistinae subfamily. For the completion of that study, the best tentative tree available at that time was used in order to establish the evolution of these microsatellite repeats in an explicitly phylogenetic perspective.

With this study we investigate two phylogenetic hypotheses. First, we wanted to test whether or not the subgenera of Polistes of Richards [79] are monophyletic. For this, we used 33 species of Polistes and 36 species in the outgroup, the Epiponini, Ropalidiini and Mischocyttarini. Second, we evaluated the phylogenetic relationships within and among the other three Polistinae tribes, Epiponini, Ropalidiini and Mischocyttarini, using the 33 species of the tribe Polistini as an outgroup.

### Results

#### Polistini

Morphological characters alone showed little resolution in the Polistini at the species level, instead supporting most of the subgenera as monophyletic – but not all (Fig. 1). New World species as a whole were monophyletic with a bootstrap support of 75% while Old World species were paraphyletic. Of the subgenera, Polistes sensu stricto, Megapolyistes, Polistella, Epicnemius and Fuscopolyistes were all supported as monophyletic with bootstraps of 64%, 95%, 74%, 63% and 61% respectively. But there was no resolution within these groups, and Aphanilopterus was not resolved as a group.

The COI, and microsatellite combined characters supported the subgenera Megapolyistes, Fuscopolyistes, and Aphanilopterus with bootstraps of 93%, 93%, and 70% respectively (Figure 2). Both Polistella and Polistes sensu stricto were mostly supported at 50% and 66%, but each left out one species. There was also considerable resolution within the supported subgenera. Thus, the independent molecular and morphological trees offered resolution in both different and overlapping parts of the tree and, where they were resolved, were not in conflict.

| Tribe       | Species          | Locality       |
|-------------|------------------|----------------|
| Epiponini   | Polybia ignobilis| Venezuela      |
| Epiponini   | Polybia emacita  | Venezuela      |
| Epiponini   | Polybia rejecta  | Venezuela      |
| Epiponini   | Polybia ruficeps | Argentina      |
| Epiponini   | Polybia scrobilis| French Guiana  |
| Epiponini   | Polybia affinis  | French Guiana  |
| Epiponini   | Synoeca septentrionalis | Venezuela |
| Epiponini   | Metapolybia cingulata | Venezuela |
| Epiponini   | Epipona niger    | Venezuela      |
Bootstrap analysis of morphological characters with 1000 replications for the 33 species of the Polistini, using 36 species of Mischocyttarini, Ropalidiini, and Epiponini as outgroups.

Figure 1
Bootstrap analysis of morphological characters with 1000 replications for the 33 species of the Polistini, using 36 species of Mischocyttarini, Ropalidiini, and Epiponini as outgroups.
The overall combined tree topology (Figure 3) is much more resolved than what was observed with the independent-character phylogenies alone (Figs. 1 and 2). The New World clade is monophyletic with 62% bootstrap
support. Although the support for this same clade was 75% in the morphological characters tree (Fig. 1), the combined tree provides a much more resolved phylogeny at the species level. The clade for the Old World Polistes is paraphyletic. All the subgenera for which we have 2 or more exemplars are supported. The Old World subgenera, Megapolystimus, Polistella, and Polistes sensu stricto are supported with bootstraps of 99%, 93% and 92% respectively. The New World subgenera, Fuscopolistes, Palisotius, and Epicnemius are supported with bootstraps of 99%, 96% and 70% respectively.

**Epiponini, Mischocyttarini, Ropalidiini**
Morphological characters alone supported the Old World Ropalidiini and the New World Epiponini and Mischocyttarini with bootstraps of 64%, 85% and 91% respectively (Figure 4). All of the genera for which we had more than one species were also supported except for Polystia. There was little resolution within genera except for Brachygastra. However most genera were only represented by a few species.

The combined COI and microsatellite characters were largely uninformative on their own, though there was some grouping of epiponine species and resolution within the monophyletic Mischocyttarins (Figure 5).

The combined tree is much more resolved than either separate analysis, and bootstrap supports are greater (Figure 6). Monophyly of Ropalidiini, Mischocyttarini, and Epiponini is supported at bootstrap levels of 83%, 95%, and 90% respectively. However the relationships between the tribes are not resolved. All genera for which we have more than one species are supported except for Polystia which is broken into three groups, among which relationships are not resolved (Figure 6). Relationships among genera within Ropalidiini and Epiponini are generally not well resolved. The basal genus of the Epiponini is unresolved.

Figure 3
Bootstrap analysis of all characters with 1000 replications for the 33 species of the Polistini, using 36 species of Mischocyttarini, Ropalidiini, and Epiponini as outgroups.
Bootstrap analysis of morphological characters with 1000 replications for 36 of the species of Mischocyttarini, Ropalidiini, and Epiponini using 33 species of Polistini as outgroups.

**Figure 4**
Bootstrap analysis of morphological characters with 1000 replications for 36 of the species of Mischocyttarini, Ropalidiini, and Epiponini using 33 species of Polistini as outgroups.
Bootstrap analysis of molecular characters with 1000 replications for the 36 species of Mischocyttarini, Ropalidiini, and Epiponini using 33 species of Polistini as outgroups.

**Figure 5**
Bootstrap analysis of molecular characters with 1000 replications for the 36 species of Mischocyttarini, Ropalidiini, and Epiponini using 33 species of Polistini as outgroups.
Discussion
Relationships in the Polistini
The combined character phylogeny is partly congruent with the earlier phylogeny proposed by Carpenter in 1996, though there are some differences (Figure 7). First, note that the earlier tree is not bootstrapped; Figure 1 is the bootstrapped tree for morphological characters, based on the present sample of species. Results from both studies on the monophyletic origin of the New and Old World Polistini subgenera only found support for the monophyletic origin of the New World Polistini (Fig. 3). In both studies, the Old World Polistini subgenera represent a paraphyletic group. However, the present study showed Polistes sensu stricto as the most basal subgenus (with a bootstrap support of 98%), in contrast to Carpenter [53], where it was the sister group of the New World subgenera (with a bootstrap support of 89%, Fig. 7). The present study found the sister group of the New World subgenera to be the clade of Polistella, Nygmopolistes, and Megapolistes. Several subgenera were absent from the present study (viz. Gyrostoma, Stenopolistes and Sulcopolistes) and it is possible that their absence affected character polarizations obtained within Polistes.

In contrast to Carpenter’s [53] study, the present results support monophyly of the subgenus Aphanolopterus in the strict sense, and of the subgenus Epicnemius, albeit based on a smaller number of species.

Polistes sensu stricto is unique in the Polistinae for having social parasites, species that kill their host queen so her workers can rear the brood of the social parasite [12,80-82]. The three social parasite species are monophyletic and most closely related to P. nimphus and P. dominulus, and less close to P. gallus and P. biglumis [12,82]. The present study agrees that the latter two species are sister
species, less close to *P. nimphus* and *P. dominulus*, but this study does not put these two as sister species. The earlier study did not include morphological characters, and did not include either *P. chinensis* or *P. marginalis* which could account for the differences [82].

**Figure 7**
Comparison of the subgenera of *Polistes* phylogeny obtained by Carpenter’s 1996 morphological study (Fig. 2.8) and the phylogeny obtained in the present study, including all morphological and molecular characters. Species included from each subgenus are listed in Table 1.

**Relationships in the Epiponini, Mischocyttarini, Ropalidiini**
The combined character phylogeny is largely congruent with the earlier phylogeny proposed by Carpenter in 1991 (Figure 8). Depending on the tribe, one or the other is
more resolved, but not in major disagreement. Within the tribe Epiponini, of the polytomies observed in the Carpenter’s 1991 topology, one was resolved: *Agelaia* and *Angiopolybia* are sister-groups (Fig. 8). The genus *Polybia* in the present study does not disagree with Carpenter et al.’s [67] conclusion that it was paraphyletic: in the present study it is unresolved, with *P. (Formicicola) rejecta*, *P. (Trichinothorax) affinis* and *P. (Pedotheoca) emaciata* sepa-
rated from the main \textit{P. sensu stricto} clade for the genus. Note that the present study is based on a broader sample of genera than in [67].

In a series of papers Strassmann and Queller argued that a form of worker control of sex ratios was general for the Epiponini [40,83]. This form of worker control meant that colonies did not produce a new generation of queens until the colony had reduced in queen number to one or perhaps no queens. This is important to sociality because such a queen reduction before requeening maintains high relatedness within colonies. The species that they studied this cyclical oligogyyn in were \textit{Parachartergus colobopterus}, \textit{Protopolybia exigua}, \textit{Brachygastra mellifica}, \textit{Polybia emaciata} and \textit{Polybia occidentalis} [63,83-86]. In addition West-Eberhard documented either cyclical oligogyny or some of its features in three additional genera, \textit{Agelaia}, \textit{Synoeca} and \textit{Metapolybia} [87-89]. Taken together, these studies come from all of the four basal-most lineages of the Epiponini, and also cover 4 of the 6 lineages in our best sampled clade, the one containing \textit{Polybia} (Figure 7). Thus, it seems they are justified in claiming the generality of this form of worker control for the tribe.

\textbf{Importance of molecular and morphological characters}

The inclusion of several genetic markers (mitochondrial and nuclear), improved the resolution of the trees obtained from morphology alone. The COI sequence alone did not resolve \textit{Polistes} as a group; combining it with the microsatellites did. While COI and the microsatellite flanking regions did not resolve \textit{Polistella}, adding the repeat regions did so. The use of microsatellite flanking regions and the presence/absence coding system for the repeated regions of the microsatellites proved to be useful in the resolution of some of the branches.

For the last fifteen years, a considerable rise in the use of molecular markers has been observed. The majority of the studies have exclusively included molecular data from mtDNA genes. Even when the information gathered for those studies has been informative, problems with the use of this type of molecular data (saturation synonymous sites with mtDNA; constrains imposed on RNA sites by secondary structure, length variation, etc.) have been considered [25,90]. Unfortunately, few nuclear markers have been used in phylogenetic systematic studies in the past. The search for new, powerful, informative-single copy nuclear markers has increased our wealth of information in systematic studies. At the present time just a few more nuclear markers have been used for insect phylogenetic studies [26,91]. More genes are needed for testing phylogenetic hypotheses. Meanwhile, the use of microsatellite flanking regions represents a valuable tool as nuclear single-copy markers. The wealth of microsatellite data is constantly growing across different taxa. The use of these regions will represent a powerful molecular tool in future studies.

\section*{Methods}

\textbf{Species and specimens used}

We used a total of 69 species representing the four tribes of the subfamily Polistinae (Table 1). Specimens were either donated by colleagues, or were in our own collections. These specimens included 33 species from the tribe Polistini, 7 from the Mischocyttarini, 7 from the Ropalidiini and 22 from the Epiponini. We used the Polistini as the outgroup for the Epiponini+Mischocyttarini+Ropalidiini. We used the latter three tribes as the outgroup for the analysis of the Polistini.

\textbf{Morphological and molecular characters}

We used a total of 932 characters, including 95 morphological and behavioral characters (all morphological characters and character states from Carpenter’s studies [2,53,67] examined on pinned specimens for this study; larval and behavioral characters and states from citations given in appendixes 1 and 2); a 488 base pair (bp) fragment of the mtDNA COI gene (appendix 3) [92]; a 311 bp fragment of single-copy nuclear flanking sequence of three microsatellite loci (Pbe216AAG [137 bp], Pbe269bAAG [75 bp], and, Pbe411AAT [99 bp] (appendix 4) [52,93], and 38 characters representing the different trinucleotides in the repeat region (not their number; appendix 5 [78]). The repeat regions varied considerably across the four different tribes and across the different subgenera in the tribe Polistini. The coding system for the repeat region was based on presence or absence of each repeat motif and resulted in 16 characters for Pbe216AAG, 15 characters for Pbe269bAAG, and 7 characters for Pbe411AAT. For example, the first four characters at Pbe216AAG were GGA, AAA, TAA, GCA, all substituted for AAG within the repeat region (not their number; appendix 5 [78]). We did not count times a specific motif was repeated because this character is so variable its evolutionary history would be unclear.

\textbf{DNA extraction, amplification and sequencing for COI, and microsatellites}

We extracted DNA using protocol Strassmann I [94]. Prior to DNA extractions, we stored all specimens either at -80°C or in 100% ethanol at room temperature. PCR reactions followed protocols as in Wilcox et al. [95] for the mtDNA fragment, and Zhu et al. [78] for the microsatellite loci. Sequencing PCR products were completed directly with thermo sequenase radiolabeled \textsuperscript{33}P (Thermo Sequenace radiolabeled terminator cycle sequencing Kit, Amersham Life Science). DNA sequences were run on denaturing polyacrylamide gels, dried and exposed to X-ray film. All new DNA microsatellite
sequences were deposited in Genbank under accession numbers AF119623-AF119660, AF120355-AF120455. COI sequences were deposited in Genbank under accession numbers AY382200-AY382265. The aligned DNA matrix is available at TreeBASE http://www.treebase.org/treebase/index.html, submission ID number SN1595.

**Sequence alignment**

The sequences for both mtDNA (COI) and for nuclear regions (microsatellite flanking regions) were aligned based on the criterion of maximum parsimony, using the program Malign, specifying a 3:1 change:gap ratio [96]. We occasionally adjusted alignments by eye when Malign gave poor results.

**Systematic analysis techniques**

We used Paup* 4.0b10 [97], Nona [98], and Winclada [99] to construct the phylogeny, with equal weight for all base-pair substitutions (including the mtDNA COI sequence, [100]). For the COI gene, the analysis was completed with the use of the three coding positions weighted equally and with the first and second coding positions only. The phylogenetic trees were not in conflict by excluding the third coding position.

We completed the phylogenetic analyses following the principles of maximum parsimony, maximum likelihood and distance analysis. As the overall topology was not in disagreement by applying different algorithms, we are only presenting the results obtained based on maximum parsimony. All parsimony analyses were run using approximate searches, due to the large number of species and characters. The runs consisted of 1000 replications of the parsimony ratchet [101] as implemented in Winclada, with 10% character reweighting and holding one tree; 500 replications at the "default amb-poly-" setting and 500 at "amb = poly-" (TBR branch swapping). Nona was also used to carry out 1000 random addition sequences as a check. Maximum parsimony trees were first estimated individually for each of the different data sets (morphological and molecular), and then in combination. Gaps were treated as missing characters. Bootstrap analyses with 1000 replications were also run for each data set. Paup* runs gave similar results as those with Winclada and Nona.

Table 2 shows the number of parsimoniously informative characters and constant characters for the each different set of data used in the study, for the 69 species used.

**Authors' contributions**

EA carried out a portion of the molecular genetic studies, participated in the sequence alignment, analysis of the data, and drafted the manuscript. YZ carried out a portion of the molecular genetic studies and participated in the sequence alignment. JMC provided the morphological data, and drafted the manuscript. YZ carried out a portion of the molecular genetic studies, and participated in the analyses and writing. All authors read and approved the final manuscript.

**Acknowledgements**

We are thankful to all our colleagues who donated specimens to the study (see Table 1). We also thank Barry Sullender for his assistance in the analysis of the characters and David Queller for his insightful comments. EA wants to thank Liam Donohoe for comments on early drafts of the manuscript. This material is based upon work supported by the National Science Foundation under Grant Nos. BIR-9612687, IBN-9808809, IBN-9900975, and DEB-9975351.

**References**

1. Bidrêgô N, Cracraft J: Phylogenetic Patterns and the Evolutionary Process. Columbia University Press, NY; 1980.
2. Carpenter JM: Phylogenetic Relationships and the Origin of Social Behavior in the Vespinae. In: The social biology of Wasps Cornell University Press, Ithaca, NY; 1991:7-32.
3. Cameron SA, Mardulyn P: Multiple molecular data sets suggests independent origins of highly eusocial behavior in bees (Hymenoptera:Apinae). Syst Biol 2001, 50:194-214.
4. Nichols R: Gene trees and species trees are not the same. Trends Ecol Evol 2001, 16:358-364.
5. Wenzel JW: Evolution of Nest Architecture. In: The social biology of Wasps Cornell University Press, Ithaca, NY; 1991:7-32.
6. Brooks DR, Mcleman DA: Systematics, ecology, and behavior. Bioscience 1995, 45:687.
7. Henning W: Phylogenetic Systematics. University of Illinois Press Urbana; 1966.
8. Hillis DM, Wiens JJ: Molecules versus morphology in systematics. In: Phylogenetic Analysis of Morphological Data Smithsonian Institution Press; 2000.
9. Brown WM: Evolution of Animal Mitochondrial DNA. In: Evolution of Genes and proteins Sinauer, Sunderland, MA; 1983:62-88.
10. Cameron SA, Mardulyn P: Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. Proc Natl Acad Sci 1993, 90:8687-8691.
11. Arévalo E, Davis SK, Sites JW: Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosomes races of the Sceloporus grammicus complex (Phrynosomatidae) in central Mexico. Syst Biol 1994, 43:387-418.

12. Choudhary M, Strassmann JE, Queller DC, Turillazzi S, Cervo R: Social parasites in polistine wasps are monophyletic: implications for symbiotic speciation and the evolution of wasp sociality. Proc Natl Acad Sci USA 1999, 96:25731-35.

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

14. Hillis DM, Moritz C, Mable BK: Molecular Systematics. Sinauer, Sunderland, MA, 1996.

15. MTR Frost DR, Grant T, Titus TA: Phylogenetics of the lizard genus Tropidurus (Squamata: Tropiduridae: Tropidurini): direct optimization, descriptive efficiency, and sensitivity analysis of congruence between molecular data and morphological evidence. Cladistics 1998, 14:141-147.

16. Richard M, Thorpe RS: Can microsatellites be used to infer phylogenies? Evidence from population affinities of the Western Canary Island lizard. Mol Phy Evol 2001, 20:

17. Schmitz J, Ohme M, Suryasentosa A, Suryasentosa E: Are the fossil data really at odds with the molecular data? Morphological evidence for cetartiodactyla phylogeny reexamined. Systematic Biology 2001, 50:444-453.

18. Graur D: Evolutionary relationships of the insect order Hymenoptera. Mol Phylogenet Evol 1996, 5:221-241.

19. Wieners J, Singlauf JL: How lizards turn into snakes: A phylogenetic analysis of body-form evolution in anguid lizards. Evol 2001, 55:2303-2318.

20. Gatesy J, Amato G, Moritz C, DeSalle R, Hayashi C: Molecular Phylogenetics and Evolution 2002, 1999:

21. Nei M: Molecular Evolutionary Genetics. Columbia University Press, NY, 1987.

22. Dixon J, Carpenter JM: On simultaneous analysis. Cladistics 1996, 12:221-241.

23. Arévalo E, Cervo R, Zacchi F, Turillazzi S: Workers produce males on the ground (Hymenoptera, Vespidae) as evidence of highly variable microsatellites. Mol Ecol 1995, 4:389-393.

24. Pamilo P: Genetic relatedness of lineages from size information. Mol Phylogenet Evol 2002, 254-259.

25. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

26. Derr JN, Davis SK, Woolley JB, Wharton RA: Variation and the allocation of kin selection in the wasp (Vespidae: Polistinae). J Hered 1997, 88:113-114.

27. Tools of the trade for inferring allelic ancestry. Mol Phylogenet Evol 2001, 19:119-134.

28. Queller DC, Strassmann JE, Hughes CR: Microsatellites and kinship. Trends Ecol Evol 1993, 8:285-292.

29. Choudhary M, Strassmann JE, Solis CR, Queller DC: Microsatellite variation in social insects. Biochem Genet 1993, 31:87-95.

30. Queller DC, Strassmann JE, Hughes CR: Genetic relatedness in colonies of tropical wasps with multiple queens. Science 1988, 242:1153-1157.

31. Estoup A, Scholl A, Pouvreau A, Solignac M: Monoandry and polyandry in bumble bees (Hymenoptera, Bombinae) as evidenced by highly variable microsatellites. Mol Ecol 1995, 4:389-393.

32. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

33. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

34. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

35. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

36. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

37. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

38. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

39. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

40. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.

41. Pamilo P, Crozier RH, Van Oven M: A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinate isozymes in anguid lizards (Sauria: Sauridae). Mol Phylogenet Evol 1996, 13:594-604.
