Multigene phylogeny reveals eusociality evolved twice in vespid wasps

Heather M. Hines†, James H. Hunt†, Timothy K. O’Connor†, Joseph J. Gillespie§¶, and Sydney A. Cameron†

†Department of Entomology, University of Illinois at Urbana–Champaign, Urbana, IL 61801; ‡Department of Biology, University of Missouri, St. Louis, MO 63121; §Virginia Bioinformatics Institute at Virginia Tech, Blacksburg, VA 24061; and ¶Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved December 5, 2006 (received for review November 27, 2006)

Eusocial wasps of the family Vespidae are thought to have derived their social behavior from a common ancestor that had a rudimentary caste-containing social system. In support of this behavioral scenario, the leading phylogenetic hypothesis of Vespidae places the eusocial wasps (subfamilies Stenogastrinae, Polistinae, and Vespinae) as a derived monophyletic clade, thus implying a single origin of eusocial behavior. This perspective has shaped the investigation and interpretation of vespid social evolution for more than two decades. Here we report a phylogeny of Vespidae based on data from four nuclear gene fragments (18S and 28S ribosomal DNA, abdominal-A and RNA polymerase II) and representatives from all six extant subfamilies. In contrast to the current phylogenetic perspective, our results indicate two independent origins of vespid eusociality, once in the clade Polistinae+Vespinae and once in the Stenogastrinae. The stenogastrines appear as an early diverging clade distantly related to the vespines and polistines and thus evolved their distinctive form of social behavior from a different ancestor than that of Polistinae+Vespinae. These results support earlier views based on life history and behavior and have important implications for interpreting transitional stages in vespid social evolution.

Hymenoptera | molecular phylogenetics | social behavior | social insect evolution | Vespidae

Insect societies are among the most complex systems in nature. Although diverse in kind, the most distinctive and widely known insect societies are those characterized by a reproductive “queen” and an effectively sterile “worker” caste that assists in raising the next generation of reproductive offspring. These eusocial societies (1), which comprise a small but ecologically successful fraction of insect species, are believed to have evolved several times independently within the insect order Hymenoptera: once in ants (2), once in wasps of the family Sphecidae (3), once in wasps of the family Vespidae (4, 5), and several times within bees (6, 7). What were the circumstances that gave rise to eusociality (8–11), and is this rare state inevitably the result of a stepwise progression through the transitional stages, from solitary nesting through “primitive” sociality (12–14) to a derived state of eusocial complexity (15, 16)? To answer these intriguing questions requires, minimally, two conditions: (i) the existence of a socially diverse group of taxa, and (ii) a robust estimate of their phylogeny to provide the historical framework from which to investigate changes from solitary to social behavior (17).

Vespidae is one of the few groups that has retained the necessary transitional stages to elucidate social evolution, encompassing solitary, presocial, facultatively eusocial (18), and eusocial taxa (17, 19). There are ~4,200 described vespid species currently classified into 6 subfamilies based on morphological evidence (4, 5). Euparagiinae (9 species, found only in southwestern North America and northern Mexico) (20) and Masarinae (“pollen wasps,” ~300 species) (21) are solitary; Eumeninae (“potter and mason wasps,” ~3,000 species worldwide) (22, 23) exhibit both solitary and presocial behavior (23); Stenogastrinae (“hover wasps,” ~50 species found in the Indo-Pacific tropics) (24) are facultatively eusocial (25); and Polistinae (“paper wasps,” ~800 species) (5) and Vespinae (“hornets and yellowjackets,” ~60 species) (26) are eusocial (27). However, relationships among the vespid subfamilies have been controversial (28–35), and the alternative proposed relationships have strikingly different implications for the evolution of social behavior.

The current leading phylogenetic hypothesis for Vespidae (Fig. 1A) (4, 5, 31) implies a single origin of social behavior in the common ancestor of a clade comprising Stenogastrinae + (Polistinae+Vespinae). This hypothesis was initially based on parsimony analysis of morphological and behavioral characters, but the three putative synapomorphies found to support the clade exhibit homoplasy (4) and may not reflect common ancestry. Additional behavioral characters proposed as shared, derived traits (synapomorphies) for the clade (5, 35) appear to be ambiguously characterized (19, 37). Indeed, earlier investigators of vespid social behavior argued that the often dramatic trait differences between stenogastrines and polistines+vespines in behavior (egg-laying, larval provisioning, nest architecture) and morphology (wing folding, pupal posture) reflect separate origins of sociality (19, 28, 29, 32–34). Moreover, a separate origin of stenogastrine sociality was inferred from a preliminary molecular analysis (Fig. 1B) of ~600 nucleotides from 16S and 28S rDNA (30), an analysis that remains controversial because of the absence of some ingroup subfamilies and the inappropriate selection of outgroups, which resulted in uncertainties in rooting the phylogeny (31).

Although interest in vespid relationships has expanded our understanding of these wasps, the proliferation of multiple phylogenies has placed a burden on students of social behavior to select the one that best reflects the true history. An incorrect phylogeny hampers a realistic understanding of hymenopteran social evolution, for example, by incorrectly inferring the number of social origins (6, 7, 38), misinterpreting homologies and transitional stages among social traits (6, 7, 38), and providing an inappropriate test of behavioral models of eusocial evolution (5, 14). Here we present a multigene phylogeny of Vespidae, including representatives of all currently recognized extant subfamilies. This phylogeny provides robust support for two independent origins of social behavior in these wasps.
resolved as a paraphyletic sister clade to Vespinae+Polistinae with maximum support (posterior probability = 1.0, ML and MP bootstrap values = 100). Eumeninae divided into two separate monophyletic clades: “Zethinae” (39) was sister to Polistinae+Vespinae, and the remaining “Eumeninae” were sister to Zethinae+ (Polistinae+Vespinae). “Eumeninae” divided into three principal groups: Pachodynerus-Stenodynerus, Symmorphus, and Alphamenes+Eumenes. Polistinae tribes (40) were recovered as monophyletic, with Ropalidiini (Ropalidia and Belonogaster) as sister to the remaining Polistinae, rather than the currently recognized Polistini (5).

**Discussion**

Our result that eusocial behavior evolved twice in two distantly related lineages is incongruent with the conventional single-origin hypothesis (4). Given our phylogeny, a dual-origin scenario is more parsimonious than a single origin of eusociality for Vespidae, which would require one gain of sociality along the branch leading to Vespinae, and three independent losses of sociality. Potential sources of error not otherwise addressed that could have misled these analyses include base composition bias and long-branch attraction. Base composition bias is an unlikely explanation for the grouping of Eumeninae with Vespinae+Polistinae, to the exclusion of Stenogastrinae, because eumenine and stenogastrine base composition is similar across genes. For instance, AT base composition of parsimony informative characters from the combined data set is 52.0% for Eumeninae and 52.1% for Stenogastrinae, and differs between the two subfamilies by only 1.5% for *abd-A*, the most informative gene for uniting Eumeninae with Polistinae+Vespinae. Long-branch attraction among clades is an unlikely source of error given that ingroup clades exhibit similar rates of substitutional change.

From this robust phylogenetic perspective, we can more confidently explore the hierarchical framework in which eusociality evolved independently from solitary ancestors and elucidate traits that may be driving vespid social evolution and their remarkable behavioral variation. A key group in which to search for early stages in the evolution of eusocial behavior is the solitary and presocial Eumeninae. The perspective that eumenines are the nearest relatives of eusocial Vespinae+Polistinae is not new. Earlier observers argued for this relationship on the basis of behavioral and morphological traits (41–45), such as the longitudinal folding of wings, a commonly used diagnostic feature of Vespidae that occurs only in eumenines, polistines, and vespines. Examining patterns of trait evolution within the morphologically and behaviorally diverse Eumeninae, which comprise the vast majority of vespid species, can shed light on whether and under what conditions a stepwise progression toward increasingly social traits occurred en route to the highly eusocial Polistinae+Vespinae. Our finding that Eumeninae is paraphyletic accords well with the distribution of their trait variation and supports an earlier taxonomic classification (39) of two subfamilies, Zethinae and Eumeninae. Eumenine relationships similar to our results were obtained in morphological analyses of Eumeninae by Vernier (46) and Carpenter (22), both of whom found a zethine taxon to be sister to the remaining Eumeninae.

“Zethinae,” the sister group to the eusocial taxa, exhibits traits that may be transitional between those of the ancestral eumenines and the eusocial Vespinae+Polistinae. For instance, rather than the typical eumenine nest construction with mud, the zethine genera *Zethus* and *Calligaster* are known to construct nests of plant material, a behavior that could precede the construction of nests from long-fiber wood pulp in the manner of Vespinae and Polistinae (41, 47, 48). Furthermore, *Zethus miniatus* will oviposit into an incompletely constructed nest cell (41), a behavior otherwise unique to polistines and vespines. *Z. miniatus* (41) progressively provisions larvae with intact prey, a behavior intermediate between the solitary vespid condition of mass provisioning with intact prey and the eusocial condition of

---

**Fig. 1.** Previous hypotheses of subfamily-level relationships of Vespidae. (A) The conventional hypothesis based on morphological and behavioral characters (4). (B) An alternative hypothesis based on limited molecular data (30), including the placement of Masarinae from a later study (36). Eusocial clades are indicated with gray shading.
progressive provisioning with macerated prey (17). *Z. miniatus* also exhibits communal behavior on occasions, with multiple adult females present on a multicelled aggregate nest. Females on these communal nests show a high degree of plasticity in whether they build and oviposit in a new cell, oviposit in a vacant cell, or usurp and oviposit in the cell of a less aggressive female; usurped females may depart and initiate a new nest (23). A similar reproductive flexibility characterizes the eusocial Polistinae. Indeed, *Z. miniatus* is believed to exemplify a social system from which polistine eusociality could have evolved (14, 49, 50). Cowan (ref. 23, p. 73) notes that *Zethus* and *Calligaster* “are regularly cited as exemplifying the critical evolutionary stages of subsocial and communal behavior that connect solitary and eusocial wasps,” a perspective that dates to de Saussure (51). The phylogeny presented here strongly supports this line of thinking and lends strong support to Cowan’s (ref. 23, p. 73) plea that “thorough reinvestigations of these insects, including careful attention to their life histories, nesting and mating behaviors, and population structures, are badly needed.” A more comprehensive eumenine phylogeny is also needed, both to investigate other potential instances of eumenine paraphyly and to identify basal character states that may pertain to the acquisition of social behavior in Zethinae and the higher grouping (Eumeninae/H11001 Polistinae/H11001 Vespinae)).

Our results highlight the distinct origin of social behavior in Stenogastrinae, a group that provides an independent source of information concerning traits that promote eusociality. Stenogastrine sociality is flexible and rudimentary, with nests that usually contain <10 adults and eusocial traits that are facultatively expressed (52, 53). Numerous authors have noted that the eusociality of Stenogastrinae differs in key aspects from that of Polistinae+Vespinae (24, 28, 29, 33, 34, 54). For example,
Materials and Methods

We selected a broad sample of 27 species of Vespidae representing all 6 currently recognized subfamilies (4), including taxa from the 4 known tribes of Polistinae (40), the historically recognized eumenine subfamilies Zethinae and Eumeninae (4), and 3 of the 7 described genera of Stenogastrinae (58) (SI Table 1). Outgroup taxa included exemplars from 3 other vespid families: Rhopalosomatidae, Scoleidae, and Tiphidae. Voucher specimens are retained at the Illinois Natural History Survey (Urbana, IL).

We generated DNA sequences from four gene fragments: the highly conserved nuclear 18S rDNA (variable regions V3–5 and related core elements), nuclear 28S rDNA (D2-D3 expansion regions and related core elements), and two intron-free protein-encoding genes, RNA pol II and abd-A. We extracted DNA from thoracic muscle or legs using the Dneasy Tissue Kit (Qiagen, Valencia, CA) and PCR-amplified each gene fragment with standard protocols by using Eppendorf HotMaster Taq and the following PCR conditions: initial denaturation at 94°C for 5 min; 35 repetitions of 94°C denaturation for 1 min, 48–58°C annealing for 1 min, and 72°C elongation for 1 min, 10 sec; and final elongation at 72°C for 5 min. Primers for each gene and their annealing temperatures are provided in SI Table 2. We purified the resulting products by using either the QIAquick PCR Purification Kit (Qiagen) or the QIAquick Gel Extraction Kit (Qiagen). We carried out sequencing reactions for both forward and reverse strands using BigDye v3.1 (ABI PRISM; Applied Biosystems, Foster City, CA). Automated sequencing was implemented at the W. C. Keck Center for Comparative and Functional Genomics (University of Illinois).

Alignment. We initially aligned sequences from the protein-encoding genes using the default parameters of Clustal W in Bioedit (59). We further refined the alignment of a few amino acid indels in abd-A (RNA pol II contained no indels) using protein translation. We aligned 28S and 18S rDNA sequences to secondary structure (60), following the methods of Gillespie (61), based on recent structural models for arthropod rRNA (62) (alignments are available in SI Materials and Methods). We excluded ~95 bp across regions of rDNA alignments where positional homology could not be established using structural criteria, including regions of alignment ambiguity, expansion and contraction, and slipped-strand compensation.

Phylogenetic Analyses. We estimated phylogenies for each gene region individually and for a combined-gene data set using both MP implemented in PAUP* v.4.0b10 (63) (heuristic search, 1,000 random additions, TBR branch swapping) and Bayesian inference implemented in MrBayes v3.1.2 (64) (4 independent runs at 3,000,000 generations, 4 chains, saving trees every 100 generations; consensus trees and posterior probability values obtained after removing trees from the first 300,000 generations, a point of convergence across all analyses). For the combined data set, we also performed an ML analysis in PAUP* [heuristic search, neighbor-joining starting tree, TBR branch swapping, SYM+1+G model and parameter values selected by using Modeltest v3.7 (65)].

For estimating clad support under parsimony, we implemented nonparametric bootstrap in PAUP* (400 replicates, 10 random additions per replicate, TBR branch swapping). For ML bootstrap estimates, we applied the same conditions used for obtaining the phylogeny, repeated for 100 replicates. We also calculated Bremer support values for the nodes in the combined-gene tree by using the strict consensus of the two most parsimonious trees as an input tree in the program TreeRot v2 (66). Bremer support values represent the number of extra steps required under parsimony if a given node is collapsed.

All Bayesian analyses were run with flat priors and different models for partitions, including variable rates and unlinking model parameters across partitions. For abd-A and RNA pol II, we combined codon positions 1 and 2 into a single partition and treated third positions separately. We partitioned 28S and 18S by stem and nonstem regions. Nucleotides that form pairs in 18S and 28S rRNA stems were treated as nonindependent units by using the doublet secondary structure model in MrBayes (64). Nucleotide substitution models were selected based on the Akaike information criterion in Modeltest and included GTR+I+G for 28S rRNA stem and nonstem regions, 18S rRNA stems, and abd-A position 3; GTR+I for 18S rRNA stems, RNA pol II positions 1 + 2, and abd-A positions 1 + 2; and HKY+I+G for RNA pol II position 3.

We especially thank R. Wallon, who assisted with the 28S rDNA sequencing, and C. K. Starr, J. Neff, and S. Turillazzi, who assisted with important specimen collections, along with A. R. Deans, E. J. Hernan-
dez, M. G. Keeping, J. P. Pitts, J. N. Seal, and A. Toth. We also thank M. J. Yoder for assistance with RNA-related Perl scripting and R. Gadagkar, A. Toth, J. Whitfield, and two anonymous reviewers for insightful comments on the manuscript. This research was supported by a U.S. Department of Agriculture (National Research Initiative CSREES 2004-35302-15077) grant (awarded to S.A.C.). J.H.H. was supported by a University of Missouri-St. Louis Research Leave Grant and the University of Illinois Sociogenomics Initiative (G. Robinson, Principal Investigator). J.J.G. was supported by National Institute of Allergy and Infectious Diseases Contract HHSN266200400035C (awarded to Bruno Sobral, Virginia Bioinformatics Institute at Virginia Tech, Blacksburg, VA).

51. de Saussure H (1875) *Smithson Misc Collect* 254.