A Phylogenetic Analysis of Soldier Evolution in the Aphid Family Hormaphididae
Author(s): David L. Stern
Source: *Proceedings: Biological Sciences*, Vol. 256, No. 1346 (May 23, 1994), pp. 203-209
Published by: The Royal Society
Stable URL: http://www.jstor.org/stable/50356

A phylogenetic analysis of soldier evolution in the aphid family Hormaphididae

DAVID L. STERN

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, U.S.A.

SUMMARY

Aphid soldiers, altruistic larvae that protect the colony from predators, are an example of highly social behaviour in an insect group with a natural history different from the eusocial Hymenoptera and Isoptera. Aphids therefore allow independent tests of theory developed to explain the evolution of eusociality. Although soldiers have been discovered in five tribes from two families, the number and pattern of origins and losses of soldiers is unknown due to a lack of phylogenetic data. Here I present a mtDNA based phylogeny for the Hormaphididae, and test the hypothesis that soldiers in the tribe Cerataphidini produced during two points in the life cycle represent independent origins. The results support this hypothesis. In addition, a minimum of five evolutionary events, either four origins and one loss or five origins, are required to explain the distribution of soldiers in the family. The positions of the origins and losses are well resolved, and this phylogeny provides an historical framework for studies on the causes of soldier aphid evolution.

1. INTRODUCTION

All known soldier aphids are members of two closely related families: the Hormaphididae and the Pemphigidae. The taxonomic distribution of soldier-producing species suggests that soldiers have been gained or lost at least six times; at least three evolutionary events are required for the Hormaphididae, and three for the Pemphigidae (Stern & Foster 1994). However, the current lack of phylogenetic estimates for these groups prevents resolution of the numbers and positions of origins and losses. Here I present a phylogeny for 14 species of the Hormaphididae, focusing on the Cerataphidini because most soldiers are found in this tribe, and include specimens from the Nipponaphidini and Hormaphidini to allow preliminary investigation of soldier evolution in these tribes.

The Cerataphidini are an interesting group for phylogenetic study for two reasons. First, many species have been examined for soldier production and these soldiers are the most aggressive and morphologically specialized known (Aoki 1987). Second, cerataphidine aphids produce soldiers at two points in the life cycle. All species produce soldiers in galls on the primary host, *Styrax*, and species of *Pseudoregma* and many *Ceratovacuna* also produce soldiers on the herbaceous secondary hosts (table 1). There are two hypotheses to explain this distribution of soldiers. First, primary- and secondary-host soldiers might represent a single evolutionary event with subsequent morphological and behavioural divergence. However, Aoki (1987) and Aoki & Kurosu (1989*a*) have argued that these two soldiers represent independent evolutionary events. First, they pointed out, primary-host soldiers are second-instar larvae, whereas secondary-host soldiers are first-instar larvae; secondly, primary- and secondary-host soldiers possess different weapons for attacking predators; and thirdly, they use different attacking behaviours. Primary-host soldiers use their stylets to pierce the cuticle of predators. Secondary-host soldiers possess enlarged forearms and horns which they jab into predators. Aoki & Kurosu (1989*a*) also noted the limited taxonomic distribution of secondary-host soldiers, and suggested that secondary-host soldiers therefore represent an independent and later evolutionary innovation.

A different type of secondary-host soldier has recently been reported from *Astegopteryx bambucifoliae* (Aoki & Kurosu, 1989*b*). These soldiers attack predators with their stylets, like primary-host soldiers. This is the only species of the genus *Astegopteryx* known to produce secondary-host soldiers, and its phylogenetic position is, therefore, of some interest.

2. MATERIALS AND METHODS

(*a*) *Data*

DNA spanning the mitochondrial Cytochrome Oxidase I (CO I) and II (CO II) genes from all species in table 1 was sequenced. Aphids collected in 95% ethanol were freeze-dried, and total DNA was extracted and purified by using standard procedures (Sambrook *et al.* 1989).

Polymerase chain reaction (PCR) amplification was done under standard conditions (Saiki *et al.* 1988) using mt2793+ (5′-ATACCTCGACGTTATTCAGA) and mt3660− (5′-CCACAAATTTCTGAACATTGACCA). These and the following primers were used for sequencing (redundant sites are indicated by a solidus and set off with hyphens): mt2853+ (5′-TGGATCAATAATTTC-T/A-ACATT), mt2993+ (5′-CATTCATATTCAGAATTACC), mt3175+ (5′-CATGA-C/T-CATACAATTTTTATTAT), mt3396+

Table 1. *Taxonomic position, current scientific names, collection locality, abbreviations used in this paper (table 2) and presence (P) or absence (A) of primary- and secondary-host soldiers for the species used in this study*

(Taxonomic positions are from Ghosh (1981). Data on soldier production are from Foster & Northcott (1994) and S. Aoki & U. Kurosu (personal communication). A question mark (?) indicates that the species is not known from this host plant.)

species	collection locality	abbreviation	primary-host soldiers	secondary-host soldiers
Family Hormaphididae				
Tribe Cerataphidini				
Aleurodaphis blumeae	Sayama, Chiba, Japan	AL	?	A
Astegopteryx bambucifoliae	Puli, Taiwan	AB	P	P
Astegopteryx unimaculata	Sun Moon Lake, Taiwan	AU	?	A
Cerataphis bambusifoliae	Kuangyinshan, Taiwan	CB	?	A
Ceratoglyphina bambusae	Sun Moon Lake, Taiwan	CG	P	A
Ceratovacuna japonica	Kumagaya, Japan	CJ	P	P
Ceratovacuna nekoashi	Niiza, Saitama, Japan	CN	P	A
Glyphinaphis bambusae	Gombak, Malaysia	GB	?	A
Pseudoregma alexanderi	Sun Moon Lake, Taiwan	PA	?	P
Pseudoregma koshunensis	Kuangyinshan, Taiwan	PK	P	P
Tuberaphis takenouchii	Sun Moon Lake, Taiwan	TT	P	?
Tribe Hormaphidini				
Hormaphis cornu	Princeton, New Jersey, U.S.A.	HC	A	A
Tribe Nipponaphidini				
Neothoracaphis yanonis	Kunitachi, Tokyo, Japan	NY	A	A
Nipponaphis distyliicola	Fuchu, Tokyo, Japan	ND	P	?
Family Pemphigidae				
Tribe Pemphigini				
Pemphigus microsetosus	Sapporo, Japan	PM	A	A

(5′-AATTATTGG-T/A-CATCAATGAT), mt3119− (5′-TTGTTCTATTAA-T/A-GG-T/A-GAAT), mt3339− (5′-GG-A/G-GATTTAATTTCATCTATT), mt3553- (5′-TT-GTTCATGAATGAAT-T/A-ACATC). Primer names indicate position relative to the *Drosophila yakuba* mitochondrial sequence (Clary & Wolstenholme 1985).

DNA was prepared for sequencing by using two methods. Either the initial fragment was reamplified by using one kinased primer and digested with lambda exonuclease (Higuchi & Ochman 1989), or the fragment was cloned by using the TA Cloning™ System (Invitrogen). Primers were removed from exonucleased reactions by using Kreitman & Landweber's (1989) protocol.

The prepared ssDNA or denatured supercoiled plasmids were sequenced by using the Sequenase kit (United States Biochemical). The reactions were electrophoresed on 6% polyacrylamide gels with a sodium acetate gradient (Sheen & Seed 1988). Gels were soaked in a 10% methanol/10% acetic acid bath, dried onto 3MM paper (Whatman's) and autoradiographed. Sequences were aligned manually by using MacClade version 3.0 (Maddison & Maddison 1992).

(*b*) *Phylogenetic analysis*

Data sets producing tree-length distributions that are strongly left skewed usually contain phylogenetic information, and the true tree tends to be close to the most parsimonious tree (Hills & Huelsenbeck 1992). Therefore, the skewness of the tree-length distribution was calculated for 10000 randomly generated trees by using PAUP version 3.0k (Swofford 1990).

Three methods were used to estimate the phylogeny. First, maximum parsimony (Felsenstein 1983) was used with the Heuristic algorithm of PAUP (tree bisection–reconnection branch swapping (TBR-swapping) random addition sequence with ten replicates). Secondly, neighbour-joining (Saitou & Nei 1987) was used, as implemented in PHYLIP, version 3.53c (Felsenstein 1993). Aligned sequences were converted into distance data by using the maximum-likelihood option of DNADIST with a transition/transversion ratio of 0.7191. The transition/transversion ratio was calculated over the most parsimonious tree by using the State Changes & Stasis option of MacClade. Thirdly, the maximum-likelihood tree was searched for by using the DNAML program of PHYLIP, using the same parameter values as for the neighbour-joining analysis.

Robustness of particular clades was examined by using bootstrapping (Felsenstein 1985). In addition, decay indices (Donoghue *et al.* 1992) were calculated for the branches of the most parsimonious tree.

(*c*) *Testing alternative hypotheses of soldier evolution*

Two major sources of error affect the reconstruction of trait evolution on a phylogeny, ignoring errors in data acquisition. The first involves errors in estimating the true phylogeny, and the methods used for estimating this error were described above. A second source involves errors in reconstructing the history of a trait on a given phylogeny. This reconstruction is typically done using parsimony (Maddison & Maddison 1992), although it is worth comparing reconstructions generated by alternative hypotheses of trait evolution to allow some basis for choice of one reconstruction over another.

A three-step procedure was used to examine two alternative hypotheses of soldier evolution (one origin against two origins of soldiers). First, the four most parsimonious topologies were found that required fewer total evolutionary events under the single origin hypothesis than the most parsimonious tree and the maximum-likelihood tree.

Secondly, the set of statistically equivalent phylogenies from these six that best explained the mtDNA data was circumscribed by using both a cladistic test (Templeton 1983) and a maximum-likelihood-based test (Kishino &

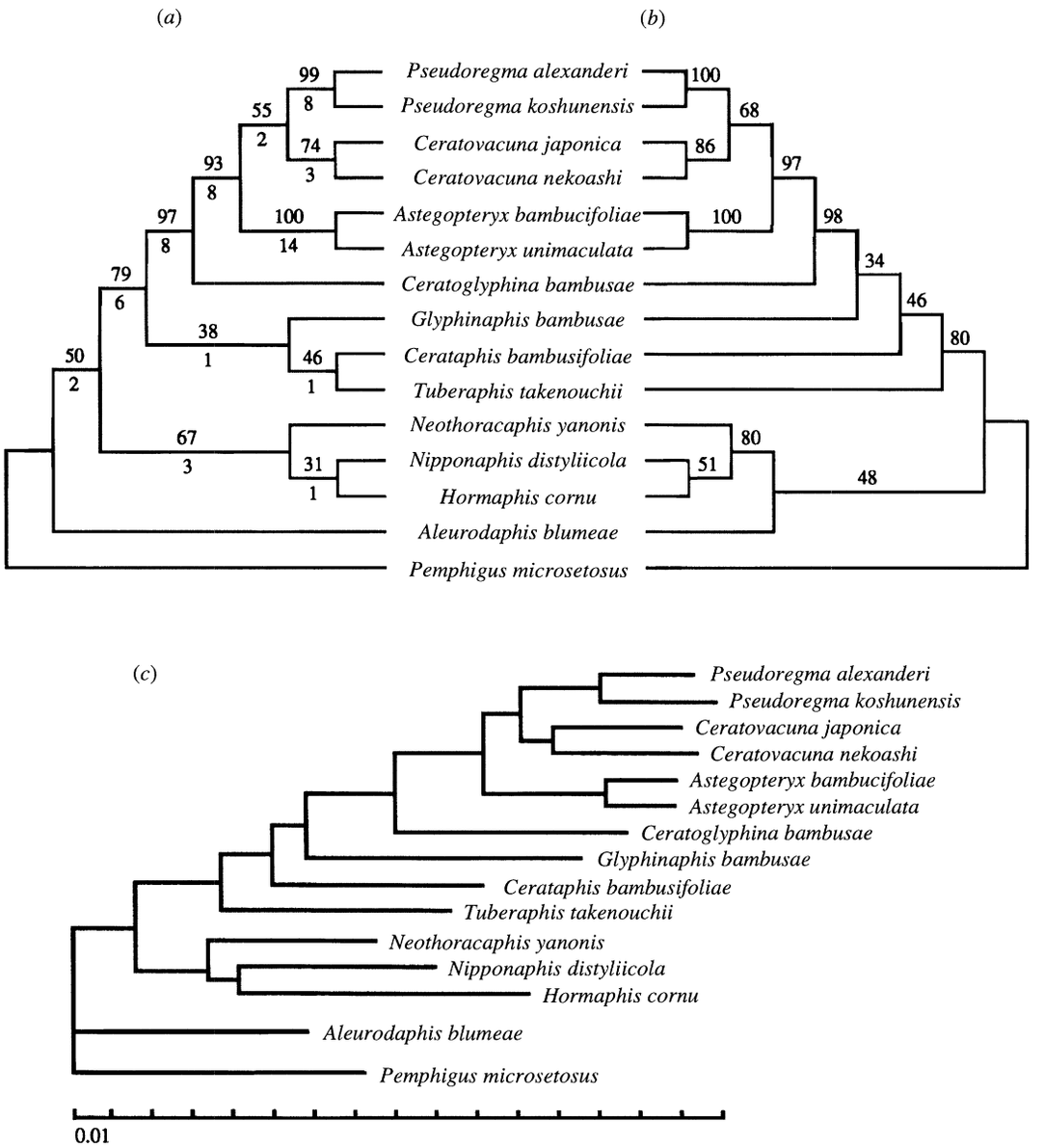


Figure 1. Phylogenetic reconstruction of 14 taxa of the Hormaphididae and one taxa from the Pemphigidae, using three different methods. All trees are rooted using the outgroup *Pemphigus microsetosus*. (*a*) The single most parsimonious tree (PAUP; Swofford 1990). Length for 209 informative sites = 752 steps; consistency index = 0.544; retention index = 0.458 (*b*). The neighbour-joining tree (PHYLIP; Felsenstein 1993). Numbers above branches for the parsimony and neighbour-joining trees are the percentage of bootstrap replicates (10000 for parsimony and 1000 for neighbour-joining) that recovered this branch. Numbers below branches on the parsimony tree are the decay indices (Donoghue *et al.* 1992). (*c*) The maximum-likelihood tree (PHYLIP; Felsenstein 1993). Horizontal branch lengths are proportional to the amount of expected change. The scale is marked in units of 0.01 expected substitutions per nucleotide position.

Hasegawa 1989) of alternative topologies. The differences between trees used in Templeton's test were calculated by using most parsimonious mappings of characters over alternative topologies provided by MacClade. Kishino & Hasegawa's test (1989), was performed by using PHYLIP.

Finally, the total number of evolutionary events required on the remaining statistically equivalent trees was compared and, by using a parsimony criterion, the model that required fewer total events on each tree was accepted.

3. RESULTS

(*a*) *Data*

DNA sequence data were obtained for 758–850 base pairs (b.p.) for each species. Aligned sequences can be obtained from GenBank (accession numbers L27324–L27338) or the author. Sequences were easily aligned by hand, except between the 3′ end of the CO I gene and the 5′ end of the leucine tRNA gene and in portions of the leucine tRNA. In addition, one species, *Ceratovacuna japonica*, possessed a 3 b.p. insertion in the CO II gene. This insertion was coded as a single codon

Table 2. *Wilcoxon signed ranks tests (Templeton 1983) and Kishino & Hasegawa (1989) paired sites tests of topologies providing differential support for the alternative hypotheses of a single origin and two origins of soldiers in the tribe Cerataphidini*

(Tree 1 is the maximum-parsimony tree and tree 2 is the maximum-likelihood tree. Trees 3–6 are the most parsimonious trees subject to the constraints imposed by the single-origin hypothesis (figure 2). For example, trees 3 and 4 are the most parsimonious trees that require one origin (O) and three losses (L) of soldiers, rather than the four losses required by the most parsimonious tree. The Wilcoxon test is performed for each topology relative to the most parsimonious topology, and the Kishino & Hasegawa test is performed for each topology relative to the maximum-likelihood topology.)

		number of events (single origin)[b]			number of events (two origins)[b,c]			Kishino & Hasegawa[d]				
tree[a]	tree length	O	L	total	O	L	total	Wilcoxon p	LL	ΔLL	s.d.	p
1	752	1	4	5	2	1	3	—	−4682.87278	−1.02681	10.4374	n.s.
2	753	1	6	7	2	1	3	n.s.	−4681.84597	—	—	—
3	760	1	3	4	2	1	3	< 0.05	−4706.30370	−24.45773	15.0500	n.s.
4	760	1	3	4	2	1	3	n.s.	−4709.98322	−28.13725	18.7773	n.s.
5	772	1	2	3	2	1	3	< 0.01	−4744.18970	−62.34373	22.4675	< 0.01
6	772	1	2	3	2	1	3	< 0.005	−4744.44053	−62.58456	20.1145	< 0.01

[a] Tree 1: ((AL,((((TT,CB),GB),(((AU,AB),((CJ,CN),(PA,PK))),CG)),((HC,ND),NY))),PM).
Tree 2: ((AL,((TT,(CB,(GB,(CG,((AU,AB),((CJ,CN),(PA,PK)))))))),((HC,ND),NY))),PM).
Tree 3: ((AL,(((((TT,CB),GB),CG),((AU,AB),((CJ,CN),(PA,PK)))),((HC,ND),NY))),PM).
Tree 4: ((AL,(((((((TT,CB),GB),CG),((AU,AB),((CJ,CN),(PA,PK)))),ND),NY),HC)),PM).
Tree 5: ((AL,(((((((TT,CB),GB),CG),(AU,AB)),((CJ,CN),(PA,PK))),ND),NY),HC)),PM).
Tree 6: ((AL,((((((TT,CB),GB),CG),(AU,AB)),((CJ,CN),(PA,PK))),((HC,ND),NY))),PM).
For abbreviations, see table 1.

[b] O = origins, L = losses, total = sum of origins and losses.

[c] All topologies also provide equal support for the equally parsimonious mapping of three origins of soldiers and zero losses. The alternative hypotheses of two origins and one loss against three origins cannot be distinguished with the current data.

[d] LL = log likelihood, ΔLL = difference in log likelihoods, s.d. = standard deviation of difference in log likelihoods, n.s. = not significant.

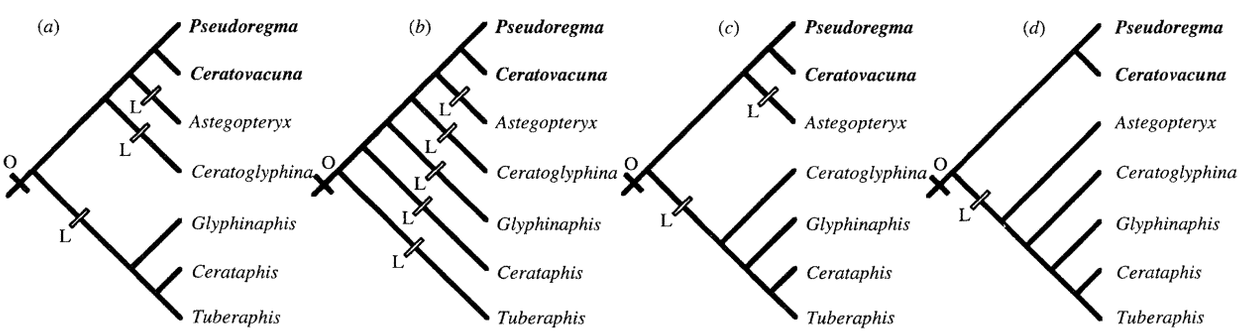


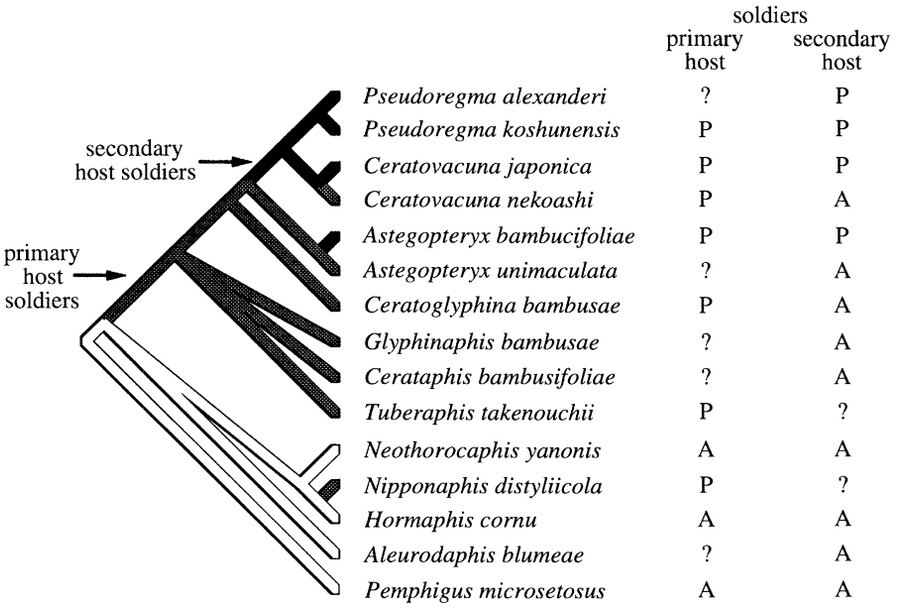
Figure 2. Alternative topologies of soldier-producing Cerataphidini compared in table 2 with the origins (O) and losses (L) of secondary-host soldiers under the single-origin hypothesis indicated. To simplify presentation, one species from each genus was selected and the relevant genera with secondary-host soldiers, *Pseudoregma* and *Ceratovacuna*, are in bold type. The single loss of soldiers in *Ceratovacuna* and single origin in *Astegopteryx* are not shown. (*a*) Tree 1; (*b*) tree 2; (*c*) trees 3 and 4; (*d*) trees 5 and 6.

insertion (positions 670–671 deleted), whereas length variation at the following 15 positions could not be interpreted unambiguously, and was deleted before analysis: 216–228, 246–247.

(*b*) *Phylogenetic analysis*

The skewness of the tree length distribution for 10000 randomly generated trees shows that the data contain significant phylogenetic structure ($g_1 = -0.5092$; $p < 0.01$; Hillis & Huelsenbeck 1992). The three methods of phylogeny estimation, maximum parsimony (figure 1*a*), neighbour-joining (figure 1*b*). and maximum likelihood (figure 1*c*), found similar topologies. In addition, bootstrapping tests using parsimony or neighbour-joining gave similar results. Branches with low bootstrap intervals and decay indices are the same branches for which the alternative phylogenetic methods disagree. In general, older splits are less well resolved, and their arrangement will require further data. However, many clades are well supported, which allows testing of hypotheses for soldier origins.

All of the soldier-producing Cerataphidini form a well-supported clade (figures 1 and 3) with a bootstrap interval of approximately 80%. Surprisingly, *Aleurodaphis blumeae*, which had been placed in the Cerataphidini based on morphological evidence (Ghosh 1985), appears distantly related to the other cerataphidines, and may represent a lineage basal to



	soldiers: primary host	soldiers: secondary host
Pseudoregma alexanderi	?	P
Pseudoregma koshunensis	P	P
Ceratovacuna japonica	P	P
Ceratovacuna nekoashi	P	A
Astegopteryx bambucifoliae	P	P
Astegopteryx unimaculata	?	A
Ceratoglyphina bambusae	P	A
Glyphinaphis bambusae	?	A
Cerataphis bambusifoliae	?	A
Tuberaphis takenouchii	P	?
Neothorocaphis yanonis	A	A
Nipponaphis distyliicola	P	?
Hormaphis cornu	A	A
Aleurodaphis blumeae	?	A
Pemphigus microsetosus	A	A

Figure 3. Parsimony mapping of soldier origins and losses on a strict consensus tree of maximum-parsimony, neighbour-joining, and maximum-likelihood trees. The presence (P) or absence (A) of primary- and secondary-host soldiers is indicated along the right. Question marks indicate that the species is unknown from this host, and it is unknown whether these species would produce soldiers on this host. To simplify the mapping, I have assumed that *Pseudoregma alexanderi*, *Astegopteryx unimaculata*, *Glyphinaphis bambusae* and *Cerataphis bambusifoliae* would produce soldiers on the primary host because there is no evidence that any lineage has lost primary-host soldiers. Stippled branches represent presence of primary-host soldiers. Black branches represent presence of both primary- and secondary-host soldiers.

all remaining Hormaphididae. The monophyly of *Astegopteryx* is well supported, with a bootstrap interval of 100%. Because this genus contains over 15 species, it is highly likely that soldiers of *A. bambucifoliae* represent a separate evolutionary event. This origin of soldiers is, therefore, not considered in the test below.

(*c*) *Testing alternative hypotheses of soldier evolution*

Table 2 presents results of comparisons of the maximum-parsimony and maximum-likelihood topologies with alternative topologies that more strongly support an hypothesis of one origin of soldiers in the Cerataphidini. The tests suggest that trees 1 and 2 are not statistically different, and that trees 3 and 4 are probably statistically equivalent to 1 and 2, although the cladistic test suggests that tree 3 is a significantly worse explanation of the data (table 2 and figure 2). Trees 1 to 4 all require more total evolutionary events under the single-origin hypothesis than under the two-origins hypothesis, supporting the two-origins hypothesis. Trees 5 and 6 require three events under either hypothesis, the same number required by all of the trees under the two-origins hypothesis. However, trees 5 and 6 can be rejected because they are significantly worse phylogenetic explanations of the mtDNA data than trees 1 and 2, lending further support to the two-origins hypothesis.

4. DISCUSSION

This is the first phylogenetic analysis of a group of soldier-producing aphids and illustrates the benefits of studying the evolution of social behaviour within an historical framework. Phylogenetic studies of vespid wasps (Carpenter 1989), halictid bees (Packer 1991), and apid bees (Cameron 1993) have similarly yielded insight into the frequency and mode of social evolution. These studies have illustrated how alternative hypotheses of social evolution can be tested within a phylogenetic framework, and have indicated the phylogenetic positions of independent origins of eusociality, information necessary for robust comparative tests of the causes of social evolution (Harvey & Pagel 1991). Evidence of independent origins is usually based on parsimony arguments and, because we currently lack a statistical approach to mapping traits, it is important to explore alternative methods for examining the support for particular mappings.

To discriminate between the hypotheses of one origin and two origins of soldiers in the Cerataphidini, I rejected the hypothesis requiring more total evolutionary events on the set of statistically equivalent trees. Origins and losses were weighted equally because no information was available to allow potentially more realistic weights. If origins and losses were weighted differently, then very different conclusions might be drawn. For example, if losses were assumed to be three times as likely as origins (so that origins received a weight of three and losses a weight of one), then we would be unable to discriminate between the hypotheses because trees 3 and 4 would favour one origin, tree 2 would favour two origins, and tree 1 would be equivocal.

Soldier evolution in the Hormaphididae is reconstructed on a consensus tree in figure 3. This reconstruction suggests that there were four origins of

soldiers in this family and one loss in the genus *Ceratovacuna*. An alternative equally parsimonious mapping (not shown) would require five origins, and no losses, of soldiers: the two primary-host origins and one secondary-host origin each in *Pseudoregma*, *Ceratovacuna* and *Astegopteryx*. This is probably an underestimate of the true number of events for the following reasons. First, I have sampled a relatively small number of species from the family, and many species have not yet been examined for soldier production. This is particularly true for the Nipponaphidini and Hormaphidini, which have been less well studied. Secondly, many species are still known from only one host plant, and data on soldier production from the alternate host may change the reconstruction illustrated in figure 3. Finally, the evolution of soldiers on the secondary host in *Pseudoregma* and *Ceratovacuna* is almost certainly more complicated than illustrated in figure 3. All species of *Pseudoregma* produce secondary-host soldiers, but only half of the approximately 15 species of *Ceratovacuna* do so. Further phylogenetic analysis of these two genera will almost certainly reveal interesting patterns of soldier evolution.

This analysis supports the hypothesis of Aoki (1987) and Aoki & Kurosu (1989*a*) that primary- and secondary-host soldiers of the Cerataphidini represent separate evolutionary events. However, do these origins represent independent evolutionary events?

It seems likely that the presence of primary-host soldiers influenced, and probably increased, the probability of secondary-host soldier evolution. This hypothesized non-independence between primary- and secondary-host origins could be examined by using a test developed by Maddison (1990). The currently small data set precludes the use of this test now, but such a test may be possible as more phylogenetic information is gathered for aphids. However, because secondary-host soldiers are only found in species with primary-host soldiers in species where both hosts are known (Foster & Northcott 1994), it is likely that the presence of soldiers in one part of the life cycle increased the probability of soldiers evolving in another part of the life cycle in aphids in general. In addition, the behaviour and natural history of soldiers supports this hypothesis. An important characteristic of all soldiers is that they attack rather than run away from predators. The morphological and behavioural details that have accompanied this switch and allowed specialization of soldiers to particular tasks are, in this context, peripheral (Stern & Foster 1994). It is possible that aphids on the secondary host co-opted the general set of behavioural traits for attacking already present in aphids on the primary host. The evolution of secondary-host soldiers in *Astegopteryx bambucifoliae* which behave like primary-host soldiers, using their stylets, rather than their horns, for attack, provides further support for this hypothesis.

Strictly, therefore, the origin of soldiers on the primary and secondary hosts probably cannot be considered statistically independent, although for many purposes it is probably reasonable to consider them uniquely informative about the causes of soldier evolution. For example, although the first origin might have increased the probability of the lineage gaining secondary-host soldiers, this is not a sufficient explanation for the origin of secondary-host soldiers. The presence of primary-host soldiers may have allowed the evolution of secondary-host soldiers when the appropriate ecological conditions arose. That is, secondary-host soldiers might not have evolved in this lineage if not for the presence of primary-host soldiers, but the origin and later loss of secondary-host soldiers was probably due to relevant changes in the selective régime, and future research will profitably be focused on elucidating these historical events.

I thank S. Aoki, J. Bonner, W. Foster, C. Godfray, C. Martinez del Rio, N. Moran, D. Stratton and an anonymous referee for comments on the manuscript. For invaluable assistance collecting and identifying aphids I thank S. Aoki, U. Kurosu and C. van Dohlen. I thank T. Gullison for use of his computer. For financial assistance and use of laboratory facilities, I thank M. Kreitman (NSF Grant #DMB-8857947), N. Pierce, C. Martinez del Rio (NSF Grant), D. Stratton and S. Hubbell. In addition, this research was supported by Sigma Xi, a NSF Graduate Fellowship, and a NSF Doctoral Dissertation Improvement Grant to D.L.S. and C. Martinez del Rio.

REFERENCES

Aoki, S. 1987 Evolution of sterile soldiers in aphids. In *Animal societies: theories and facts*, (ed. Y. Itô, J. L. Brown & J. Kikkawa), pp. 53–65. Tokyo: Japanese Scientific Society Press.

Aoki, S. & Kurosu, U. 1989*a* Two kinds of soldiers in the tribe Cerataphidini. *J. Aphidol.* **3**, 1–7.

Aoki, S. & Kurosu, U. 1989*b* A bamboo horned aphid attacking other insects with its stylets. *Jap. J. Entomol.* **57**, 663–665.

Cameron, S. A. 1993 Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8687–8691.

Carpenter, J. M. 1989 Testing scenarios: wasp social behavior. *Cladistics* **5**, 131–144.

Clary, D. O. & Wolstenholme, D. R. 1985 The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organization, and genetic code. *J. molec. Evol.* **22**, 252–271.

Donoghue, M. J., Olmstead, R. G., Smith, J. F. & Palmer, J. D. 1992 Phylogenetic relationships of dipsacales based on rbcL sequences. *Ann. Mo. bot. Gdn.* **79**, 333–345.

Felsenstein, J. 1983 Parsimony in systematics: Biological and statistical issues. *A. Rev. Ecol. Syst.* **14**, 313–333.

Felsenstein, J. 1985 Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. 1993 PHYLIP, Phylogenetic Inference Package, version 3.53c. University of Washington, Seattle.

Foster, W. A. & Northcott, P. A. 1994 Galls and the evolution of social behaviour in aphids. In *Plant galls: organisms, interactions, populations* (ed. M. A. J. Williams). Oxford University Press. (In the press.)

Ghosh, A. K. 1985 Hormaphidinae: distribution, phylogeny and systematics. In *Evolution and biosystematics of aphids. Proceedings of the International Aphidological Symposium at Jablonna, 1981*, pp. 303–336. Warsaw: Polska Academia Nauk, Ossolineum.

Harvey, P. H. & Pagel, M. D. 1991 *The comparative method in evolutionary biology*. Oxford University Press.

Higuchi, R. G. & Ochman, H. 1989 Production of single-

stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucl. Acids Res.* **17**, 5865.

Hillis, D. M. & Huelsenbeck, J. P. 1992 Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.* **83**, 189–195.

Kishino, H. & Hasegawa, M. 1989 Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J molec. Evol.* **29**, 170–179.

Kreitman, M. & Landweber, L. 1989 A strategy for producing single-stranded DNA in the polymerase chain reaction: A direct method for genomic sequencing. *Genet. analyt. Techn.* **6**, 84–88.

Maddison, W. P. 1990 A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* **44**, 539–557.

Maddison, W. P. & Maddison, D. R. 1992 *MacClade: Analysis of phylogeny and character evolution*, version 3.0. Sunderland, Massachusetts: Sinauer Associates.

Packer, L. 1991 The evolution of social behavior and nest architecture in sweat bees of the subgenus *Evylaeus* (Hymenoptera: Halictidae): a phylogenetic approach. *Behav. Ecol. Sociobiol.* **29**, 153–160.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R. J., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1988 Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science, Wash.* **239**, 487–491.

Saitou, N. & Nei, M. 1987 The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. New York: Cold Spring Harbor Laboratory Press.

Sheen, J. & Seed, B. 1988 Electrolyte gradient gels for DNA sequencing. *BioTechniques* **6**, 942–944.

Stern, D. L. & Foster, W. A. 1994 The evolution of sociality in aphids: a clone's-eye-view. In *Social competition and cooperation in insects and arachnids: vol. 2. Evolution of sociality* (ed. J. C. Choe & B. J. Crespi). Princeton University Press. (In the press.)

Swofford, D. L. 1990 PAUP, Phylogenetic Analysis Using Parsimony, version 3.0k. Champaign: Illinois Natural History Survey.

Templeton, A. R. 1983 Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* **37**, 221–244.

Received 14 January 1994; accepted 1 February 1994