5-2003

Phylogenetic Position of Pterocommatinae and Cavariella, and Implications for the Origins of Host Alternation in Aphids

Carol A. Rowe
Utah State University

PHYLOGENETIC POSITION OF PTEROCOMMATINAE AND *CAVARIELLA*, AND IMPLICATIONS FOR THE ORIGINS OF HOST ALTERNATION IN APHIDS

by

Carol A. Rowe

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Biology

Approved:

Dr. Carol D. von Dohlen Major Professor	Dr. Karen Mock Committee Member
Dr. James H. Cane Committee Member	Dr. Thomas L. Kent Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2003

ABSTRACT

Phylogenetic Position of Pterocommatinae and *Cavariella*, and Implications for the Origin of Host Alternation

by

Carol A. Rowe, Master of Science

Utah State University, 2003

Major Professor: Dr. Carol D. von Dohlen
Department: Biology

Aphids are morphologically simple. Their numerous hypothesized convergent reductions, such as reduced siphunculi length in association with ant attendance, have made it difficult to define morphological synapomorphies that are necessary for phylogenetic studies. Thus, I used molecular characters both to reexamine the phylogenetic relationships of *Cavariella* and Pterocommatinae within Aphididae, and to further map host associations and life cycles onto these phylogenies to better understand the evolutionary lability of host alternation within Aphididae. Independent and combined analyses were performed under unweighted parsimony and maximum likelihood criteria for sequences of mitochondrial cytochrome oxidase II plus tRNA-Leucine plus partial cytochrome oxidase I (COII + trnL), and nuclear elongation factor-1α (EF1α). Shimodaira-Hasegawa likelihood ratio tests were also employed to test for statistically significant differences between: (1) the tree topologies obtained from the

analyses in this study; and (2) topologies supporting the traditional phylogenetic hypotheses based upon morphological data. These analyses recovered various relationships contradicting the current morphology-based phylogeny: 1) a highly supported sister relationship of Pterocommatinae to *Cavariella*; 2) paraphyly of Myzinae and Anuraphidinae, as well as paraphyly in some genera within Dactynotinae; and 3) support for the sister relationship of Pterocommatinae/*Cavariella*/*Liosomaphis* to the remaining macrosiphines. There was also evidence within Aphididae for an evolutionary rapid radiation and multiple origins of host-alternation. These results imply the need for further molecular analyses in resolving relationships within Macrosiphini, and for defining the morphological attributes that characterize these relationships.

(47 pages)

ACKNOWLEDGMENTS

I would like to thank my committee, Drs. Carol von Dohlen, Karen Mock, and Jim Cane, for their support and assistance throughout this process. I would also like to thank Erik Maw and Bob Foottit for their invaluable work on the identification of aphid specimens, John Hanks for running my large data sets on the computer cluster, Colin Brammer for the use of a reliable (no crashes yet) Mac, and Erik Pilgrim and Carrie Drake for immoral support. My unlimited appreciation, debts, and sympathy go to Paul Wolf for putting up with me. This project was funded by the National Science Foundation (DEB-9807076 to Carol von Dohlen).

Carol A. Rowe

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The classification and systematics of aphids (family Aphididae *sensu* Remaudière and Remaudière 1997; suborder Sternorrhyncha; order Hemiptera) have been problematic issues dating back to the 18th century. The first aphid descriptions by Linnaeus included approximately 25 species, most of which were placed in the genus *Aphis* (Heie 1980). Currently, there are over 4700 described aphid species (Remaudière and Remaudière 1997) with almost as many proposed classifications as there have been practicing aphid taxonomists (Heie 1980). The lack of consensus, however, is mostly over the taxonomic level of different aphid groups rather than over the species that compose these groups. For example, Heie (1980) has placed all true aphids into the superfamily Aphidoidea and has further divided them into ten families. Blackman and Eastop (1994) lowered the taxonomic status such that all true aphids belong to the family Aphididae that is, in turn, divided into eleven subfamilies. The assignment of aphid species within these taxonomic groupings has, for the most part, remained the same.

Discrepancies within taxonomic groupings do exist, however, and are becoming more prevalent as the use of molecular characters has proliferated. The family Aphididae (sensu Heie 1980), for example, traditionally consists of the subfamilies Pterocommatinae and Aphidinae (Fig. 1). The subfamily Aphidinae is composed of the tribes Macrosiphini (which includes the genus *Cavariella*) and Aphidini. In a recent pilot study implementing the use of molecular characters (C. von Dohlen unpublished data), however, representatives of Pterocommatinae and *Cavariella* did not occupy these traditional positions, but rather formed a common lineage that was sister group to the rest

of Macrosiphini (Fig. 2). The high level of phylogenetic support for these sister relationships of Pterocommatinae to *Cavariella* and Pterocommatinae/*Cavariella* to Macrosiphini may have been an artifact of either small sample size (two *Pterocomma*, one *Cavariella*, various Macrosiphini, and few outgroup species) and/or problematic branch lengths in the tree. However, this molecular-based pilot study gains credibility in light of past taxonomic studies that were typically limited to morphological characters. Morphological synapomorphies in aphids are difficult to define due to the paucity of informative characters; in addition, those that exist have probably experienced convergent reductions. This problem is exemplified by Heie's (1994) re-examination and alteration of the Macrosiphini phylogeny. Heie (1994) postulated that previous estimates of Macrosiphini phylogeny were based upon convergent reductions in several morphological characters, such as siphunculi length, associated with ant attendance (myrmecophyly). Thus, I felt that the pilot study provided intriguing evidence warranting further investigation into the phylogenetic position of Pterocommatinae and *Cavariella*.

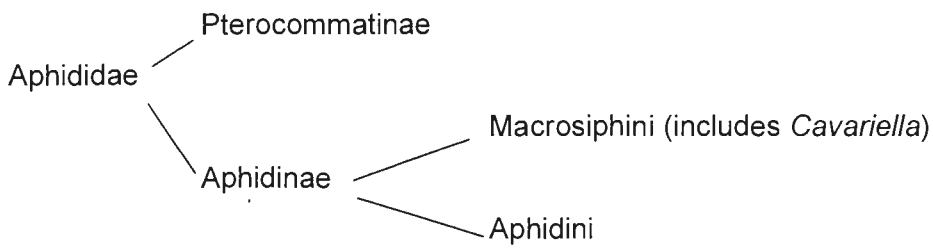


FIG. 1. Traditional hypothesis of Aphididae classification based on analyses of morphological characters by Heie (1980).

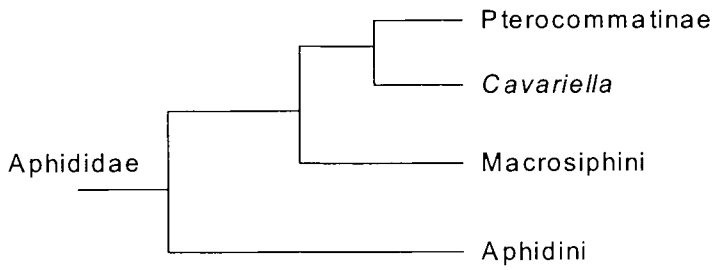


FIG. 2. Proposed hypothesis of Aphididae phylogeny based on preliminary analysis of molecular characters.

The objectives of this study were to test the two hypotheses that: 1) *Cavariella* is a sister group to Pterocommatinae and 2) Pterocommatinae plus *Cavariella* is a sister group to Macrosiphini. I tested these hypotheses by using molecular characters from both mitochondrial and nuclear genes to construct a phylogeny of Aphididae. In addition, the intriguing complexity of aphid life cycles was examined in the context of the resulting phylogeny. I mapped host associations and life cycles on the tree(s) to understand the evolutionary lability of host alternation, and the plasticity or rigidity in host plant associations. This, in turn, may also provide insight as to how host plant shifts are linked to aphid diversification.

BACKGROUND

The family Aphididae is the most successful of all extant aphid families, comprising over 60% of aphid species (Heie 1998). Prior to their Miocene radiation, the Aphididae accounted for approximately 4% of all Tertiary species (Heie 1994). Much of the success of Aphididae is, in part, due to their simplified morphology and their ability to exploit new ecological niches. In the following sections, I will briefly discuss aphid life cycles and host-plant associations, and how their evolutionary lability contributed to the predominance of the family Aphididae. The significance of this background information will be demonstrated when applied to groups within Aphididae (e.g., *Cavariella*, Pterocommatinae, Macrosiphini) and will be shown to be consistent with my hypotheses.

Aphid Life Cycles and Polyphenism

The intriguing complex life cycles of aphids are typically characterized by a wide range of polyphenisms (morphological differences between genetically identical individuals), several parthenogenetic generations, and a single sexual generation per year. The vast majority of aphid species are highly specific to only one or a few host plant species (Blackman and Eastop 1994). Aphid life cycles that are limited to these few, specific host plant species are referred to as non host-alternating or monoecious. However, ten percent of all aphid species have obligate, seasonal host shifts between two discrete groups of host plant taxa. These life cycles are referred to as host-alternating or dioecious.

A typical life cycle for a host-alternating species within Aphidinae can be divided into two phases with respect to the plant host. The woody primary host is the one or few closely related species of plants on which sexual reproduction takes place, eggs are laid, and the emerging females reproduce. The herbaceous secondary host plants, upon which only parthenogenetic reproduction takes place, are unrelated to the primary host and are typically more taxonomically diverse (Hille Ris Lambers 1966; Blackman and Eastop 1994). The single sexual generation is born in the fall (Heie 1998): winged males born on the secondary hosts fly to the primary host, where they mate with flightless, sexual females produced by winged, asexual females (gynoparae) that fly from the secondary to the primary host. Fertilized eggs are deposited on the primary host where they overwinter. The eggs hatch in the spring coinciding with the flush of their host's leaf buds, giving rise to highly fecund females called fundatrices. After a few or several parthenogenetic generations, winged females are produced that emigrate to the secondary hosts. This usually occurs during the summer, after the leaves of the primary host have matured and a dense colony of aphids has formed. These summer aphid generations may remain on a single secondary host, or winged females may be produced that can migrate to other secondary hosts. The fall migration back to the primary host coincides with seed set and leaf senescence. A monoecious aphid life cycle is essentially a simplified version of host-alternation in which both the sexual and parthenogenetic phases of the life cycle occur on one or a few closely related plant host species.

For most aphid lineages, polyphenism is most highly pronounced between the fundatrices and sexuals versus summer females (Moran 1992; Heie 1994). The degree of

polyphenism or specialization seems to be directly correlated with the evolutionary duration that the morph, especially the fundatrix, has been associated with its host (Moran 1988, 1992).

If too specialized, the fundatrix may prohibit the acquisition of more nutritive host plants as they become available over evolutionary time, potentially subjecting such aphid species to extinction (Hille Ris Lambers 1966; Moran 1988, 1992). Aphididae, however, tend to exhibit less extreme polyphenism than other aphid families. The lesser specialization or modification of the fundatrix by Aphididae has apparently allowed them to exploit a greater range of host plants, thus leading to the overwhelming success of Aphididae.

Host Specificity

Most aphid families are oligophagous, with aphid genera and species feeding on particular plant families and genera, respectively (Hille Ris Lambers 1966; Eastop 1973; Blackman and Eastop 1984). Unlike monoecious aphid species, host-alternating aphids have botanically distinct primary and secondary hosts, but are still specific in their host-plant relationships (Hille Ris Lambers 1966). Aphididae, however, are less narrowly restricted to their hosts, feeding on plants within a few different families. This host-plant lability is exaggerated in the many genera of Aphididae that no longer have their primary hosts, but instead are monoecious on their former secondary hosts, plus a range of reserve hosts (Hille Ris Lambers 1966; Moran 1992; Heie 1994).

Aphididae evolution is apparently associated with that of the Rosales (Hille Ris Lambers 1966; Blackman and Eastop 1984; Heie 1994, 1996). Hille Ris Lambers (1966)

proposed parallel evolution of the Rosales and Aphidinae: aphid taxa thought to be more primitive lived only on more primitive Rosales. The aphid taxa thought to be more derived lived on more advanced Rosales and have wider host ranges. Some genera of Aphididae, however, use host plants outside (and widely unrelated to) the Rosales. *Cavariella*, like the monoecious Pterocommatinae, use Salicaceae as their primary hosts. The recently hypothesized sister relationship of Pterocommatinae to *Cavariella* is consistent with the conservative evolution of host plant relationships. Other individual species such as *Aphis farinosa* and *Macrosiphum californicum* have likewise captured Salicaceae as their sole hosts, but are presumably unrelated to *Cavariella* and Pterocommatinae.

Evolutionary Lability of Host-Alternation

Host-plant plasticity and evolutionary lability in host-alternation have been key prerequisites for the diversification of aphid species. These gains, and probably multiple losses, of host-alternation and subsequent acquisition of new host plants seem to be responsible for Aphididae being the most species-rich of all aphid families. Most members of Aphidinae are monoecious on herbaceous plant families, with host-alternating species widely scattered throughout the entire family. It is hypothesized that the monoecious aphid species are derived from host-alternating ancestors that have subsequently lost their primary hosts and became monoecious on their former secondary hosts (Blackman and Eastop 1984; Moran 1988, 1992). Thus, the long-standing view is that there was a single origin and several losses of host-alternation within the subfamily Aphidinae. This hypothesis is supported by the common mechanism within all Aphidinae

of returning to the primary hosts via winged males and winged females that give rise to sexual females (versus migration back to the primary host by a non-sexual female that produces both sexual males and females on the primary host).

An alternative explanation for the distribution of host alternating life cycles in Aphidinae is that there were several origins, and fewer losses, of this life cycle. The phylogenetic position of *Cavariella* will have important implications for the evolutionary lability of host-alternation itself versus the plasticity or its lack in host plant associations. If *Cavariella* retains its current classical, taxonomic position within Macrosiphina (sensu Heie 1994), then there will be more support for lability in host plant associations, while the number of origins of host alternation will remain constant. In contrast, a sister group relationship between *Cavariella* and Pterocommatinae would be more likely to support the evolutionary lability of host-alternation with more conservation in host-plant associations. The total number of origins of host-alternation within Aphidinae will be better ascertained with improved phylogenetic understanding of this group.

MATERIALS AND METHODS

Selection of Species and Genes for Molecular Analysis

Table 1 lists the 56 aphid species included in this study. Among these species were world-wide representatives of the major clades of Aphidinae (as represented by Börner and Heinze [1957]), *Cavariella* (Remaudière and Remaudière 1997), Pterocommatinae (Remaudière and Remaudière 1997), and the outgroup clade of Pemphigini sensu Heie (1980). The aphid species selected for use in this study were initially based upon Heie's (1980) classification of the family Aphididae (Fig. 1). The classifications of Börner and Heinze (1957), as well as Remaudière and Remaudière (1997) were also consulted for their more inclusive representations of the subfamilies Aphidinae and Pterocommatinae, respectively (Table 2). All specimens were collected and stored in 80% ethanol (for identification and voucher) and 100% ethanol (for DNA extraction). Voucher specimens were deposited in the Canadian National Collection in Ottawa, Ontario, Canada, and the Utah State University Insect Collection in Logan, Utah, U.S.A.

I selected two gene regions for use in this study: 1) mitochondrial cytochrome oxidase II plus tRNA-Leucine plus partial cytochrome oxidase I (COII + trnL); and 2) nuclear elongation factor-1α (EF1α). These independent sources of data have been shown to effectively resolve branching at the subfamily and family levels, respectively, in aphids and other insects (Normark 1999; Rokas et al. 2002; von Dohlen et al. 2002).

TABLE 1. Aphid species included in this study.

Genus	species	Ref#	Host plant	Location	Collector	Collect. Date
Aphis	*farinosa* Gmelin	94-13	*Salix?*	USA: ?, Clearwater	CvD	29-Jun-90
Aphis	*helianthi* Monell	06.00.35	*Heracleum lanatum*	USA: UT, Logan Canyon	CR	18-Jun-96
Aphis	*nigratibialis* Robinson	06.00.36	*Cornus sericea*	USA: UT, Logan Canyon	CR	18-Jun-96
Aphis	*oenotherae* Oestland	06.00.42	*Ribes aureum*	USA: UT, Logan Canyon	CR	20-Jun-96
Aphis	*pomi* De Geer	09.00.44	*Cotoneaster*	USA: UT, USU	CR	26-Sep-96
Aphis	*spiraecola* Patch	01-36	*Malus* (Adirondack)	USA: IL, Lisle	CvD	3-Jun-97
Aphis		CR0141	Apiaceae	USA: NC, Clay CO	CR	29-May-98
Toxoptera	*citricida* (Kirkaldy)	E96-0644a	*Citrus x paradisi*	USA: FL, Coconut Grove		6-Mar-92
Hyalopterus	*pruni* (Geoffroy)	01-39	*Prunus sp.* (wild plum)	USA: UT, Logan	CvD	28-Jun-97
Rhopalosiphum	*maidis* (Fitch)			DNA gift of P. Baumann		
Rhopalosiphum	*padi* (L.)			DNA gift of P. Baumann		
Schizaphis	*graminum* (Rondani)			DNA gift of P. Baumann		
Brachycaudus	*cardui* (L.)	1500	*Carduus sp.*	USA: ID, Bingham Co.		28-Sep-89
Brachycaudus	*helichrysi* (Kaltenbach)	95-9	*Prunus (domestica?)*	USA: ID, Pocatello	CvD	1-Jun-91
Brachycaudus	*tragopogonis* (Kaltenbach)	94-08	*Tragopogon sp.*	USA: ID, Pocatello	CvD	28-Jun-90
Dysaphis	*plantaginea* (Passerini)	01-35	*Malus coronaria*	USA: IL, Lisle	CvD	3-Jun-97
Aphthargelia	*symphoricarpi* (Thomas)	94-12	*Symphoricarpos*	USA: ID, Pocatello	CvD	28-Jun-90
Brevicoryne	*brassicae* (L.)	94-48	*Brassica oleracea*	USA: ID, Pocatello	CvD	30-Jul-90
Cavariella	*aegopodii* (Scopoli)	04.00.03	*Salix sp.*	FRANCE: Grenoble	CR	26-Apr-96
Cavariella	*konoi* Takahashi	92EM-214	*Cicuta bulbifera*	CANADA: Ontario, Ottowa	EM	13-Sep-88
Cavariella	*pastinacae* (Linnaeus)	93EM-256	*Salix discolor*	CANADA: Ontario, Havelock	EM	3-Jul-89
Cavariella	*theobaldi* (Gillette and Bragg)	CR0124	Apiaceae	FRANCE: Rennes	CR	4-Sep-97
Cavariella	*archangelicae* (Scopoli)	CR0131	*Salix*	UK: Norwich (Univ. of E. Anglia)	CR	19-May-98
Diuraphis	*noxia* (Kurdjumov)			DNA gift of P. Baumann		
Hayhurstia	*atriplicis* (Linnaeus)	94-3	*Chenopodium*	USA: ID, Pocatello	CvD	22-Jun-90
Hyadaphis	*tataricae* (Aizen)	01-37	*Lonicera sp.*	USA: UT, Logan	CvD	17-Jun-97
Hyperomyzus	*lactucae* (Linnaeus)	94-86	*Ribes*	USA: ID, Pocatello	CvD	25-Sep-90
Liosomaphis	*berberis* (Kaltenbach)	CR0140	*Berberis sp.*	USA: NC, Franklin (Wendy's)	CR	29-May-98
Myzus	*varians* Davidson	02-36	*Clematis*	USA: NC, Barnard (Madison City)	CvD	1-Jun-98
Myzus	*persicae* (Sulzer)	94-90	*Brassica oleracea*	USA: ID, Aberdeen (from culture)	SH	10-Nov-90
Nasonovia	*ribisnigri* (Mosley)	02-52	*Lactuca*	NEW ZEALAND: Christchurch, Harewood	EM	29-Apr-98
Illionia	*linodendri* (Monell)	02-11	*Linodendron tulipifera ?*	USA: SC	CvD	28-May-98
Macrosiphoniella	*ludovicianae* (Oestlund)	06.00.40	*Artemisia ludoviciana*	USA: UT, Logan Canyon	CR	20-Jun-96
Macrosiphoniella	*millefolii* (de Geer)	CR0125	Apiaceae	FRANCE: Rennes	CR	4-Sep-97
Macrosiphoniella	*tenacetaria* (Kaltenbach)	94-30	*Tenacetum* sp.	USA: ID, Idaho Co., Lochsa R.	CvD	2-Jul-90
Macrosiphum	*aetheocornum* Smith & Knowlton	94-52	*Geranium*	USA: ID, Sulphur Canyon	CvD	4-Aug-90
Macrosiphum	*albifrons* Essig	94-4	*Lupinus*	USA: ID, Pocatello	CvD	22-Jun-90
Macrosiphum	*californicum* (Clarke)	CR0118	*Salix* sp.	USA: NY, Gilbertsville	CR	15-Jul-97
Macrosiphum	*rosae* (L.)	93-2	*Rosa* sp.	USA: ID, Pocatello	CvD	18-Oct-89
Metopeurum	*fuscoviride* Stroyan	02-89	*Tanacetum ?*	GERMANY	CvD	
Metopolophium	*dirhodum* (Walker)	94-05		USA: ID, Aberdeen (from culture)	SH	22-Jun-90
Sitobion	*avenae* (Fabricius)	94-91	*Triticum* sp.	USA: ID, Aberdeen (from culture)	SH	10-Nov-90
Uroleucon	*gigantiphagum* Moran	95-23	*Solidago*	USA: ID, Pocatello	CvD	10-Aug-91
Uroleucon	*russellae* Hille Ris Lambers	01-41	*Gnaphalium obtusifolium*	USA: NH, Pittsburg	CvD	28-Jul-97
Uroleucon	*sonchi* (L.)			DNA gift of P. Baumann		
Uroleucon	*tanaceti* (Linnaeus)	02-88	*Tanacetum ?*	GERMANY	CvD	
Plocamaphis	*flocculosa* (Weed)	CR0147	*Salix*	USA: UT, Logan Canyon (Temple Fork)	CR	12-Jul-98
Pterocomma	*beulahense* (Cockerell)	96-16	*Populus tremuloides*	USA: ID, Bannock Co. (outside Pocatello)	CvD	29-Jun-92
Pterocomma	*bicolor* (Oestlund)	CR0105	*Salix*	USA: UT, Logan	CR	22-Jun-97
Pterocomma	*populeum* (Kaltenbach)	CR0136	*Populus alba?*	UK: Norwich (city center-Cow Tower)	CR	20-May-98
Pterocomma	*salicis* (Linnaeus)	CR0130	*Salix*	UK: Norwich (Univ. of E. Anglia)	CR	19-May-98
Pterocomma	*sanguiceps* Richards	2000EM-0575	*Salix*	CANADA: B.C., Hagensborg	EM/RF	18-Jul-96
Pterocomma	*smithiae* (Monell)	CR0104	*Salix sp.*	USA: UT, Logan	CR	22-Jun-97
Pachypappa	*marsupialis* Koch	99-88	*Populus maximowiczii*	JAPAN: Moshiri, Hokkaido	SA	26-Jul-95
Prociphilus	*caryae* (Fitch)	99-70	*Amelanchier alnifoliae?*	USA: UT, Logan	CvD	10-Jul-95
Prociphilus	*fraxinifolii* (Riley)	99-23	*Pinus* roots	USA: NC, Elizabethtown	CvD	28-Apr-95

[a]CR = C. Rowe; CvD = C. von Dohlen; DH = D. Hales; EM = E. Maw; RF = R. Foottit; SA = S. Akimoto; SH = S. Halbert

TABLE 2. Taxonomic structure of Aphididae. Classification follows that of Börner and Heinze (1957) except that Pterocommatinae follows Remaudière and Remaudière (1997).

SUBFAMILY	TRIBE	SUBTRIBE	GENUS
Pterocommatinae			*Pterocomma*
			Plocamaphis
			Neopterocomma
			Fullawaya
			Paducia
Aphidinae[1]	Rhopalosiphonini		
	Aphidini		
Anuraphidinae[2]	Cryptosiphonini		
	Acaudinini		
	Anuraphidini	Anuraphidina	
		Brachycaudina	
Myzinae[2]	Brachycolini	Brachycolina	
		Coloradoina	
	Myzaphidini		
	Liosomaphidini		
	Phorondontini		
	Myzini	Pentalonina	
		Myzina	
	Cryptomyzini		
	Nasonoviini		
Dactynotinae[2]	Aulacorthini	Microlophiina	
		Hottesina	
		Aulacorthina	
	Macrosiphonini	Macrosiphonina	
		Sitobiina	
	Dactynotini	Metopeurina	
		Dactynotina	
	Megourini	Megourina	
		Wahlgreniellina	

[1] Equivalent to Aphidini of Heie (1980)
[2] Composes Macrosiphini of Heie (1980)

DNA Extraction, PCR, Cloning, and Sequencing

DNA was extracted following the 'protein salting-out' protocol of Sunnucks and Hales (1996) with the two following modifications: numerous aphids were extracted in a single tube and a portion of the TNES buffer was added in addition to the proteinase K when crushing the aphids. Primers used for polymerase chain reactions (PCR) are listed in Table 3 and mapped in Fig. 3 and 4. PCRs were carried out in 25 µl reaction volumes consisting of 0.2 mM each dNTP, 1X PCR buffer with 1.5 mM $MgCl_2$ (Roche Molecular Biochemicals, Indinanapolis, IN), 1.25 units Taq polymerase (Roche Molecular Biochemicals), 5 pmols of each primer, 25 ng genomic DNA, and (in the amplification of the EF1α gene) 0.5 µg T4 gene 32 protein. A typical temperature profile for EF1α consisted of 1 cycle of 95°C for 3 min., 35 cycles of 94°C for 1 min., 55°C for 1 min., and 72°C for 1 min., and a final extension of 72°C for 10 min. in either a GeneAmp 2400 or 9600 thermalcycler. The annealing temperature for the COII + trnL reactions was 48°C. Finished reactions were run on a 0.7% agarose gel and then stained with ethidium bromide. PCR products were purified with ammonium acetate and isopropanol precipitation. Problematic PCR products were cloned into plasmid vectors using the TA Coning Kit (Invitrogen, San Diego, CA) and purified using the Quantum Prep plasmid miniprep kit (BIO-RAD, Hercules, CA). Direct sequencing of the PCR products and plasmid inserts was carried out with Perkin-Elmer BigDye Terminator chemistry (PE Biosystems, Foster City, CA) and visualized on an ABI 377 sequencer.

TABLE 3. Primers used for PCR amplification and sequencing of mitochondrial cytochrome oxidase II, tRNA-Leucine, plus partial cytochrome oxidase I (COII + trnL), and nuclear elongation factor-1α (EF1α) genes.

Locus	Name (alias)	Reference	Direction[a]	Use[b]	Sequence: 5' to 3'
EF1α	EF3	von Dohlen et al. 2002	s	P,S	GAACGTGAACGTGGTATCAC
EF1α	EF4	von Dohlen et al. 2002	s	S	GAACCACCATACAGCGAA
EF1α	EF4a	von Dohlen et al. 2002	s	S	GAACCACCGTACAGTGAAG
EF1α	EF5-rc	this study	s	S	GAAGTCAGCAGTTACATCAA
EF1α	EF7	von Dohlen et al. 2002	s	S	ATTGGAGGTATTGGAACAGT
EF1α	EF5	von Dohlen et al. 2002	a	S	TTGATGTAACTGCTGACTTC
EF1α	EF8	this study	a	S	GGGACTGTTCCAATACCTCC
EF1α	EF6	von Dohlen et al. 2002	a	P,S	TGACCAGGGTGGTTCAATAC
EF1α	EF2	Palumbi 1996**	a	P,S	ATGTGAGCAGTGTGGCAATCCAA
COII + trnL	C1-J-1859 (RonII)	Simon et al. 1994	s	P	GGAACIGGATGAACWGTTTAYCCICC*
COII + trnL	2951+	this study	s	P,S	TAATTCAATTGAATGAAT
COII + trnL	2993+	Stern 1994	s	P,S	CATTCATATTCAGAATTACC
COII + trnL	S2792	Normark 1996	s	S	ATACCTCGACGTTATTCAGA
COII + trnL	C2-N-3661 (Barbara)	Simon et al. 1994	a	S	CCACAAATTTCTGAACATTGACCA
COII + trnL	A3772 (Eva)	Normark 1996	a	P,S	GAGACCATTACTTGCTTTCAGTCATCT

[a]a = antisense, s = sense

[b]Primer used for PCR amplification (P), sequencing (S), or both

* I = inosine

** primer designed by G. Roderick

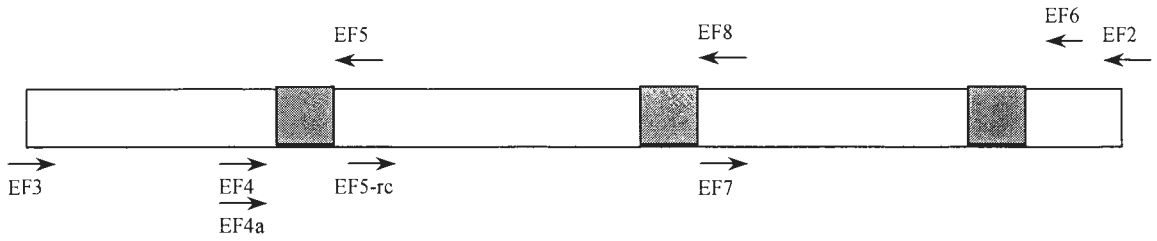


FIG. 3. Primer map for the nuclear elongation factor-1α (EF1α) gene. The shaded areas represent the relative position of introns.

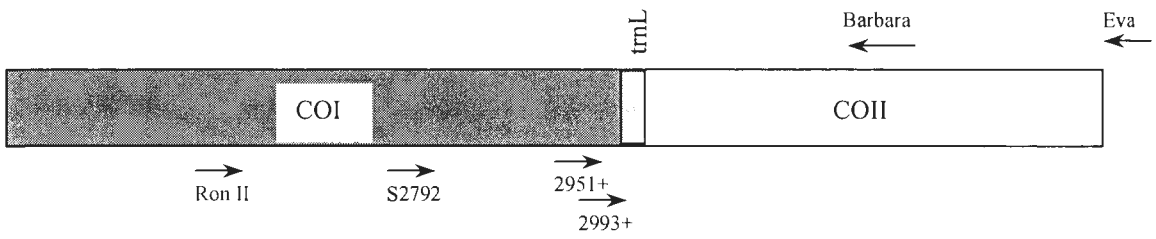


FIG. 4. Primer map of mitochondrial cytochrome oxidase II, tRNA-Leucine, and partial cytochrome oxidase I genes (COII, trnL, and COI). Primers 2951+ and 2993+ are located approximately 60 and 25 nucleotides, respectively, from trnL. RonII is located approximately 1859 nucleotides from the start of COI. The 3' end of Eva is 6 nucleotides outside of COII.

Sequence Compilation, Alignment, and Phylogenetic Analyses

Sequences were compiled and aligned using Sequencher 1.1. Other alignment programs were not necessary due to the lack of ambiguities in alignment; COII + trnL sequences had no insertions or deletions, and the variable-length introns contained in EF1α were removed before analysis. All sequences will be submitted to GenBank upon submission of this work to a peer-reviewed journal.

Phylogenetic analyses were conducted using PAUP* 4.0b8a (Swofford 1998). Maximum parsimony (MP) analyses were run for each gene separately and as a

combined data set under the heuristic search strategy with all sites equally weighted, and 100 random-addition replicates with tree-bisection-reconnection (TBR) branch swapping. The relative robustness of individual clades was assessed by nonparametric bootstrapping (Felsenstein 1985) on parsimony-informative characters only, under heuristic search strategies with 10 random-sequence addition replicates for each of 1000 bootstrap replicates.

The best-fit model of nucleotide evolution for maximum likelihood (ML) analyses was identified with Modeltest 3.06 (Posada and Crandall 1998). Typically, a single best-fit model of evolution is obtained using the two evaluation criteria (log-likelihood ratio tests and Akaike Information) in Modeltest. If this was not the case, ML analyses were performed using the simpler of the two models generated. Models were selected and ML analyses were performed for each locus and for the combined data set in PAUP* under heuristic search strategies with 10 random-sequence additions and TBR branch swapping. Bootstrap analyses were performed under the same evolution models under heuristic search strategies with one random-addition sequence for each of 500 pseudoreplicates.

Tests for mutational saturation within each locus were conducted using the method proposed by Philippe et al. (1994). Under this method, the uncorrected distances between phylogenetically independent pairs of species (N/2 – 1 pairs, where N = the number of taxa) were plotted against corrected ML-estimated distances from the same species pairs. The tree used to pick taxon pairs was arbitrarily chosen from one of the shortest trees from the unweighted MP analysis. ML distances were calculated according to the best-fit model of nucleotide evolution from Modeltest.

I used the Shimodaira-Hasegawa (SH) likelihood ratio tests (Shimodaira and Hasegawa 1999) under the best-fit evolutionary model to test for statistically significant differences between each of the ML trees obtained from EF1α and combined EF1α + COII + trnL data analyses to all corresponding MP trees, and to a set of trees constructed to test hypotheses concerning the placement of specific clades. This set included: (1) a tree consistent with the taxonomic structure of Börner and Heinze (1957); (2) a tree constraining *Cavariella* and *Liosomaphis* to be members of Macrosiphini; (3) a tree constraining the three macrosiphine subfamilies of Börner and Heinze (1957), with the exception of *Cavariella* and *Liosomaphis*; and (4) a tree in which only the Pterocommatinae clade was placed in the traditionally basal position (Heie 1980) of the otherwise completely resolved ML tree (Fig. 5D). Test trees were obtained by implementing ML searches under the listed constraint trees to find the optimal ML tree consistent with the given hypothesis, except in the latter tree where the ML value for the tree was directly obtained. For tree (1), an ML tree was obtained from implementing a Börner and Heinze-based (1957) constraint tree in which all branches were left resolved except for the basal placement of Pterocommatinae and the reduced polytomy of the macrosphines into Anuraphidinae, Myzinae, and Dactynotinae (Fig. 5A). For tree (2), an ML tree was obtained from implementing a constraint tree in which all branches were left resolved except for the placement of the *Cavariella* clade and *Liosomaphis* into a completely polytomous macrosiphine clade (Fig. 5B). For tree (3), an ML tree obtained from implementing a macrosiphine constraint tree in which all branches were left resolved except for the reduced polytomy of the macrosiphines into Anuraphidinae,

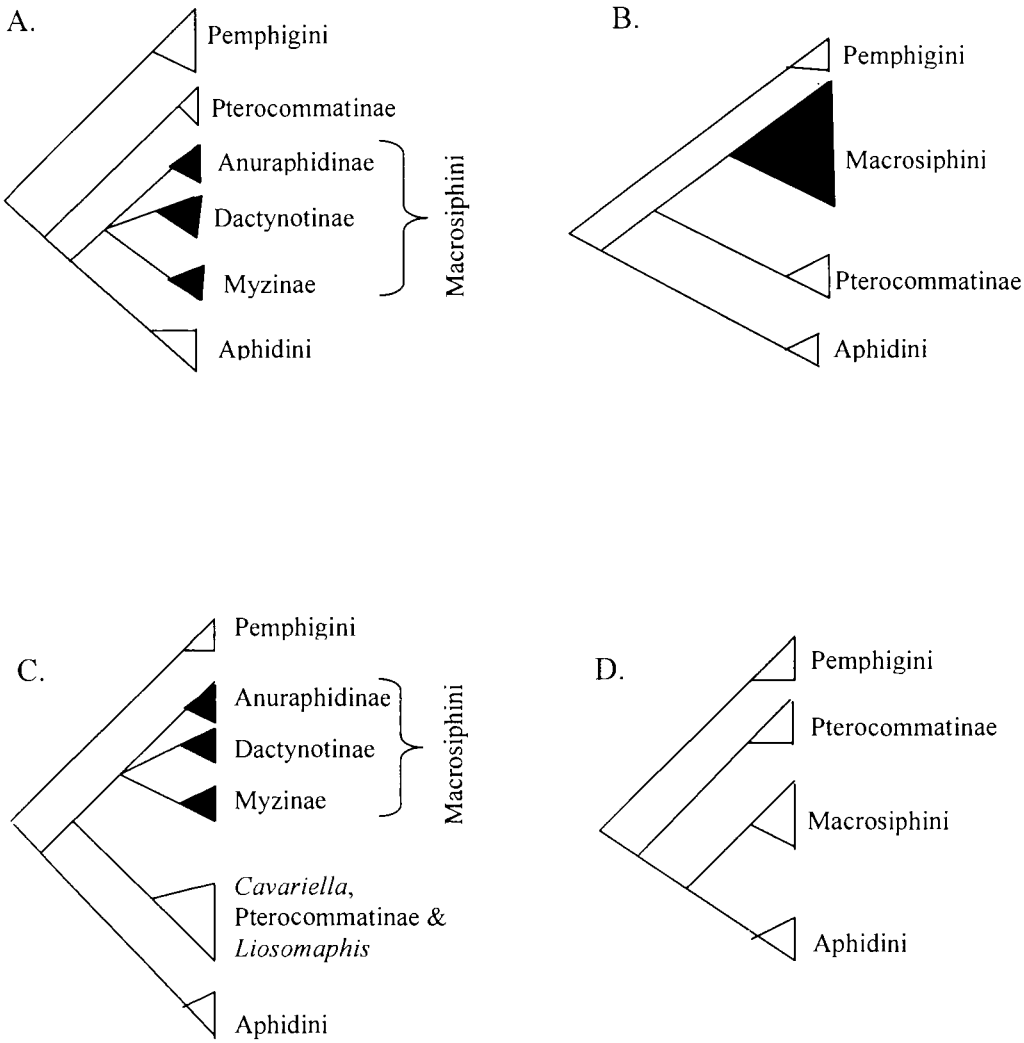


FIG. 5. Four test trees (A-D) used to test hypotheses of Aphididae phylogeny. Three constraint trees (A-C) were used for ML analyses of EF1α, and the combined data set of EF1α +COII + trnL under identical models and search strategies as the unconstrained searches in this study. One test tree (D) was used directly to obtain ML values for the EF1α, and the combined data sets. Shaded triangles represent unresolved clades. Unshaded clades were left as resolved in the unconstrained ML trees.

Myzinae, and Dactynotinae (i.e., the Pterocommatinae/*Cavariella* clade remained as-is in the original ML analyses) (Fig. 5C). The four constraint trees obtained from the EF1α data set, in addition to the ML tree and the MP trees, were used as source trees from which new likelihoods were determined under the EF1α + COII + trnL data set. This was also done in reverse with the trees from the combined data set as the source trees from which new likelihoods were determined under the EF1α data set.

The SH test is appropriate for comparisons of an *a posteriori* tree to other *a priori* or *a posteriori* trees (i.e., when multiple topologies generated from the same data are being tested) and requires all feasible alternative hypotheses such that what might be the true topology is always available for comparison against the ML topology (Shimodaira and Hasegawa 1999; Goldman et al. 2000; Buckley et al. 2001). The expected differences in log likelihoods are adjusted to the expectation of the null hypothesis that the topologies are not significantly worse (versus the expectation of the difference in log likelihoods equals 0 in the Kishino-Hasegawa [1989] test). This is because the difference in log likelihoods will always be > 0 since the ML tree will always have the highest likelihood score (Shimodaira and Hasegawa 1999; Goldman et al. 2000; Buckley et al. 2001). The SH test, therefore, is a one-tailed test.

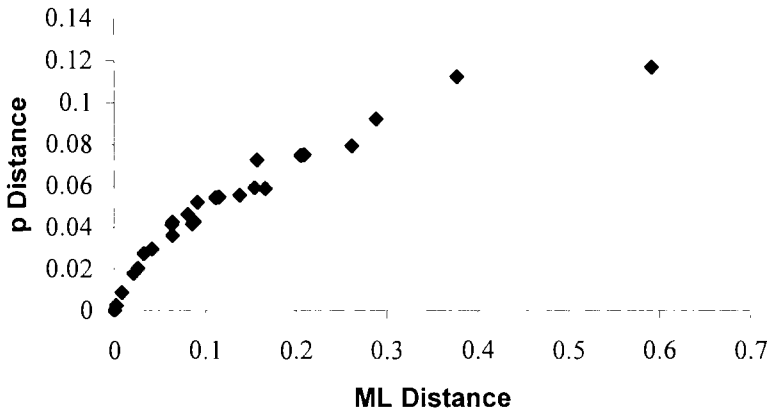
RESULTS

Mitochondrial (COII + trnL) Data

The aligned data set contained 778 characters (excluding primers), of which 217 were parsimony-informative. These sequences were biased toward A (40.1%) and T (39.6%) nucleotides, as observed for insects in previous studies (Simon et al. 1994; von Dohlen et al. 2002). The plots of uncorrected p-distances against ML distances showed saturation beginning at ~10% to 18% ML distances, where the slope of the comparison approaches an asymptote (Fig. 6A). The initial saturation corresponds to independent pairwise comparisons at or above the genus level. The point with the largest xy value in this graph corresponds to the independent pairwise comparison of *Aphis spiraecola* to *Pachypappa marsupialis* (a member of the outgroup).

The unweighted MP analysis yielded 98 most-parsimonious trees of length = 1289, consistency index (CI) = 0.320, retention index (RI) = 0.514, rescaled consistency index (RC) = 0.164, and homoplasy index (HI) = 0.680. The majority of the trees (Fig. 7) exhibited the monophyly of *Cavariella*, Aphidina, and Rhopalosiphina. Three sister relationships were also found: (1) *Cavariella* to Pterocommatinae; (2) *Cavariella*/ Pterocommatinae/*Liosomaphis* to Macrosiphini; and (3) Aphidina to Rhopalosiphina. Due to a general lack of resolution at the deeper nodes (above species level), bootstrap support (BS) for the above relationships was limited. The relationships with >50% of bootstrap pseudoreplicates included the monophyly of Rhopalosiphina (BS = 65%), and *Cavariella* (BS = 78%). Also supported was the questionable relationship of *Pterocomma*

A. COII + trnL



B. EF1α

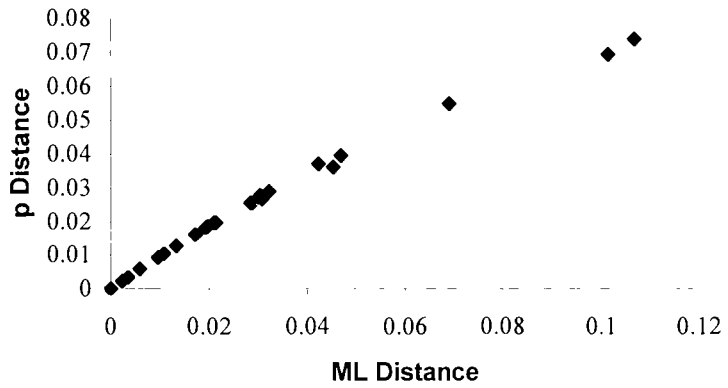


FIG. 6. Uncorrected (p) pairwise distances plotted against ML-corrected pairwise distances for data sets (A) COII + trnL, and (B) EF1α.

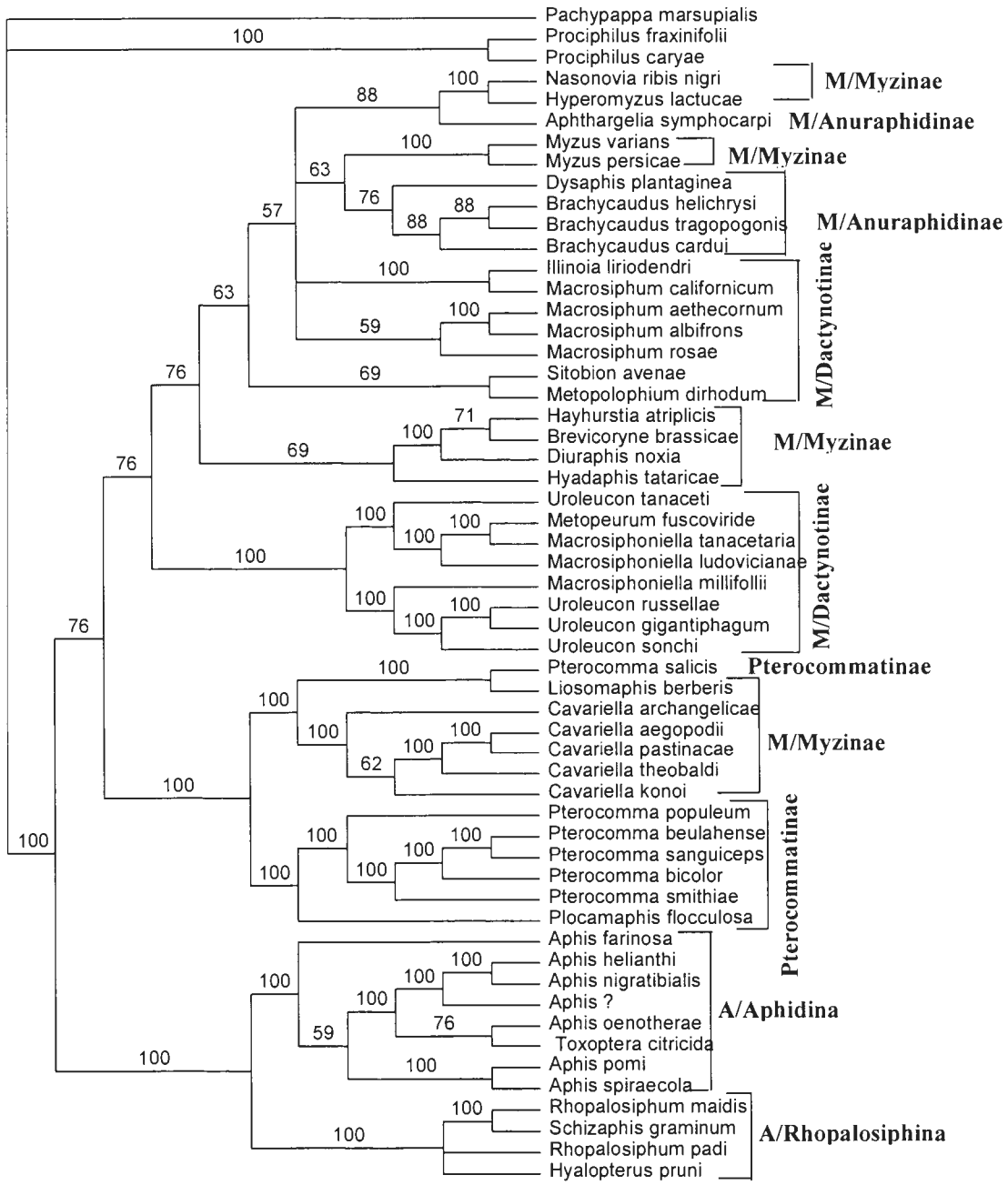


FIG. 7. The 50% majority rule tree from the consensus of the 98 most parsimonious trees obtained from the unweighted maximum parsimony analysis (heuristic search with 10 random-additions and TBR branch swapping) of the COII + trnL data set. A = Aphidinae. M = Macrosiphini sensu Heie 1980.

salicis with *Liosomaphis berberis* to form a clade sister to the *Cavariella* plus remaining pterocommatines clade (BS = 66%).

Maximum likelihood analysis of the COII + trnL data set under the best-fit model (GTR+I+G, Lanave et al. 1984; Swofford et al. 1996) yielded one tree which showed the same monophyletic groups and relationships as in the unweighted MP analysis, except for the sister relationship of *Cavariella*/Pterocommatinae/*Liosomaphis* to Macrosiphini. The *Cavariella*/Pterocommatinae/*Liosomaphis* clade was sister to Aphidini, and their common lineage was positioned within the macrosiphines. Boostrapping under ML was not performed given the lack of resolution at deeper nodes in this data set.

Nuclear (EF1α) Data

The aligned EF1α data set consisted of 865 characters, of which 172 characters were parsimony-informative. This set excludes primers and the three variably sized introns of approximately 60-70 nucleotides each. The base composition was more uniform than that of COII + trnL: A = 28.7%, T = 26.3%, G = 23.6%, and C = 21.4%. The plots of uncorrected p-distances against ML distances showed only very slight evidence of saturation at the greatest distances (Fig. 6B).

Unweighted MP analysis of the EF1α data set yielded six most-parsimonious trees of length = 714, CI = 0.422, RI = 0.749, RC = 0.316, and HI = 0.578. Boostrapping supported many clades. All trees included a monophyletic *Cavariella*, Pterocommatinae (BS = 97%), Aphidina (BS = 96%), Rhopalosiphina (BS = 64%), and Dactynotinae (BS = 64%). Also supported were the sister relationships of *Cavariella* to Pterocommatinae (BS = 100%), *Liosomaphis* to *Cavariella*/Pterocommatinae, Aphidina to Rhopalosiphina

(BS = 91%), *Cavariella*/*Liosomaphis*/Pterocommatinae to all other macrosiphines, and *Liosomaphis* to Pterocommatinae/*Cavariella*.

Maximum likelihood analysis of the EF1α data set under the best-fit model yielded one tree (Fig. 8), which included nearly all of the relationships in the unweighted parsimony analysis with the main exception of the placement of *Liosomaphis* with respect to *Cavariella* and Pterocommatinae. Many of the same relationships were supported by the bootstrap. *Liosomaphis* was sister to Pterocommatinae (BS = 58%), and *Liosomaphis*/Pterocommatinae was sister to *Cavariella* (BS = 100%) (Fig. 8).

Combined Mitochondrial and Nuclear Data

An unweighted MP analysis yielded five most-parsimonious trees of length = 2036, CI = 0.350, RI = 0.616, RC = 0.216, and HI = 0.650. All parsimony trees were similar to the single tree obtained from the ML analysis under the best-fit model (Fig. 9). Bootstrap analyses under both MP and ML methods supported many of the previous monophyletic groups and sister relationships that were observed and well-supported within the individual data sets (Fig. 9). The two major differences between trees obtained from the individual versus the combined data analyses were that the COII + trnL data set had slightly higher resolution at the species level, while the EF1α displayed higher resolution between genera. The complementary areas of resolution observed in the independently analyzed genes resulted in a more highly resolved tree with better-supported relationships from the combined genes analysis (Fig. 9).

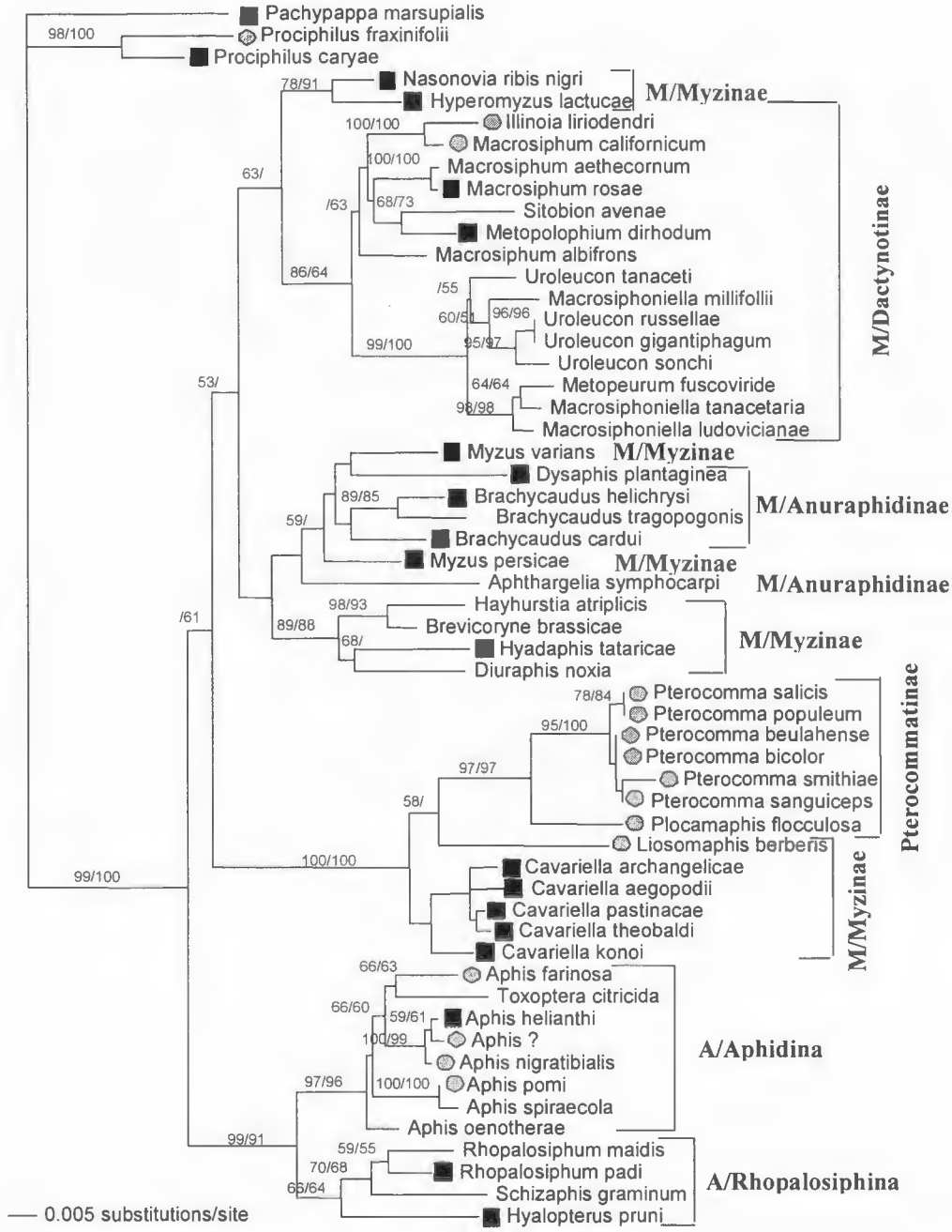


FIG. 8. Maximum likelihood (-ln likelihood = 5041.26) phylogeny of EF1α sequences (heuristic search with 10 random-additions and TBR branch swapping) under the best-fit evolutionary model (TrN+I+G, Tamura and Nei 1993). At each node, nonparametric bootstrap values (>50%) for maximum likelihood/maximum parsimony are given (500 and 1000 replicates, respectively). A = Aphidinae. M = Macrosiphini sensu Heie 1980. ■ = host-alternating. ◎ = monoecious on a woody host.

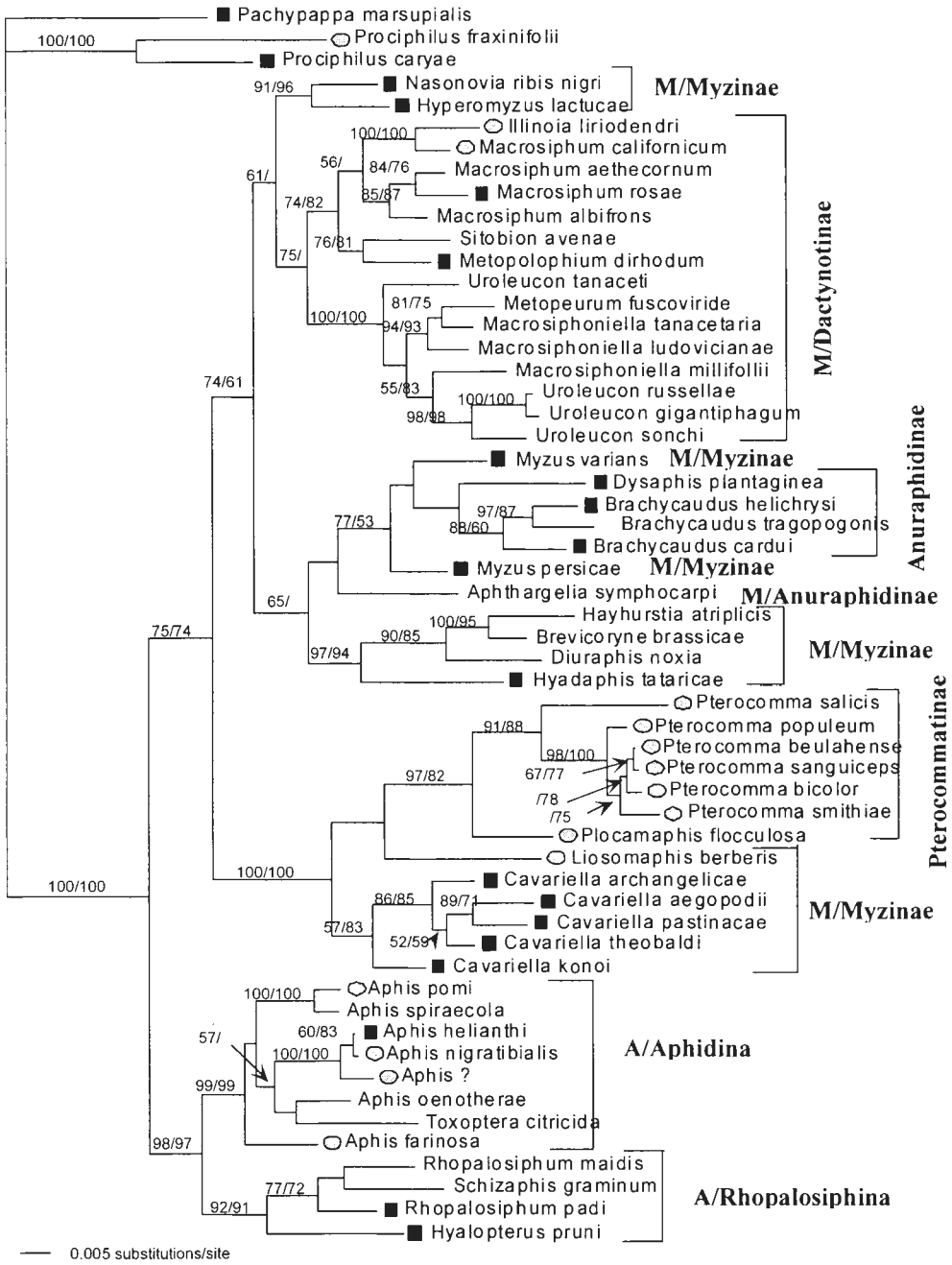


FIG. 9. Maximum likelihood (-ln likelihood = 12179.78) phylogeny of the combined EF1α + COII + trnL sequences (heuristic search with 10 random-additions and TBR branch swapping) under the best-fit evolutionary model (GTR+I+G, Lanave et al. 1984; Swofford et al. 1996). At each node, nonparametric bootstrap values (>50%) for maximum likelihood/maximum parsimony are given (500 and 1000 replicates, respectively). A = Aphidinae. M = Macrosiphini sensu Heie 1980. ■ = host-alternating. ○ = monoecious on a woody host.

Shimodaira-Hasegawa Likelihood Ratio Tests

The results from the SH test (Table 4) show support for the relationship of Pterocommatinae to *Cavariella* and *Liosomaphis*, as well as this clade's sister relationship to all other macrosiphines. All trees obtained from the MP analyses were not significantly different ($P > 0.05$) from the corresponding ML tree of that given data set for EF1α and the combined EF1α + COII + trnL data sets. All other alternative phylogenetic scenarios within a given data set were significantly different ($P \leq 0.05$) from the best ML tree, with the exception of the topology derived from the analysis of the EF1α data set under the constraint tree found in Figure 5C ($P = 0.210$). This topology, however, was significantly different ($P = 0.002$) from the ML topology of the combined EF1α + COII + trnL data set.

The difference in the placement of taxa at the species level observed from the MP and ML results between the EF1α data set and the combined EF1α + COII + trnL data set did not, in general, result in significantly different topologies (Table 4).

TABLE 4. Shimodaira-Hasegawa test results for the phylogenetic position of Pterocommatinae, *Cavariella*, and *Liosomaphis* to Macrosiphini and Aphidini, the relationships of Börner and Heinze's (1957) macrosiphine subfamilies, and the significance of finer level topologies. Log likelihoods were calculated under the same best-fit evolutionary models used in the ML analyses in this study.

		Data used to determine the new likelihood:			
		EF1α		EF1α + COII + trnL	
Source of Tree	Topology[a]	Ln likelihood	P -value	Ln likelihood	P - value
EF1α	ML tree	5041.26	optimal	12225.63	0.20
EF1α	Pterocommatinae	5095.19	0.00*	12297.96	0.00*
EF1α	Figure 5B	5090.53	0.01*	12291.90	0.00*
EF1α	Figure 5C	5073.68	0.12	12273.14	0.00*
EF1α	Figure 5A	5143.48	0.00*	12339.51	0.00*
EF1α	MP #1	5056.33	0.44	12236.07	0.09
EF1α	MP #2	5056.26	0.44	12232.45	0.11
EF1α	MP #3	5057.45	0.40	12235.40	0.09
EF1α	MP #4	5056.60	0.43	12245.10	0.05
EF1α	MP #5	5056.50	0.43	12241.39	0.06
EF1α	MP #6	5057.69	0.39	12244.40	0.05*
COII + trnL	ML tree	5061.28	0.49	12179.78	optimal
COII + trnL	Pterocommatinae	5115.19	0.00*	12246.59	0.01*
COII + trnL	Figure 5B	5146.07	0.00*	12286.50	0.00*
COII + trnL	Figure 5C	5121.77	0.00*	12245.15	0.01*
COII + trnL	Figure 5A	5193.62	0.00*	12346.76	0.00*
COII + trnL	MP #1	5076.40	0.16	12199.51	0.47
COII + trnL	MP #2	5074.96	0.21	12196.02	0.55
COII + trnL	MP #3	5075.18	0.18	12200.04	0.47
COII + trnL	MP #4	5078.93	0.12	12191.76	0.66
COII + trnL	MP #5	5078.93	0.12	12190.95	0.69

[a] ML = tree obtained from the ML analysis; Pterocommatinae = ML tree with the Pterocommatinae clade placed in the traditional basal position (Heie 1980); Figure 5A – 5C = tree topologies obtained from implementing the corresponding constraint trees from Figure 5 in an ML analysis under identical criteria used in the original searches in this study; MP = all MP trees obtained from the unweighted parsimony analysis.
* = $P \leq 0.05$

DISCUSSION

Phylogenetic Relationships Within Aphidinae

Phylogenetic analyses of mitochondrial (COII + trnL) and nuclear (EF1α) sequences, as well as SH tests between various tree topologies, supported the two proposed hypotheses in this study regarding the taxonomy of *Cavariella* and Pterocommatinae. First, the data unambiguously supported the sister relationship of *Cavariella* to Pterocommatinae by bootstrap replicates in both ML and MP analyses of the EF1α and the combined data set of EF1α + COII + trnL. The significant differences in tree topologies under the SH test provided further support for this sister relationship (Table 4). Significant differences were observed between the optimal ML tree (for both the EF1α and combined data sets), and all topologies in which *Cavariella* was constrained within Macrosiphini, and Pterocommatinae was excluded (Fig. 5A, B, D). Second, a sister relationship was supported for *Cavariella*/Pterocommatinae/*Liosomaphis* to other macrosiphines, but only by ≤ 75% of bootstrap replicates. Likewise, the SH tests between the optimal ML trees and the constraint topology depicted in Figure 5C gave significant differences in all comparisons, except in that between the EF1α ML tree and the constraint topology obtained under the EF1α data set ($P = 0.116$) (Table 4).

Highly supported relationships in this study that are consistent with current aphid taxonomy were the monophyly of Aphidina, Rhopalosiphina, Dactynotinae, and Rhopalosiphina, as well as the sister relationship of Aphidina to Rhopalosiphina (monophyly of Aphidini). The phylogenetic relationships within these clades, however, were not congruent across analyses of all data sets. These discrepancies (within tribes and

at the species level) were observed in various other clades within Macrosiphini, and are probably, in part, a result of the suitability of each gene for resolving the different phylogenetic levels. The EF1α sequences showed little evidence of saturation at the deeper nodes, whereas the COII + trnL sequences exhibited evidence of saturation beginning at an ML distance of approximately 0.10 (Fig. 6). A hint of saturation in EF1α corresponded to pairwise distances between *Pachypappa* (an outgroup member) and *Hyalopterus*, and between *Liosomaphis* and *Hyperomyzus*. However, the patterns of this saturation did not affect resolution at the phylogenetic levels of interest in this study. Conversely, saturation in the COII + trnL data corresponded to pairwise distances between genera and between most deeper relationships in general. Thus, homoplasy most likely accounted for the lack of phylogenetic support from the COII + trnL data set at these deeper levels of branching.

Several molecular phylogenetic relationships within Macrosiphini supported in this study contrast with the current morphological taxonomy. In addition to the relationship between *Cavariella* and Pterocommatinae (contradicting its current taxonomic position within Macrosiphini), there was bootstrap support for the paraphyly of Anuraphidinae, Myzinae, *Macrosiphoniella*, *Uroleucon*, *Myzus*, and *Macrosiphum*. The monophyly of Dactynotinae had strong bootstrap support in the MP and ML analyses of EF1α, and in the ML analysis of the combined data set, but there was no bootstrap support ≥50% in the MP analysis of the combined data sets. Furthermore, topologies constraining the monophyly of each of these three subfamilies were significantly different under the SH test. The topology constraining the strict monophyly of these three

clades (Fig. 5A) resulted in significantly different scores in all comparisons to the optimal ML tree obtained from the EF1α and combined data sets (Table 4). Comparisons of these ML trees to the topologies obtained under the constraint of Figure 5C (monophyly of Anuraphidinae, Dactynotinae, and Myzinae with the exception of *Cavariella* and *Liosomaphis*) resulted in significantly different scores, except in the comparison under EF1α. Further taxon sampling and more specific topology comparison tests are needed to better resolve the relationships within Macrosiphini.

Rapid Radiation Within Aphididae

A rapid evolutionary radiation event within Aphididae was suggested by several short, internal branch lengths observed in the ML trees in Figures 8 and 9. These short branch lengths corresponded to connections at the sub-tribal/tribal level (e.g., between *Nasonovia*/*Hyperomyzus* and Dactynotinae, between *Myzus varians* and Anuraphidinae, etc.). These results were also observed in the MP trees (not shown) from the analyses of the EF1α, COII + trnL, and combined data sets. Although branch lengths on an unweighted MP tree are not a direct function of rate and time as in ML trees (Felsenstein 1981), they are indicative of such events. The branch lengths on an unweighted MP tree represent the number of base substitutions between the given clades.

Additional support for the hypothesized rapid radiation event from the observed data was the suitability of the data as indicators of the evolution within Aphididae (i.e., the data did not obscure phylogenetic signal as a result of saturation). The plot of uncorrected pairwise distances against ML-corrected distances showed no evidence of saturation in the EF1α data set. There was some evidence of saturation in the COII + trnL

data set between genera and at much deeper branching events, however, COII + trnL was still useful for resolving relationshps at the species level and provided valuable information when used in conjunction with the EF1α data set.

Historical reconstructions of aphid diversification and life cycles, and host-plant diversification, provided further evidence for a rapid radiation event within Aphididae. Exhaustive work done by Heie (1990, 1994, 1996, 1998) on fossil aphids provided the historical context from which we can examine the relative time and means of aphid radiations. Under Heie's scenario, Aphididae, represented by only one (*Aphidocallis caudatus*) out of 63 known species from the Upper Cretaceous, underwent a rapid radiation between the beginning of the Miocene to the end of the Pliocene (Heie 1994, 1996, 1998). Prior to this radiation, Aphididae were assumed to be monoecious on various woody hosts, particularly the Rosales. In the mid-Tertiary, herbaceous angiosperms (grasses and forbs) displaced existing woody plant species in many parts of the world as the climate became drier and cooler. The success of Aphididae, now comprising over 60% of extant aphid species (Heie 1998), was attributed to their ability to initially acquire these herbs as secondary hosts and, in most cases, to eventually transfer their entire life cycle over to these new hosts as cases of secondary monoecy (Heie 1994). This ability was attributed to the reduced specialization of Aphididae (there is little variation between morphs which may allow them to more easily adapt with the changing environment), and their unique kind of host alternation. Within Aphididae, winged males and winged females return to the primary host. The winged females bear sexual females, as opposed to a winged sexuparae that give rise to wingless males and

mating females on the primary host. The two morphs that fly back to the primary host migrate independently of each other and at different times. This increases the chance of outbreeding and thus, the ability to acquire and maintain genetic variation necessary for evolution and radiation (Futuyma 1998).

Multiple Origins of Host Alternation

The scattered distribution of host-alternating taxa throughout Aphididae (Figs. 8, and 9), in addition to the proposed rapid radiation, has implications for the number of origins of host-alternation (von Dohlen and Moran 2000). Host-alternation may be pleisiomorphic (a shared ancestral character) or it may have had multiple origins within Aphididae. Aphid ecologists have traditionally accepted host-alternation as pleisiomorphic with subsequent, multiple losses in Aphidinae (Heie 1994). The universal mechanism of returning to the primary host via winged males and winged females that later bear sexual females, which is unique to Aphididae among aphids, was the basis of support for this hypothesis. Recent alternative hypotheses, however, support multiple origins of host-alternation throughout Aphidoidea (Moran 1988; von Dohlen and Moran 2000). Support for the latter hypothesis was drawn from the fundatrix constraint hypothesis (Moran 1988, 1992), which postulates that after a long evolutionary period of association with her host, the fundatrix became highly specialized on that host. The successive generations of the life cycle may then acquire more nutritive hosts, but must return to the primary host to accommodate the fecund and locomotively challenged fundatrix in the next generation (Shaposhnikov 1985; Moran 1988). Thus, the evolutionary transfers would be unlikely by specialized fundatrices in host-alternating

cycles between the unrelated host groups observed within Aphididae (Table 5), such as Salicaceae, Rosaceae, and Ericaceae. Rather, these woody hosts were more likely to have been acquired independently by ancestors with simple life cycles and unspecialized fundatrices, and, in numerous cases, in parallel. These events were then followed by multiple origins of host-alternation with subsequent monoecy on the secondary hosts.

My analyses support the hypothesis for multiple origins of host-alternation within Aphididae. Evidence for a rapid radiation contradicts a hypothesis for deep coevolution of a common Aphididae ancestor with the primary hosts, followed by the transfer of the fundatrices to unrelated hosts, as would be necessary to invoke if there were a single origin of host-alternation within Aphididae. Current evidence indicates that the fundatrix becomes specialized on the primary host only after a long, evolutionary association (Moran 1988). The extreme fundatrix specialization observed in other host-alternating aphid lineages has not yet evolved within Aphididae, and may account for their rapid radiation coupled with frequent host shifts. This lesser specialization of fundatrices has allowed Aphididae to capture a wider range of primary hosts than their aphid ancestors (Heie 1994). Thus, it is more conceivable that the relatively few, but taxonomically diverse woody hosts of Aphididae (Table 5) were acquired independently and, in many cases in parallel, associated with multiple origins of host-alternation.

The simple life cycles of all Pterocommatinae on woody hosts, as well as their newly observed phylogenetic position within Aphididae (Figs. 8 and 9), provide a further basis of support for the multiple origins of host-alternation within Aphididae. Secondary monoecy on the primary woody host, as would be necessary to invoke if the ancestor of

TABLE 5. Aphid genera and species within Macrosiphini (Börner and Heinze 1957) included within this study, life cycles, and major host associations. HA = host alternating. M = monoecious.

Anuraphidine

Aphthargelia : M; Asterids-Dipsicales-Caprifoliaceae (*Symphoricarpos*)

A. symphoricarpi : M; Caprifoliaceae

Brachycaudus : HA and M; HA from Rosales(Amygdalaceae) to Asterids (mostly Asterales-Asteraceae and Solanales-Boraginaceae).

B. cardui : HA; *Prunus* to Asteraceae (*Carduus*)

B. helichrysi : HA; *Prunus* to Asteraceae

B. trgpopogonis : M; Asteraceae (*Tragopogon)*

Dysaphis : HA and M; HA from Rosales (Pomaceae) to Asterids. M on Rosales or Asterids. Form leaf curls or galls.

D. plantaginea : HA from *Malus* to Lamiales (*Plantago*)

Dactynotinae

Illinoia : HA and M mostly M); Asterales (Asteraceae), Ericales (Ericaeae), Rosales (*Rubus*)

I. liriodendri : M; Magnoliales-Magnoliaceae (*Liriondendron tulipifera*)

Macrosiphoniella : M; Asteraceae

M. ludoviciane : M; *Artemisia*

M. millefolii : M; *Achillea*

M. tenacetaria : M; *Tenacetum*

Macrosiphum : HA and M (mostly M); Rosales (*Rosa*, *Rubus*) to herbaceous; Most M on herbs and shrubs.

M. aethecorunum : M: Geraniaceae (*Geranium*)

M. albifrons : M; Fagaes- Fagaceae (*Lupinus*)

M. californicum : M; Malpighiales-Salicaceae (*Salix*)

M. rosae : HA (sometimes M); Rosales (*Rosa*) to Dipsacaeae or Valerianaceae (M on *Rosa*)

Metopeurum : M; Asteraceae (*Tanacetum* spp.)

M. fuscoviride : M; *Tanacetum vulgare, Achillea millefolium*

Metopolophium : HA and M; Rosales (*Rosa*) to Poaceae

M. dirhodum : HA; Rosales (*Rosa*) to Poaceae

Sitobion : HA and M; Rosales(*Rosa*, *Rubus*) or Ericales (Ericaeae) to Poaceae, Polypodiophyta (ferns), or Equisetophyta (horsetails) (M typically on Poaceae)

S. avenae : M; Poaceae

Uroleucon : M; Asterales (Asteraceae and Campanulaceae)

U. gigantiphagum : M; Asteraceae (*Solidago*)

U. russellae : M; Asteraceae (*Gnaphalium*)

U. sonchi : M; Asteraceae (*Sonchus*)

U. tanaceti : M; Asteraceae (*Tanacetum*, *Chrysanthemum*)

TABLE 5 continued.

Myzinae

Brevicoryne : M; Brassicales (Brassicaceae)

B. brassicae : M; Brassicaceae (*Brassica*)

Cavariella : HA and M (one spp. M); Salicaceae to Apiales-Apicaceae

C. aegopodii : HA; Salicaeae (*Salix*) to Apiaceae

C. archangelicae : HA; *Salix* spp. to Apiaceae (*Angelica*)

C. konoi : HA; Salicaceae (*Salix*) to Apiaceae (*Angelica*, *Myrrhis*)

C. pastinacae : HA; Salicaceae (*Salix*) to Apiaceae (*Heracleum*, *Pastinaca*, *Angelica*)

C. theobaldi : HA; Salicaceae (*Salix*) to Apiaceae (*Pasinaca*, *Heracleum*)

Diuraphis : M; Poaceae

D. noxia : M; Poaceae (barley and wheat)

Hayhurstia : M; herbaceous

H. atriplicis : Caryophyllaceae (*Chenopodium*, *Atriplex*)

Hyadaphis : HA and M; Caprifoliaceae to Apiaceae (M on either primary or secondary host)

H. tataricae : Caprifoliaceae (*Lonicera*)

Hyperomyzus : HA and M; Saxifragales (*Ribes*) to Asterales (Asteraceae) or Laminales (Scrophulariaceae)

H. lactucae : Saxifragales (*Ribes*), Asteraceae (*Sonchus*)

Liosomaphis : M; Rannunculids-Berberidaceae (*Berberis* and *Mahonia*)

L. berberis :M; *Berberis*

Myzus : HA and M (mostly M); Rosales (Amygdalaceae, *Prunus*) to various herbaceous plants

M. persicae : HA; Rosales (*Prunus*) to herbaceous hosts (over 40 families)

M. varians : HA or M on secondary host; Rosales (*Prunus*) to herbaceous hosts (Ranunculaceae (*Clematis*))

Nasonovia : HA and M (mostly M and on secondary hosts); Saxifragaceae (*Ribes*) to Asterales, Laminales, and Solonales

N. ribisnigri : HA; *Ribes* spp. to Asteraceae (*Cichorium*, *Crepis*, *Hierachium*, *Lactuca*, *Lamsana*), Scrophulariaceae, and Solanaceae

Pterocommatinae had been host-alternating, has apparently never led to species radiations in other aphid lineages.

Conclusions

The data strongly support the proposed hypothesis of the sister relationship of Pterocommatinae to *Cavariella*, and contradict their affiliations based on morphological taxonomy. Support for the sister relationship of Pterocommatinae/*Cavariella*/*Liosomaphis* to the other macrosiphines, however, was not conclusive across all of the bootstrap analyses, particularly under the EF1α data set. Phylogenetic relationships within macrosiphine tribes were also ambiguous across the data sets. For instance, the relationships of *Metopolophium*, *Macrosiphum albifrons*, *Myzus varians*, etc. was contradicted between EF1α and the combined data sets, and not well supported by any bootstrap analyses. This ambiguity and lack of molecular resolution at the finer branching events is also observed in topological comparisons using the SH test. Comparisons between the optimal ML trees and all MP trees resulted in only one significant difference (between the optimal ML tree under the combined data set and one [out of six] MP tree obtained from the EF1α data set) (Table 4). Thus, not only must the relationships within Macrosiphini be subject to further molecular analyses, but the morphological attributes characterizing traditionally recognized clades must also be revisited.

A rapid radiation at the tribal/sub-tribal level, and multiple origins of host-alternation within Aphididae were also supported by the data. Most host alternating macrosiphines use hosts primarily in Rosaceae, Caprifoliaceae, Grossulariaceae, Poaceae, and Asteraceae (Table 5). Members of Aphidini use hosts in over 30 angiosperm

families. The rapid diversification within Aphididae did not allow for specialization of the aphid tribes on specific hosts. Thus, the ancester of Aphididae had a simple life cycle with several gains of host-alternation, possibly with more numerous gains within Aphidini as evidenced by the greater primary-host diversity and shorter period of time in which to specialize on these hosts.

The observed paraphyly of previously recognized subgroups within Macrosiphini (e.g., Börner and Heinze 1957) lends this intriguing tribe to further study. Improved resolution of their evolutionary relationships, through more intensive taxon sampling, could elucidate the broader evolutionary patterns in plant-host shifts and life cycle transitions. Further phylogenetic and morphological investigations of Pterocommatinae and their relationship to macrosiphines could also help answer questions such as: *why did pterocommatines remain on woody hosts when there are no apparent morphological constraints in their fundatrices?*

LITERATURE CITED

Bckman, R. L., and V. F. Eastop. 1984. Aphids on the world's crops: an identification and information guide. John Wiley and Sons, New York.

———. 1994. Aphids on the world's trees: an identification and information guide. CAB International. New York, NY.

Bner, C., and K. Heinze. 1957. Aphidina - Aphidoidea. Pp. 1-402 *in* P. Sorauer, ed. Aphidina - Aphidoidea. Paul Parey, Berlin.

Bkley, T. R., C. Simon, H. Shimodaira, and G. K. Chambers. 2001. Evaluating hypotheses on the origin and evolution of the New Zealand Alpine Cicadas (*Maoricicada*) using multiple-comparison tests of tree topology. Mol. Biol. Evol. 18(2): 223-234.

Etop, V. F. 1973. Deductions from the present day host plants of aphids and related insects. Pp. 157-178 *in* H. F. van Emden, ed. Deductions from the present day host plants of aphids and related insects. John Wiley and Sons, New York.

Fsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17: 368-376.

———. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Fuyma, D. J. 1998. Evolutionary biology. 3rd ed.. Sinauer Associates. Sunderland, MA.

Gdman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49(4): 652-670.

He, O. E. 1980. The Aphidoidea (Hemiptera) of Fennoscandia and Denmark. I. General Part. The families Mindaridae, Hormaphididae, Thelaxidae, Anoeciidae, and Pemphigidae. Scandinavian Science Press Ltd., Klampenborg, Denmark.

———. 1990. Recent advances in palaeoaphidology. Acta Phytopath. Ent. Hung. 25: 253-260.

———. 1994. Aphid ecology in the past and a new view on the evolution of Macrosiphini. Intercept Ltd., Andover, Hampshire, UK.

——. 15. The evolutionary history of aphids and a hypothesis on the coevolution of ıids and plants. Bollettino di Zoologia agraria e di Bachicoltura, Series II 28: 1455.

——. 13. Aphids of the past (Hemiptera, Sternorrhyncha). Proceedings of the First Inıational Palaeonetomological Conference, Moscow. AIA/AM/PFICM98/1.99:49-55.

Hillis nbers, D. 1966. Polymorphism in Aphididae. Annu. Rev. Entomol. 11: 47-78

Kis), Iand M. Hasegawa. 1989. Evaluation of the maximum likelihood estimates of thevolutionary tree topologies from sequence data, and the branching order in Hcnoidea. J. Mol. Evol. 29: 170-179.

Lan. C). Preparata, C. Saccone, and G. Serio. 1984. A new method for calculating evtionary substitution rates. J. Mol. Evol. 20: 86-93.

Mo:N. 1988. The evolution of host-plant alternation in aphids: evidence for spalization as a dead end. Amer. Natur. 132: 681-706.

—— 12. The evolution of aphid life cycles. Annu. Rev. Entomol. 37: 321-348.

Nork, B. 1996. Phylogeny and evolution of parthenogenetic weevils of the *4rigus tessellatus* species complex (Coleoptera: Curculionidae: Naupactini): evnce from mitochondrial DNA sequences. Evolution 50: 734-745.

—— 1). Evolution in a putatively ancient asexual aphid lineage: recombination and ·aıkaryotype change. Evolution 53: 1458-1469.

Palu, S. 1996. Nucleic acids II: the polymerase chain reaction. Pp. 205-247 *in* D. M.llis, C. Mortiz, and B. K. Mable, eds. Molecular systematics. Sinauer Ass:iates, Sunderland, MA.

Phil· HJ. Sörhannus, A. Baroin, R. Perasso, F. Grasse, and A. Adoutte. 1994. Comparison of molecular and paleontological data in diatoms suggests a major gaıı the fossil record. J. Evol. Biol. 7:247-265.

Pos:Dnd K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioformatics 14: 817-818.

Remè:G., and M. Remaudière. 1997. Catalogue des Aphididae du monde. Homptera Aphidoidea. INRA, Paris.

R, A., J. A. A. Nylander, F. Ronquist, and G. N. Stone. 2002. A maximum-likelihood analysis of eight phylogenetic markers in gallwasps (Hymenoptera : Cynipidae): Implications for insect phylogenetic studies. Mol. Phylogenet. Evol. 22: 206-219.

Sshnikov, G. C. 1985. The main features of the evolution of aphids. Pp. 19-99. Evolution and biosystematics of aphids. Proceedings of the International Aphidological Symposium, Fablonna, 1981. Polska Academia Nauk, Warsaw.

Sdaira, H., and M. Hasegawa. 1999. Multiple Comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16: 1114-1116.

S, C., F. Frati, A. Beckenbach, B. Crespi, H. Lui, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals Entomol. Soc. Am. 87(6): 651-701.

S D. L. 1994. A phylogenetic analysis of soldier evolution in the aphid family Hormaphididae. Proc. R. Soc. Lond. B 256: 203-209.

Scks, P., and D. F. Hales. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol. Biol. Evol. 13: 510-524.

Sord, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.

Sord, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pp. 407-514 *in* D. M. Hillis, C. Mortiz, and B. K. Mable, eds. Molecular systematics, 2nd ed. Sinaur Associates, Sunderland, MA.

T a, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10: 512-526.

vohlen, C. D., and N. A. Moran. 2000. Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. Biol. J. Linn. Soc. 71: 689-717.

vohlen, C. D., U. Kurosu, and S. Aoki. 2002. Phylogenetics and evolution of the eastern Asian-eastern North American disjunct aphid tribe, Hormaphidini (Hemiptera: Aphididae). Mol. Phylogenet. Evol. 23: 256-267.