Mitochondrial DNA Sequences and Multiple Data Sets: A Phylogenetic Study of Phytophagous Beetles (Chrysomelidae: *Ophraella*)

Daniel J. Funk, Douglas J. Futuyma, Guillermo Ortí, and Axel Meyer
Department of Ecology and Evolution, State University of New York at Stony Brook

This paper presents the phylogenetic infrastructure for an integrated historical and experimental study of host use evolution in the chrysomelid leaf beetle genus *Ophraella*. We report the collection of sequence data from the 16S ribosomal RNA (446 bp) and the cytochrome oxidase subunit I (420 bp) mitochondrial genes from 12 species of *Ophraella* and two outgroups. Sequence analysis revealed a strong A+T nucleotide bias, high interspecific COI sequence divergences (up to 21.4%) that greatly exceeded those for 16S (up to 5.9%), high intraspecific COI divergences (up to 3.8%), a dearth of amino acid substitutions in COI, and differing substitution patterns in ribosomal stems and loops. Intraspecific variation in COI haplotypes generally supported the genealogical coherence of *Ophraella* lineages, while suggesting two cases of paraphyletic species. Separate phylogenetic analyses of 16S and COI data sets yielded largely congruent trees. A combined 16S + COI analysis yielded a single shortest tree under maximum parsimony that was identical to trees provided by successive approximations, neighbor-joining, and maximum-likelihood methods. This topology proved robust to various forms of weighting and most nodes were highly supported (by bootstrap analysis). Separate parsimony analyses of mtDNA and previously collected morphological and electromorphic data sets revealed congruent estimates of all cladistic relationships except those within one clade. Analysis of the pooled data sets in a combined approach additionally provided support for the basal placement of two species from this clade, although the topology for the remaining species was weakly supported and incongruent with the mtDNA tree. Each data set contained significantly structured phylogenetic signal with respect to this clade, and data sets exhibited limited conflict (character incongruence) with each other. The combined data set, however, was found to lack phylogenetic signal. These observations may imply that pooling heterogeneously evolving classes of data obscured the phylogenetic signal in each, a potential limitation of the combined approach.

Introduction

Recent empirical and theoretical advances have provided increased confidence in our phylogenetic estimates for many groups of organisms. With this confidence has come a willingness to use phylogeny to generate and test evolutionary hypotheses. Such is the goal of the research program for which we have undertaken this phylogenetic study of the phytophagous beetle genus *Ophraella*. In this program, *Ophraella* phylogeny has been used to reconstruct the history of host plant use by these herbivorous insects. This history, in turn, has been employed to investigate the role of genetic variation as a constraint on host range evolution (Futuyma et al. 1993, 1994, 1995), to evaluate the likelihood of cospeciation of these beetles and their host plants (Futuyma and McCafferty 1990; Funk et al. 1995), and to generate intraspecific hypotheses for investigating the origin of particular *Ophraella* species and host shifts (Funk et al. 1995).

Ophraella (Wilcox 1965) includes 14 exclusively North American species of chrysomelid leaf beetles (LeSage 1986; Futuyma 1990, 1991). Both larvae and adults feed on the foliage of composites (Asteraceae), with various beetle species recorded from 10 genera belonging to 4 asteraceous tribes. Some *Ophraella* species appear to be strictly monophagous (using a single host plant species) while others feed on particular species from up to five host genera. (See table 1 in Funk et al. 1995 for distributions and host affiliations of particular species.)

A phylogeny for *Ophraella* was first provided by the analysis of morphological and electromorphic data sets (Futuyma and McCafferty 1990; see fig. 1). That study reported the resolution of *Ophraella* into three major clades but failed to resolve relationships within one clade of closely related species. As this clade includes

Key words: Chrysomelidae, combined approach, congruence analyses, insect-plant interactions, mtDNA sequences, nucleotide substitution patterns, phylogeny, total evidence, weighted parsimony.

Address for correspondence and reprints: Daniel J. Funk, Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY 11794-5245. E-mail: danfunk@life.bio.sunysb.edu.

Mol. Biol. Evol. 12(4):627–640. 1995.

0737-4038/95/1204-0011$02.00

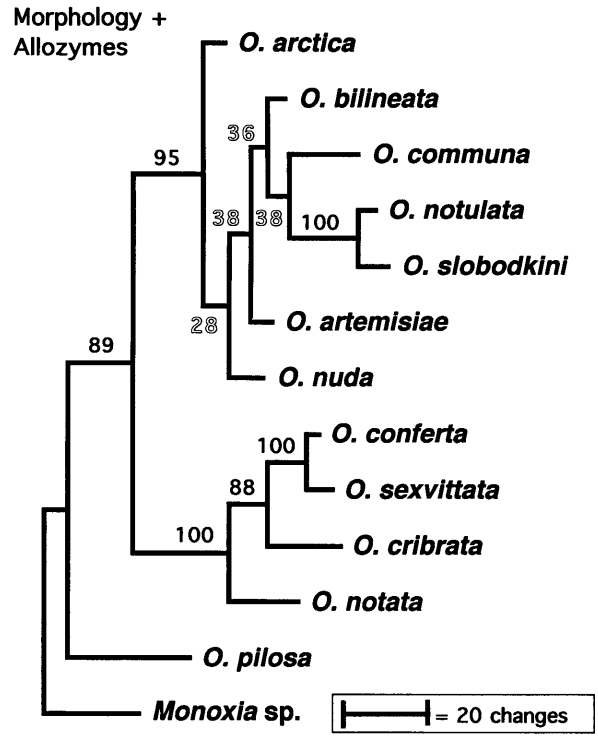

FIG. 1.—Single shortest tree from MP analysis of the morphology + allozymes data set under equal weights. Bootstrap proportions supporting each node are also presented. Note the low bootstrap support for nodes involving members of the communa subclade (see fig. 3*a*). The horizontal length of each branch is proportional to the assigned branch lengths from PAUP under ACCTRAN.

three of the focal species for the genetic studies, it was decided that a more confident estimate of these relationships was needed. To this end, we have collected and phylogenetically analyzed sequence data from portions of two genes from the rapidly evolving mitochondrial genome.

Because the phylogenetic estimates that we derive here are used elsewhere to test evolutionary hypotheses, the cladistic details of *Ophraella* relationships are of practical importance as well as taxonomic interest. Thus, we adopt a variety of phylogenetic algorithms, outgroups, and weighting schemes to evaluate the robustness of our mtDNA-derived tree. We also make use of the uncommon availability of three (presumably) independent sources of data: morphology, allozymes, and mtDNA. Whether such data sets are most appropriately analyzed separately and the resulting phylogenies compared, or pooled together and analyzed in a "combined approach," has been the subject of much controversy. We analyze our data sets both separately, using congruence analyses to evaluate agreement among data sets, and in combination. The results of these analyses illuminate potential strengths and weaknesses of combining data sets for phylogenetic analysis.

Material and Methods

Genes and Taxa

We collected and analyzed 866 base pairs of sequence data from portions of the large subunit ribosomal RNA (16S, 446 bp) and the cytochrome oxidase subunit I (COI, 420 bp) mitochondrial genes from single specimens of 12 of the 14 *Ophraella* species and from two outgroup taxa: a presumably undescribed species of *Monoxia* (the purported sister genus of *Ophraella*) (Futuyma and McCafferty 1990) and *Exema neglecta* Blatchley, from the distantly related chrysomelid subfamily Chlamisinae. Of the two *Ophraella* species excluded from our analysis, *O. americana* may be conspecific with *O. pilosa,* while *O. californiana* is known only from two specimens other than the type series and could not be found in the vicinity of the type locality by D. Futuyma. The species referred to elsewhere (Futuyma and McCafferty 1990) as *Ophraella* sp. nov. has since been designated *O. slobodkini* (Futuyma 1991).

We also collected COI sequences from additional specimens of most *Ophraella* species (including several geographic populations of *O. communa*) and from *Monoxia inornata.*

DNA Extraction, PCR, and Nucleotide Sequencing

After homogenization of entire frozen beetle specimens, total genomic DNA was extracted by proteinase K/SDS dissolution and purified by phenol-chloroform extraction and ethanol precipitation (Sambrook et al. 1989), and then "gene cleaned" to remove coprecipitating compounds which inhibited PCR (using Geneclean II, BIO 101). The polymerase chain reaction (PCR) (Saiki et al. 1988) was run for 30 cycles (45 s at 92°C, 60 s at 47°–50° C, and 90 s at 72°C) using 25-μl reactions containing 6.7mM $MgCl_2$, 1μM of each dNTP, 0.5–1.0 μM of each primer, template DNA, and 1 unit of *Taq* polymerase in Tris buffer (67 μM, pH 8.8). Primers used were 16S A and B (Palumbi et al. 1991) and S1751 (5′ GGA TCA CCT GAT ATA GCA TTC CC 3′)/A2191 (5′ CCC GGT AAA ATT AAA ATA TAA ACT TC 3′) (for COI, developed in the lab of R. Harrison, Cornell University; names correspond to the position of their 3′ end based on *Drosophila yakuba* sequence [Clary and Wolstenholme 1985]). Five microliters of the double-stranded PCR products were gel purified on 2.5% NuSieve agarose minigels in Tris-acetate buffer (0.1 M, pH 7.2) and stained with ethidium bromide. Target products were excised, dissolved in 500–1,000 μl ddH_20, and used for asymmetric PCR (Gyllensten and Erlich 1988) which was run for 35 cycles (at 55°C for annealing step) using 50-μl reactions with one primer in limiting concentration of 1:100. These products were ultrafiltrated three times using 300 μl H_20 in spin columns

(Millipore 30,000) prior to direct sequencing (Sequenase Version 2.0, U.S. Biochemical) (Sanger et al. 1977) with the limiting primer from asymmetric PCR. Radiolabeled sequencing products were resolved by vertical electrophoresis on long 6% acylamide-urea gels in TBE buffer and visualized by autoradiography. For each specimen studied, sequence was obtained from both sense and antisense strands. The considerable overlap of strands allowed confident assessment of nucleotide identity for all sequence analyzed.

Sequence Analysis

Sequence data were edited using ESEE, a multiple-sequence editor (Cabot and Beckenbach 1989). CLUSTAL V (Higgins and Sharp 1989) was used to align sequences using the default parameters. Using *Drosophila yakuba* as a model (Gutell and Fox 1988), we inferred the secondary structure of the 16S sequences and assigned individual sites to "stems" or "loops" according to whether they exhibited complementarity. Using MEGA (Molecular Evolutionary Genetic Analysis; Kumar et al. 1993), substitution patterns were characterized using the distance and statistics options for variously partitioned subsets of our sequence data. These analyses considered all possible pairwise comparisons of 16S and COI sequences from those *Ophraella* specimens sequenced for both (i.e., one per species).

Weighting Strategies for mtDNA

Given a priori evidence that various classes of characters are evolving under different rules, differential weighting of these classes may be justified (Swofford and Olsen 1990). For example, as transitions (TIs) generally accumulate more rapidly than transversions (TVs), the strength of their phylogenetic signal often decreases with increased sequence divergence due to multiple substitutions, which erase prior history and introduce homoplasy (Brown et al. 1982; Li et al. 1984; DeSalle et al. 1987). For this reason, TVs are often weighted more heavily in phylogenetic analyses (Swofford and Olsen 1990). Similarly, differences in the evolutionary tempo and mode of ribosomal RNA stem and loop structures due to functional constraints can be used as a basis for weighting (see, e.g., Vawter and Brown 1993).

As an objective guide to weighting, we attempted to infer the native TI/TV ratios in *Ophraella* COI and 16S genes by estimating the rate at which TIs and TVs accumulate in paired sequence comparisons unobscured by multiple hits (fig. 2). This was done by plotting the number of TIs versus the number of TVs for all pairwise sequence comparisons and then calculating a Model II linear regression using those comparisons in which TIs were unsaturated, that is, those on the linear portion of the curve (after Sturmbauer and Meyer 1992). This method provided TI/TV ratios that we applied as initial weighting factors in the maximum-parsimony analyses using stepmatrices.

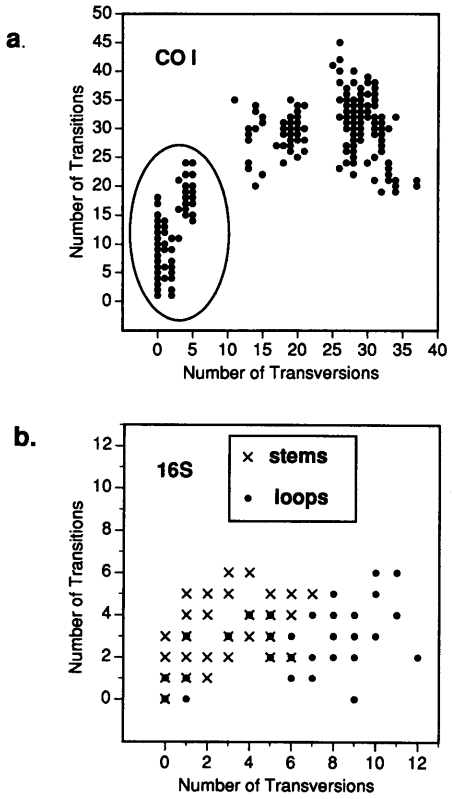

FIG. 2.—Relative substitution rates of transitions versus transversions for (*a*) COI and (*b*) 16S stems and loops. These plots consider all possible pairwise comparisons among the individual 16S haplotypes of *Ophraella* species and among all COI haplotypes. The portion of the curve used to estimate the native TI/TV ratio for COI is circled. The 16S plot includes fewer points because it was generated from fewer haplotypes and because some points are superimposed.

An alternative to deriving weights from external criteria is to let the characters decide their own weights based on their fit to an initial topology (Farris 1969; Williams and Fitch 1989). Using PAUP, we generated an mtDNA tree using equal weights to provide rescaled consistency indices for initial a posteriori weighting. These were then applied iteratively in a successive approximations approach (Farris 1969).

Phylogenetic Analysis of mtDNA

We inferred the phylogenetic relationships among *Ophraella* haplotypes using maximum-parsimony (MP) (PAUP version 3.1.1; Swofford 1993), successive approximations (Farris 1969) (PAUP), neighbor-joining (NJ) (MEGA; Kumar et al. 1993, using Kimura distances), and maximum-likelihood (ML) (PHYLIP ver-

sion 3.42; Felsenstein 1991, using TI:TV ratio of 3.0, "jumble" option, and global rearrangements) methods. For MP analyses of both the separate and combined 16S and COI data sets, shortest trees were found with the branch-and-bound algorithm. Support of the data for various clades was assessed using heuristic bootstrap searches with random addition of taxa (10 repetitions for each of 100 bootstrap replications). These searches sampled all characters with equal probability and then applied any assigned weights. The bootstrap (Felsenstein 1985) was also used with NJ and ML analyses.

The sensitivity of MP results to choice of outgroup was examined under three weighting strategies (equal weights, TV = 3TI, and TVs only; see Results) by either designating *Exema* as the outgroup or by excluding it and using *Monoxia* sp.

The sensitivity of these results to weighting was evaluated by varying weights for both genes from the inferred native TI/TV ratio of COI and observing the effect on topology. The COI ratio was used for 16S analyses as well because the 16S TI/TV ratio could not be accurately inferred (fig. 2; see below). However, guided by the observed TI/TV ratios in stems and loops (see Results), we also conducted 16S and combined (16S + COI) analyses in which TIs in loops were excluded and those in stems were downweighted (TV = 3TI).

Phylogenetic Analysis of Multiple Data Sets

Support for *Ophraella* relationships provided by the morphology, allozyme, and mtDNA data sets was evaluated using (1) bootstrap support for separate MP analyses of each data set; (2) bootstrap support for MP analysis of the pooled data sets using the combined approach (Kluge 1989); and (3) the degree of phylogenetic congruence among the topologies from (1) and (2). Observed incongruencies were investigated with reference to the presence or absence of phylogenetic signal within data sets (Hillis and Huelsenbeck 1992) and the degree of conflict (character incongruence) among data sets (Kluge 1989). The MP analyses of the morphology and allozyme data sets employed branch-and-bound searches exclusively. These treated 64 morphological characters invariant within individual *Ophraella* species and 144 alleles (all with frequency >0.05) from 19 loci, with each allele coded as a binary (presence/absence) character (Swofford and Berlocher 1987; see Futuyma and McCafferty 1990 for more details). Of these, 48 morphological and 44 allozyme characters provided support for groupings within *Ophraella.* Morphological characters were not scored for *O. sexvittata* but were set equal to those of *O. conferta* as these taxa are distinguishable by only a single minor character and may be geographic variants of the same species (LeSage 1986; Futuyma 1990). A successive approximations analysis of the combined data set was also conducted.

Intraspecific Analysis of COI Haplotypes

The phylogenetic coherence of haplotypes from specimens of individual *Ophraella* species was assessed through MP analysis of all COI sequences of all taxa from the intraspecific survey. This analysis employed heuristic searches and used equal weights as we were interested in intraspecific relationships that were largely supported by TIs, which were unsaturated at this phylogenetic level. An NJ analysis of this data set was also conducted.

Results and Discussion

Sequence Variation and Modes of Substitution

Previous studies of insect mitochondrial DNA (see, e.g., Clary and Wohlstenholme 1985; Crozier et al. 1989; Crozier and Crozier 1992; Liu and Beckenbach 1992; Pashley and Ke 1992; Tamura 1992; Beckenbach et al. 1993; Fang et al. 1993; Brower 1994; Brown et al. 1994; reviewed in Simon et al. 1994) have revealed many of the patterns exhibited by our sequence data. COI sequences included no indels (insertions/deletions), while indels totaling 4 bp were inferred in 16S sequences with reference to *Exema.* No indels were inferred between 16S haplotypes of *Ophraella* species. We found a prominent A+T nucleotide bias that was greater in 16S than in COI, somewhat greater in stems than in loops, and much greater in third-base positions than in first and second positions, in which little bias exists (table 1). A<->T substitutions accounted for 70% of TVs in COI and 81% of TVs in 16S.

The overall TI/TV ratio was roughly twice as high in COI as in 16S and much higher at low sequence divergences in COI (table 2). 16S exhibited a TV bias even at rather low sequence divergences, while the 16S TI/TV ratio was three times greater in stems, in which it was identical to that of COI, than in loops, in which TIs may be completely saturated. Figure 2 documents the saturation of TIs (at approximately 5% uncorrected sequence divergence) in COI, and its slope provides a native TI/TV estimate of 3:1. 16S stems and loops; however, both exhibit considerable scatter. This is consistent with previous observations that both regions contain sites with heterogeneous rates of evolution (Simon et al. 1994) but is incompatible with confident estimation of native TI/TV ratios for 16S.

A higher overall substitution rate in COI is largely accounted for by third-position base changes, while no second-position changes were observed. Amino acid replacements were inferred at only eight sites. Codon usage was greatly nonrandom, as has often been found in taxa

Table 1
Nucleotide Substitution Patterns for Subsets of mtDNA Sequence Data

mtDNA Subset	bp	%A	%T	%C	%G	TI/TV	% Seq. Div.	Variable Sites	Informative Sites
16S all	442	37.3	41.2	6.7	14.8	0.64	3.4	50	22
16S "stems"	244	34.7	42.0	7.5	15.8	1.21	2.4	20	7
16S "loops"	198	40.6	40.2	5.6	13.6	0.41	4.6	30	15
COI all	420	28.9	37.1	18.6	15.4	1.21	12.6	129	91
COI 1st pos.	140	29.3	26.9	18.5	25.2	∞[a]	5.7	22	14
COI 2d pos.	140	14.0	42.7	26.1	17.2	0.00	0.0	0	0
COI 3d pos.	140	43.5	41.6	11.2	3.6	0.91	32.4	107	77
16S + COI	862	33.2	39.2	12.5	15.1	1.05	7.9	179	113

NOTE.—All proportions are based on the mean of all pairwise comparisons of *Ophraella* haplotypes for specimens sequenced for both 16S and COI (i.e., one per species). TI/TV, transition/transversion ratio = (no. inferred TIs)/(no. inferred TVs). Percentage uncorrected sequence divergence = (no. of differences)/(no. base pairs compared).

[a] 22 TIs and zero TVs.

with pronounced A+T or G+C bias (see, e.g., Crozier and Crozier 1992). Interspecific *Ophraella* sequence divergences (corrected using Kimura's [1980] two-parameter model) ranged from 0.7% to 21.4% in COI and from 0.2% to 5.9% in 16S (table 2). Intraspecific COI variation (also corrected) ranged from 0.0% to 3.8%.

Support for mtDNA Phylogeny

The fruitful use of a phylogeny for testing evolutionary hypotheses demands a confident estimate of the relationships of interest. Several sources of evidence suggest the reliability of our mtDNA-derived estimate of *Ophraella* phylogeny (fig. 3*a*).

Phylogenetic Methods

As various models of phylogeny reconstruction make varying assumptions about evolutionary process, each is susceptible to different kinds of bias in the data, with the result that different methods may provide contradictory phylogenetic estimates for a given data set (Felsenstein 1988; Swofford and Olsen 1990). However, MP and successive approximation analyses of the combined (16S+COI) data set agreed on a single shortest tree that was completely resolved and topologically identical with those from NJ and ML analyses (fig. 3*a*). This agreement suggests that our data may be relatively free of "positively misleading" systematic biases (Felsenstein 1978) and increases our confidence in the recovered topology (Kim 1993).

Statistical Support

The degree to which a particular data set supports groupings within a given topology can be estimated using statistical resampling methods such as the bootstrap (Felsenstein 1985). Our MP analysis yielded reasonably high overall bootstrap proportions, and strong support for certain groups within the important "slobodkini clade" was found by all methods (fig. 3*a*). Although only a few nodes were supported by bootstrap proportions of >95%, Hillis and Bull (1993) report that bootstrap proportions of >70% generally correspond to a probability of >95% that the data consistently support a given clade. ML analysis, which provides another means of statistical assessment, found all branch lengths but one (that between *O. bilineata* and its ancestor) to be significantly positive.

Different Genes

Although genes on the nonrecombining mitochondrial genome of animals are not independently evolving loci, they exhibit considerable heterogeneity of evolutionary rate (reviewed in, e.g., Simon 1991; Meyer 1994; Simon et al. 1994) and sometimes yield incongruent phylogenetic estimates (see, e.g., Cameron et al. 1992, using 16S and COII data for *Apis*). However, 16S and COI trees were largely congruent (fig. 3*b, c*). Although the 16S slobodkini clade relationships are not resolved, the bootstrap support for these groupings offered by COI is robust to the addition of 16S data, and these data sets exhibit little apparent conflict (see below).

Weighting Strategy

The accurate inference of nucleotide substitution patterns allows weights to be assigned that enhance phylogenetic accuracy. As it is often difficult to objectively decide among competing topologies that vary according to the weights applied (see, e.g., Edwards et al. 1991; Liu and Beckenbach 1992; Helm-Bychowski and Cracraft 1993), topological robustness to varying weights provides another measure of confidence. Our parsimony analyses, adopting a variety of weighting strategies (table 3 and below), suggest both the suitability of our inferred

Table 2
Intraspecific Variation in COI Haplotypes and Interspecific Divergence among 16S (below Diagonal) and COI (above) Haplotypes

	Intraspecific Variation	Interspecific Divergence													
		arc	bil	com	art	nud	ntl	slb	con	sex	cri	not	pil	Mon	Exe
arc	...		0.7	2.1	4.8	5.3	12.8	16.2	16.7	17.5	15.5	16.2	18.1	16.4	24.4
	(1)		1/2	6/2	15/4	17/4	29/19	32/28	29/32	31/31	30/28	33/27	34/27	27/28	35/50
bil	0.5–1.0	0.2		2.3	5.0	6.4	13.2	16.8	17.1	18.5	15.9	16.8	19.9	17.1	24.4
	(3)	1/0		9/0	15/5	20/5	30/20	33/29	30/33	34/31	31/28	33/29	36/31	28/30	36/50
com	0.0–4.7	0.5	0.2		4.9	0.0	12.2	16.1	16.8	17.7	15.2	17.1	20.2	16.5	25.0
	(12)	2/0	1/0		14/5	18/5	25/20	29/29	28/32	30/31	27/28	32/29	35/31	26/30	35/50
art	2.8	1.1	0.7	0.9		4.2	12.9	16.9	17.5	18.7	15.7	16.3	18.9	16.8	26.9
	(2)	3/2	1/2	2/0		17/0	28/21	33/30	34/30	38/28	32/27	35/26	38/26	28/29	44/49
nud	0.0	0.2	0.0	0.2	0.9		14.7	18.8	18.1	19.3	16.9	17.9	21.4	18.9	28.4
	(2)	1/0	0/0	1/0	2/2		34/21	39/30	36/30	40/28	36/27	40/26	45/26	34/29	48/49
ntl	1.3–1.8	3.5	3.1	3.4	3.5	3.3		12.4	14.7	15.0	13.1	16.8	18.1	16.8	25.8
	(3)	8/7	6/7	7/7	7/8	7/7		32/15	22/33	24/30	22/28	35/27	31/31	23/34	40/50
slb	0.2–0.7	3.0	2.6	2.9	3.0	2.8	2.8		14.4	16.4	16.9	16.9	16.7	18.1	23.4
	(4)	7/6	5/6	6/6	6/7	6/6	5/7		20/34	27/32	34/29	33/30	21/37	28/33	34/49
con	...	4.7	4.2	4.5	4.4	4.4	3.5	3.3		2.9	9.6	11.6	16.2	16.9	22.6
	(1)	7/13	5/13	6/13	6/13	6/13	3/12	5/9		10/1	22/15	30/14	23/33	24/33	33/47
sex	...	4.6	4.4	4.6	4.9	4.4	4.7	3.6	1.4		9.9	12.6	16.5	17.5	22.9
	(1)	7/12	6/12	7/12	7/13	6/12	6/13	5/10	5/1		24/13	35/11	24/33	27/32	33/45
cri	1.2	5.9	5.4	5.8	5.9	5.7	5.0	4.0	3.0	2.4		11.1	17.8	17.4	24.1
	(2)	12/12	10/13	11/13	10/15	11/13	8/13	7/10	7/6	6/4		30/13	29/32	25/34	35/50
not	0.0	5.6	5.2	5.3	5.4	5.4	4.5	3.5	3.2	3.1	3.5		19.4	20.2	26.9
	(2)	10/14	8/14	9/13	9/14	9/14	5/14	4/11	5/9	5/8	6/9		32/34	30/37	42/51
pil	...	4.0	3.5	3.8	3.7	3.7	4.5	2.8	4.0	4.1	5.4	4.7		17.3	27.8
	(1)	6/11	4/11	5/11	5/11	5/11	6/13	4/8	5/12	7/10	9/14	5/15		23/34	38/51
Mon	...	7.2	6.9	7.0	7.4	6.9	7.3	5.5	6.9	6.9	7.5	7.2	7.2		27.8
	(1)	6/24	5/24	6/23	5/26	5/24	6/24	2/21	4/25	4/24	4/27	4/26	5/25		42/45
Exe	...	22.6	22.1	21.6	22.6	22.3	21.5	20.6	21.7	20.2	22.4	20.8	21.7	20.0	
	(1)	22/63	21/62	22/58	22/63	21/63	22/58	18/60	20/62	17/57	21/63	21/58	22/60	22/54	

NOTE.—Presented range of percentage sequence divergences among COI haplotypes within *Ophraella* species, corrected for multiple hits using Kimura's (1980) two-parameter model, with number of specimens sequenced given in parentheses; and percentage sequence divergences and number of inferred TIs/TVs from pairwise comparisons of haplotpyes for specimens sequenced for both 16S and COI (1 per species). Species abbreviations are as follows: arc, *O. arctica,* bil, *O. bilineata;* com, *O. communa;* art, *O. artemisiae;* nud, *O. nuda;* ntl, *O. notulata;* slb, *O. slobodkini;* con, *O. conferta;* sex, *O. sexvittata;* cri, *O. cribrata;* not, *O. notata;* pil, O. pilosa; Mon, *Monoxia* sp.; Exe, *Exema neglecta.* For particular pairwise divergences among all *O. arctica, O. bilineata, O. communa, O. artemisiae,* and *O. nuda* haplotypes, see table 2 in Funk et al. (1995).

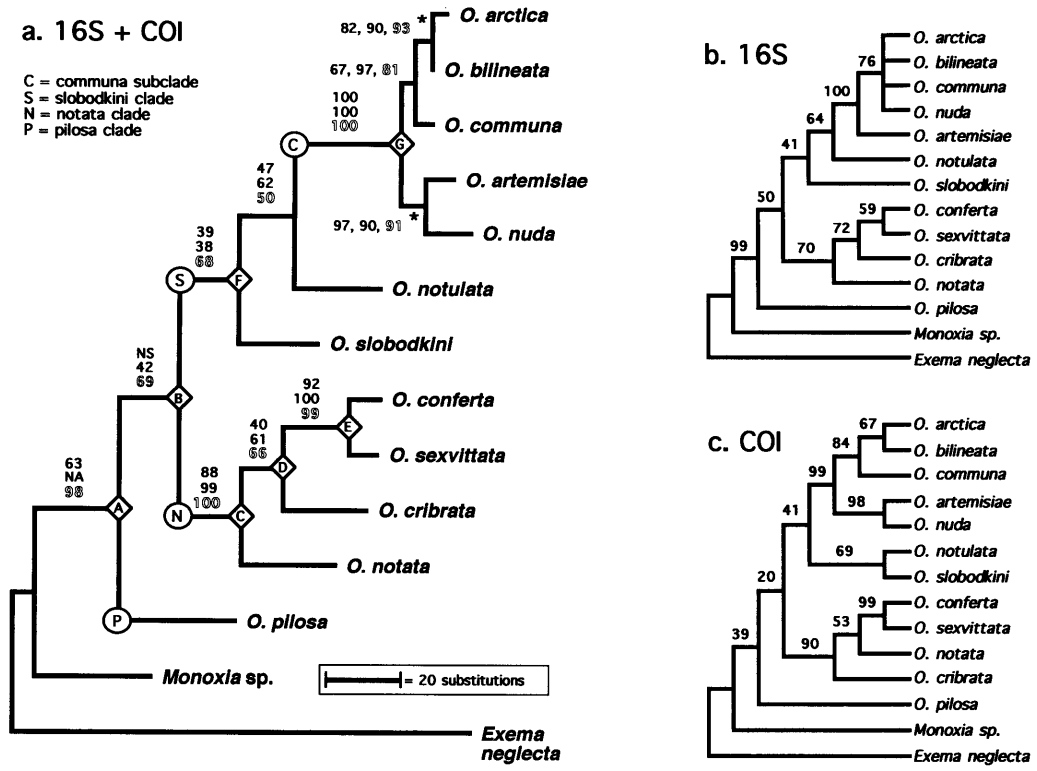

FIG. 3.—Strict consensus of shortest trees under maximum parsimony for (*a*) 16S+COI, (*b*) 16S, and (*c*) COI analyses. These analyses weighted TVs three times as heavily as TIs based on the inferred native TI/TV ratio of COI. The MP topology of *a* is identical to that of successive approximations, NJ, and ML analyses. The bootstrap proportions from the MP analyses are presented for each tree, while those from MP, NJ, and ML (*lower to upper*) analyses are provided for the 16S + COI analysis. *NA,* no value because *Monoxia* was the outgroup in NJ analysis; *NS,* no value because grouping was not supported in ML bootstrap tree. Asterisks indicate nodes largely defined by transitions, for which bootstrap proportions from the TV = 1.1TI analysis (see table 3) are presented. Circled letters define major clades referred to in the text; letters within diamonds define clades referred to in table 4. The horizontal length of each branch is proportional to the assigned branch lengths from PAUP under ACCTRAN for the equally weighted data.

TI/TV ratio (3:1, from fig. 2) as a weighting scheme and the robustness of *Ophraella* phylogeny to choice of weights.

In our analyses, increased weighting of TVs generally provided fewer, more highly resolved trees, but completely excluding TIs yielded less resolution (table 3). An observed decrease in bootstrap support for recent nodes, and the corresponding increase for more basal nodes (not illustrated here), under increased weighting of TVs, suggested that TIs provide information about

Table 3
Maximum-Parsimony Analyses of mtDNA: Effects of Differential Weighting of Transitions and Transversions on Phylogenetic Resolution and Topology

	TV = TI			TV = 1.1 TI			TV = 3TI			TV = 9TI			TV = 1, TI = 0		
	t	*n*	*i*	*t*	*n*	*i*	*t*	*n*	*i*	*t*	*n*	*i*	*t*	*n*	*i*
16S	24	7	v	12	9	v	12	9	v	12	9	v	2	9	v
COI	3	7	w	1	12	wxy	1	12	wxy	1	12	wxz	8	9	xz
16S + COI	2	11	...	1	12	...	1	12	...	1	12	xz	2	11	xz

NOTE.—Number of equally parsimonious shortest trees obtained (*t*), number of resolved nodes (*n*) (out of 12) in a consensus of shortest trees, and a list of incongruencies (*i*) with the mtDNA tree (=16S + COI, TV = 3TI topology [fig. 3*a*]) are presented for analyses from each of five different strategies for weighting transversions relative to transitions. Incongruencies supported by bootstrap proportions of ≥50% are underlined: *v,* art basal to clade composed of arc/bil/com/nud; *w,* placement of cri/not reversed; *x,* ntl/slb paired as sister taxa; *y,* Mon placed basal to slobodkini clad; *z,* bil/com paired as sister taxa. *Exema* was used as the outgroup in analyses summarized in this table.

Table 4
Maximum-Parsimony Analyses of Multiple Data Sets

Data Set	Trees	Nodes	CI	RCI	A	B	C	D	E	F	G	S	I
16S	12	9	0.871	0.677	99	50	70	72	59	41	100	42	...
COI	1	12	0.621	0.362	39	20	90	53	99	41	99	83	w1, x
16S + COI	1	12	0.688	0.420	98	69	100	66	99	68	100	88	...
Morphology	9	5	0.685	0.531	—	71	100	58	81	50	*	34	...
Allozymes	7	8	0.729	0.468	—	77	87	90	100	83	*	48	w2, y
Morph. + Allo.	1	12	0.684	0.472	—	89	100	88	100	95	*	35	w2
Combined Data	1	12	0.678	0.463	—	89	100	91	100	99	100	33	w3, z

NOTE.—Number of equally shortest, most parsimonious trees, number of nodes (out of 12) resolved by a strict consensus of these trees, consistency indices, and rescaled consistency indices are presented from analyses of each data set. Also shown are bootstrap proportions supporting nodes defined in fig. 3*a*(A–G), average of bootstrap proportions supporting groupings of communa subclade species (S), and incongruencies (I) with the mtDNA topology of fig. 3*a* that are supported by bootstrap proportions of ≥50%: *w*, ntl/slb are sister taxa (w1: bootstrap = 69, w2: *b* = 100, w3: *b* = 95); *x*, cri/not placement reversed (b = 52); *y*, arc basal to remainder of *slobodkini* clade (*b* = 61); *z*, bil/com are sister taxa (*b* = 53). For these analyses, TVs were weighted three times as heavily as TIs, and morphological and electrophoretic characters were weighted as transversions in the combined approach. Outgroup was *Exema* for mtDNA analyses and *Monoxia* for others. A dash (—) indicates absence of Exe from analysis, so no value; an asterisk (*) indicates group not recovered.

recent divergences but contribute largely noise to the inference of ancient relationships. For example, the position of *Monoxia* sp. within *Ophraella* in certain COI analyses was rectified when TVs were more heavily weighted. These observations suggest the utility of weighting TVs but not eliminating TIs. Besides the misplacement of *Monoxia,* the only other deviations from the TV = 3TI topology (fig. 3*a*) introduced by differential weighting of TVs and TIs was the pairing of *O. bilineata* with *O. communa* rather than with *O. arctica* and the pairing of *O. notulata* and *O. slobodkini* as sister taxa, both under stronger downweighting of TIs (table 3). The *O. notulata/O. slobodkini* grouping was also recovered when TIs from 16S loops were ignored. These analyses (not illustrated) yielded 16S and 16S + COI trees which were otherwise identical to figure 3*a* (for each COI weighting scheme). Successive approximations also provides the topology of figure 3*a*. The robustness of our topology to various forms of a priori and a posteriori weighting provides confidence in our phylogenetic estimate.

Choice of Outgroup

Using different outgroups sometimes yields greatly dissimilar ingroup topologies, due in part to the loss of phylogenetic signal in distantly related outgroups (Wheeler 1991; DeSalle 1992; Helm-Bychowski and Cracraft 1993; Pashley et al. 1993). However, use of the closely related *Monoxia* sp. versus the very genetically divergent *Exema neglecta* (table 2) had no effect on *Ophraella* topologies under all weighting schemes in the combined 16S + COI analyses.

Congruence Analyses and the Combined Approach

Much recent controversy has concerned how to best analyze separate data sets that represent conventionally defined "classes" of characters such as the morphology, allozyme, and mtDNA data collected from *Ophraella.* Some authors assert that the division of data into subsets is an arbitrary procedure and point out that phylogenetic information may be lost by failing to conduct a combined approach analysis of all data sets (Kluge 1989; de Queiroz 1993; Kluge and Wolf 1993). Others have observed that combining heterogeneously evolving data sets may worsen phylogenetic estimates (Bull et al. 1993), that contradictions between well-supported topologies—suggesting systematic bias and the likelihood of spurious results—can only be detected by the comparison of separate analyses (de Queiroz 1993), and that topological agreement among data sets conceivably evolving under disparate rules provides a means of phylogenetic corroboration (Swofford 1991). We have analyzed our data sets both separately and in combination to assess the contribution of each to the resolution of *Ophraella* relationships.

MP analyses of the morphology and allozyme data sets (fig. 1) yielded topologies that agree with our mtDNA estimate (fig. 3*a*) with respect to those relationships that Futuyma and McCafferty (1990) claimed to have confidently recovered, namely, the constitution of, and relationships among, the slobodkini, notata, and pilosa clades and the groupings within the notata clade (fig. 3*a*). All these relationships were supported by bootstrap proportions of ≥50% in each analysis (table 4). Analysis of the pooled data sets using the combined approach also recovered these groupings and supported them with high overall bootstrap proportions (fig. 4, table 4). This congruence suggests much explicit agreement among data sets in phylogenetic signal.

However, the consistent and confident estimation of these *Ophraella* relationships contrasts sharply with the irresolution involving slobodkini clade groupings.

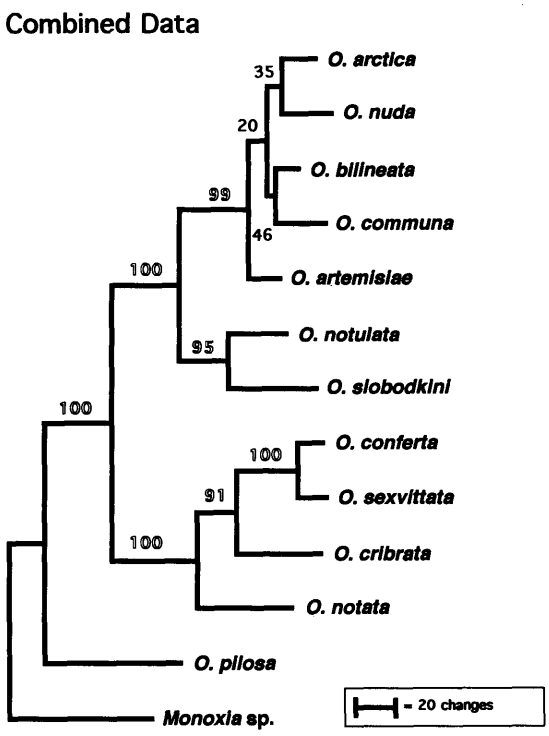

FIG. 4.—Single shortest tree from MP analysis of the combined data set in which TVs were weighted three times as heavily as TIs and morphological and allozyme characters weighted as TVs. Bootstrap proportions supporting nodes are also presented. The horizontal length of each branch is proportional to the assigned branch lengths from PAUP under ACCTRAN for the equally weighted data.

Other than the pairing of *O. notulata* and *O. slobodkini* in the allozyme and morphology analyses, no relationships within this clade were common to majority-rule bootstrap trees from any two of the separate analyses. Further, although mtDNA provided strong bootstrap support for most nodes in the slobodkini clade (fig. 3*a*), the morphology and allozyme analyses supported only one such grouping with a bootstrap proportion as high as 50% (table 4), an indication of the uncertainty Futuyma and McCafferty (1990) expressed about relationships within this clade.

The combined-approach analysis partially resolved these incongruities, providing strong bootstrap support for the union of *O. notulata* and *O. slobodkini* as sister species and for their placement basal to the remaining slobodkini clade species (the communa subclade) (fig. 4). This result suggests implicit agreement among data sets in phylogenetic signal and illustrates the capacity of the combined approach to reveal relationships unsupported by separate analyses. However, the combined approach failed to resolve relationships among communa subclade species, insofar as the addition of morphological and electromorphic characters to the mtDNA data set yielded greatly *lowered* bootstrap support for these relationships (fig. 4). Additionally, for analyses of the combined data set, the single shortest tree from the weighted data (fig. 4) and the single shortest tree yielded by both successive approximations and MP analyses of the equally weighted data (not illustrated) differ from each other and from the mtDNA tree (fig. 3*a*) in communa subclade topology.

These anomalous results provide an opportunity to explore the circumstances under which a combined approach may provide a less, rather than more, confident estimate of relationships. To this end, several plausible explanations may be offered: (1) the addition of a random signal, devoid of phylogenetic content, from the morphology/allozyme data sets (MA) might have diffused a strong signal from mtDNA; (2) the phylogenetic signal from MA may conflict with that of mtDNA due to systematic bias stemming from nonindependence within data sets; (3) a weak signal from MA may agree with that of mtDNA yet yield an incorrect topology in combination with it, due to the pooling of heterogeneously evolving classes of data (Bull et al. 1993). (See the next section for another hypothesis.)

We addressed the first possibility by conducting exhaustive searches for trees that were consistent with a constrained topology holding constant all cladistic relationships from the combined-approach analysis except those of the communa subclade, which were allowed to vary. These provided g_1 values, measures of skewness (asymmetry) of tree length distribution, that we compared to critical values provided by Hillis and Huelsenbeck (1992). As significant skewness (significantly negative g_1 values) suggests the presence of phylogenetic signal, this allowed us to ask whether individual data sets were significantly more structured than random data with respect to communa subclade relationships. We found significant structure for analyses of each of the 16S, COI, mtDNA, morphology, allozyme, and MA data sets at $P < 0.01$, suggesting that the MA data are not simply contributing noise.

This finding led us to examine whether conflicts between the signals of these data sets might explain the combined-approach analyses. To do so, we calculated the character incongruence indices (I_m) of Miyamoto (as communicated in Kluge 1989). Given two data sets and a most parsimonious tree from each, we calculated I_m by summing the number of extra steps (above that of a minimum-length tree) required to fit each data set to its own tree (giving i_w), summing the number of extra steps required to fit each data set to the tree of the other data set (giving i_t), and obtaining I_m as $(i_t - i_w)/i_t$: the proportion of incongruence due to conflict between data sets. In cases where a given data set yielded multiple equally shortest trees, we calculated extra steps using the

best-fitting topology for the data set being used, as recommended by Swofford (1991).

Bearing in mind the limitations of these indices (Swofford 1991), we calculated I_m for several data set comparisons, taking as a subjective metric Kluge's (1989, p. 16) assessment of 11.4% as a "small" proportion of incongruence between data sets. On this standard, our results suggest relatively little conflict among *Ophraella* data sets: I_m = 8.0% for 16S and COI, 10.6% for morphology and allozymes, and 8.8% for mtDNA and morphology. In a very conservative approach, we also calculated I_m by fitting data to truncated topologies including only slobodkini clade species, thus measuring character conflict with respect to only those relationships known to be incongruent among data sets a priori. These analyses revealed higher yet still moderate levels of incongruence: I_m = 28.6% for 16S and COI, 23.1% for morphology and allozymes, and 27.5% for mtDNA and morphology. In each analysis, however, the comparison of mtDNA and allozyme data sets yielded a much higher estimate of I_m: 21.3% and 50.0%, respectively.

Lacking a theoretical framework for interpreting incongruence indices, we cannot be certain that *Ophraella* data sets do not conflict. True conflict would indicate bias internal to one or more of the data sets and offer no means of choosing among estimates of communa subclade relationships. Even under very conservative assumptions, however, the values obtained for I_m provide little compelling evidence that data sets strongly disagree in their signals, a possibly important exception being the comparatively great conflict between mtDNA and allozyme data sets. Interestingly, these latter I_m values decreased to 15.4% and 38.2% when the mtDNA data were fitted to a UPGMA tree based on Rogers distances from allozyme frequencies (fig. 1 in Futuyma and McCafferty 1990) rather than to MP trees. As I_m values depend on the tree used to fit the data, these results may illustrate the difficulties inherent in phylogenetically analyzing electromorphic data (Buth 1984; Avise 1989, p. 1194) rather than the presence of systematic bias within the allozyme data set.

Together, these results imply that each data set contains some phylogenetic signal for *communa* subclade relationships and that, by and large, these signals do not strongly disagree with one another. In such a case, one might reasonably predict that an analysis using the combined approach would draw out the (sometimes weak) signals from each data set and combine them into a single strong signal, reflected in strong bootstrap support for previously ambiguous relationships, as was the case with *O. notulata* and *O. slobodkini.* However, calculation of skewness for trees generated from the combined data set yielded a positive $g1$ value (+0.22), implying a thorough lack of phylogenetic signal for communa subclade relationships.

Collectively, then, our results suggest a scenario consistent with the third explanation profferred above. Morphology and allozyme data sets contain phylogenetic signals that do not strongly conflict with the strong signal of mtDNA, but these signals are too weak to strongly support communa subclade groupings (fig. 1). The pooling of these noisy data with mtDNA may, then, have obscured the strong signal of the latter, yielding a less confident estimate of communa subclade phylogeny (fig. 4) than that offered by mtDNA alone (fig. 3*a*) (see fig. 3 in Bull et al. 1993). These analyses illustrate the value of analyzing data sets both separately and in combination and document the potential risks of combining heterogeneously evolving data sets (Bull et al. 1993).

A New Phylogeny for *Ophraella*

If we accept the plausibility of the scenario developed above, in which the signals from morphology and allozyme data sets are too weak to provide confident estimates of communa subclade topology, while the combined data set provides no signal at all, then the highly supported mtDNA topology offers the best current estimate of these relationships. We adopt, then, communa subclade relationships from mtDNA for our revised phylogeny of *Ophraella.* Given the complete agreement among data sets for non–slobodkini clade relationships and the confidence with which the combined approach places *O. notulata* and *O. slobodkini,* we adopt the combined-approach topology for all other relationships. Together, these elements form a topology which differs from the mtDNA tree (fig. 3*a*) only in that it places *O. notulata* and *O. slobodkini* as sister species, a result that appeared in several mtDNA analyses as well (fig. 3*c;* table 3) and that is also supported by several highly distinctive morphological features that appear to be uniquely derived (Futuyma 1991).

Intraspecific Relationships and Lineage Sorting

Another possible explanation for the topological incongruence among data sets is that incomplete sorting of nonrecombining mitochondrial lineages may have yielded an accurate mitochondrial gene tree that fails to depict the actual sequence of cladogenetic events (Avise and Ball 1990), a problem that arises when ancestrally polymorphic mitochondrial haplotypes are maintained for periods of time greater than that separating consecutive speciation events. In such an instance, an mtDNA gene tree would be incongruent with accurate trees generated from other data sets. Further, gene tree groupings that inaccurately estimate species relationships might be strongly supported by bootstrap, while accurate group-

ings provided by other data sets are poorly supported due to a weak signal. This scenario is consistent with the results obtained from *Ophraella* mtDNA and morphology/allozyme analyses. Its plausibility is further supported by the implied paraphyly of *O. communa* and *O. artemisiae* (fig. 5), examples of the maintenance of ancestral polymorphism for 1.7–2.5 Myr (see Funk et al. 1995).

Further review of the COI analyses (fig. 5), however, provides little evidence that incomplete sorting of ancestral polymorphisms offers a general account of the incongruence among data sets. Rather, mitochondrial diversity may be rapidly purged from *Ophraella* populations, yielding a high phylogenetic fidelity of haplotypes to their taxa: of 10 cases in which more than one individual was sequenced from a given population, the haplotypes from those populations formed a monophyletic group nine times (fig. 5); of four cases in which individuals from multiple populations were sequenced from a single species, those populations formed a monophyletic group three times (fig. 5). In no instances were haplotypes from a population or species polyphyletically distributed on the tree. Both exceptions to monophyly, rather, are cases of paraphyly in which a single lineage descends from within the paraphyletic taxon. And in both cases the biogeographic and ecological relationships of the species involved in the implied paraphyly are consistent with a model of peripatric speciation, providing a possible evolutionary explanation for these deviations from monophyly (Funk et al. 1995).

The phylogenetic coherence observed within species and even populations of *Ophraella* might be expected given the rapid loss of mitochondrial diversity associated with a historical demography that was likely affected by Pleistocene bottlenecks (see, e.g., Avise et al. 1988) and patchily distributed host plants. Indeed, considerable population structure has been discovered in phytophagous beetle species, even on a microgeographic scale (McCauley et al. 1988; McCauley 1991; Rank 1992). On the other hand, for lineage sorting to explain the wholesale disagreement between mtDNA and morphology/allozyme topologies for the 2.8- to 4.4-Myr-old communa subclade (Funk et al. 1995), *Ophraella* effective population sizes would have had to have consistently numbered on the order of 10^6 (assuming that an average of $4N_e$ generations must pass before reciprocal monophyly of initially polymorphic sister taxa is attained [Neigel and Avise 1986]).

Finally, hybridization among *Ophraella* species may have distorted estimated relationships. Counting against this possibility, however, is the strongly supported monophyly of haplotypes from both *O. notulata* and *O. slobodkini*. These morphological sibling species seem to have maintained their genetic integrity despite the seeming opportunity for gene flow due to their overlapping geographic distributions and related host plants (see table 1 in Funk et al. 1995). Similarly, *O. bilineata* haplotypes are monophyletic with respect to *O. arctica* despite the parapatric distribution and similar host plants of these species, and the haplotypes of *O. communa* avoid polyphyly although the transcontinental distribution of this species overlaps with that of all *Ophraella* species except *O. arctica*.

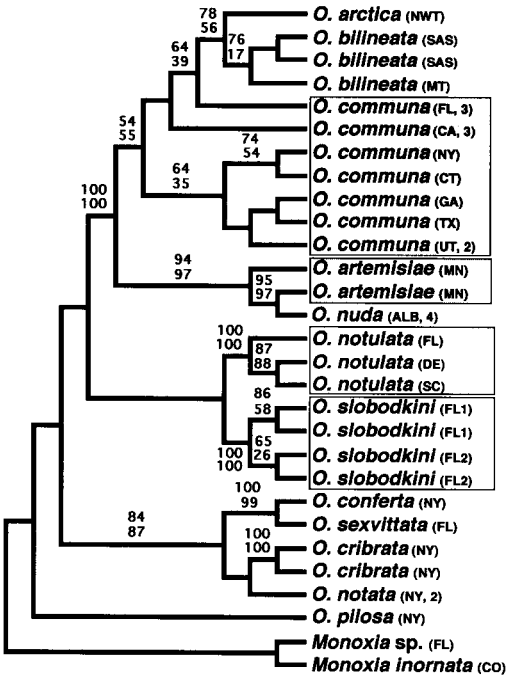

FIG. 5.—Majority rule consensus of 100 bootstrap replicates from the analysis of all COI haplotypes under equal weights using the two *Monoxia* species as an outgroup. This topology is compatible with the strict consensus of 24 equally shortest MP trees, the latter differing only in its lack of resolution among the GA, TX, UT, and NY/CT populations of *Ophraella communa*. Bootstrap proportions are provided from the MP (lower values) and NJ (upper values) analyses when one of these is greater than 50%. Following each OTU, the state or provincial abbreviation for its collection locality is presented, as is the number of specimens sequenced which share this haplotype. The seeming paraphyly of *O. communa* and *O. artemisiae* and the monophyly of *O. notulata* and *O. slobodkini* are highlighted.

Conclusion

Phylogenetic analysis of mtDNA haplotypes provides a robust estimate of *Ophraella* phylogeny and supports the genealogical coherence of *Ophraella* species lineages. Confidence in the inferred species-level topology is enhanced by substantial congruence among morphological, electromorphic, and mtDNA phylogenetic estimates and by the confident placement of two cladistically unstable species by the combined approach.

These data sets, however, offered incongruent estimates of "slobodkini clade" relationships, which only mtDNA strongly supported by bootstrap. The combined approach also provided low support for these relationships. The observation of significantly structured phylogenetic signal in the separate data sets, of relatively little conflict among them, and of lack of signal in the combined data suggest that combining even phylogenetically consistent data sets may diminish phylogenetic signal. The combination of heterogeneously evolving data sets may pose a challenge to the general applicability of the combined approach.

Sequence Availability

The nucleotide sequences reported here have been deposited in GenBank and may be recovered under accession numbers U20678–U20721.

Acknowledgments

We thank Alan de Queiroz, Paul Wilson, Leo Shapiro, and an anonymous reviewer for thoughtful comments on earlier drafts of the manuscript. Support by the National Science Foundation to D. Futuyma (BSR-8817912) is gratefully acknowledged. This is contribution 931 from Ecology and Evolution at the State University of New York at Stony Brook.

APPENDIX

Locality Data

Collection localities for specimens used in this study are given here. In parentheses are the number of sequenced specimens from a particular locality which share a COI haplotype.

Ophraella arctica: Canada, NWT, Inuvik (1). *O. artemisiae:* MN, Anoka Co., Bethel (2). *O. bilineata:* Canada, Sask., Chaplin (1,1); MT, Cascade Co., Cascade (1). *O. communa:* CA, San Diego Co., Kitchen Creek (1); CA, Inyo Co., Antelope Spring (2, same haplotype as Kitchen Creek); CT, Fairfield Co, Reading (1); FL, Leon Co., Iamonia (3); GA, Tift Co., Tifton (1); NY, Suffolk Co., Stony Brook (1); TX, Reeves Co., Balmorhea (1); UT, Uintah Co., Vernal (2). *O. conferta:* NY, Tompkins Co., Ithaca (1). *O. cribrata:* NY, Suffolk Co., Manorville (1,1). *O. notata:* NY, Tompkins Co., Ithaca (2). *O. notulata:* DE, Kent Co., Bombay Hook NWR (1); FL, Brevard Co., Merritt Is. (1); SC, Beaufort Co., Bluffton (1). *O. nuda:* Canada, Alta., Pakowki L. (2). *O. pilosa:* NY, Tompkins Co., Ithaca (1). *O. sexvittata:* FL, Dixie Co., Jena (1). *O. slobodkini:* FL, Leon Co., Iamonia (2) and Tylor Co., Steinhatchee (2). *Monoxia* sp.: FL, Wakulla Co., St. Mark's NWR (1). *Monoxia inornata:* CO, San Miguel Co., Placerville (1). *Exema neglecta:* FL, Tampa (1).

LITERATURE CITED

AVISE, J. C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. Evolution **43**: 1192–1208.

AVISE, J. C., and R. M. BALL, JR. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxford Surv. Evol. Biol. **7**:45–67.

AVISE, J. C., R. M. BALL, and J. ARNOLD. 1988. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. Mol. Biol. Evol. **5**:331–344.

BECKENBACH, A. T., Y. W. WEI, and H. LIU. 1993. Relationships in the *Drosophila obscura* group, inferred from mitochondrial cytochrome oxidase II sequences. Mol. Biol. Evol. **10**:619–634.

BROWER, A. V. Z. 1994. Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). Mol. Phylogenet. Evol. **3**:159–174.

BROWN, J. M., O. PELLMYR, J. N. THOMPSON, and R. G. HARRISON. 1994. Phylogeny of *Greya* (Lepidoptera: Prodoxidae) based on nucleotide sequence variation in mitochondrial cytochrome oxidase I and II: congruence with morphological data. Mol. Biol. Evol. **11**:128–141.

BROWN, W. M., E. M. PRAGER, A. WANG, and A. C. WILSON. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18**:225–239.

BULL, J. J., J. P. Huelsenbeck, C. W. CUNNINGHAM, D. L. SWOFFORD, and P. J. WADDELL. 1993. Partitioning and combining data in phylogenetic analyses. Syst. Biol. **42**:384–397.

BUTH, D. G. 1984. The application of electrophoretic data in systematic studies. Ann. Rev. Ecol. Syst. **15**:501–522.

CABOT, E. L., and A. T. BECKENBACH. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comp. Appl. Biol. Sci. **5**:233–234.

CAMERON, S. A., J. N. DERR, A. D. AUSTIN, J. B. WOOLEY, and R. A. WHARTON. 1992. The application of nucleotide sequence data to phylogeny of the Hymenoptera: a review. J. Hymenopt. Res. **1**:63–79.

CLARY, D. O., and D. R. WOLSTENHOLME. 1985. The mitochondrial molecule of Drosophila yakuba: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. **22**:252–271.

CROZIER, R. H., and Y. C. CROZIER. 1992. The cytochrome b and ATPase genes of honeybee mitochondrial DNA. Mol. Biol. Evol. **9**:474–482.

CROZIER, R. H., Y. C. CROZIER, and A. G. MACKINLEY. 1989. The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates. Mol. Biol. Evol. **6**:399–411.

DE QUEIROZ, A. 1993. For consensus (sometimes). Syst. Biol. **42**:368–372.

DESALLE, R., T. FREEDMAN, E. M. PRAGER, and A. C. WILSON. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. **26**:157–164.

EDWARDS, S. V., P. ARCTANDER, and A. C. WILSON. 1991. Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. Proc. R. Soc. Lond. [Biol.] **243**:99–107.

FANG, Q., W. C. BLACK, H. D. BLOCKER, and R. F. WHITCOMB. 1993. A phylogeny of New World *Deltocephalus-*

like leafhoppers based on mitochondrial ribosomal sequences. Mol. Phylogenet. Evol. **2**:119–131.

FARRIS, J. S. 1969. A successive approximations approach to character weighting. Syst. Zool. **18**:374–385.

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. **28**: 49–62.

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

———. 1988. Phylogenies from molecular sequences: inferences and reliability. Annu. Rev. Genet. **22**:521–565.

———. 1991. PHYLIP—phylogeny inference package (version 3.4). University of Washington, Seattle.

FUNK, D. J., D. J. FUTUYMA, G. ORTÍ, and A. MEYER. 1995. A history of host associations and evolutionary diversification for *Ophraella* (Coleoptera: Chrysomelidae): new evidence from mitochondrial DNA. Evolution (in press).

FUTUYMA, D. J. 1990. Observations on the taxonomy and natural history of *Ophraella* Wilcox (Coleoptera: Chrysomelidae), with a description of a new species. J. N.Y. Entomol. Soc. **98**:163–186.

———. 1991. A new species of *Ophraella* Wilcox (Coleoptera: Chrysomelidae) from the southeastern United States. J. N.Y. Entomol. Soc. **99**:643–653.

FUTUYMA, D. J., M. C. KEESE, and D. J. FUNK. 1995. Genetic constraints on macroevolution: the evolution of host affiliation in the leaf beetle genus *Ophraella*. Evolution (in press).

FUTUYMA, D. J., M. C. KEESE, and S. J. SCHEFFER. 1993. Genetic constraints and the phylogeny of insect-plant associations: responses of *Ophraella communa* (Coleoptera: Chrysomelidae) to host plants of its congeners. Evolution **47**:888–905.

FUTUYMA, D. J., and S. MCCAFFERTY. 1990. Phylogeny and the evolution of host plant associations in the leaf beetle genus *Ophraella* (Coleoptera: Chrysomelidae). Evolution **44**:1885–1913.

FUTUYMA, D. J., J. WALSH, T. MORTON, D. J. FUNK, and M. C. KEESE. 1994. Genetic variation in a phylogenetic context: responses of two specialized leaf beetles (Coleoptera: Chrysomelidae) to host plants of their congeners. J. Evol. Biol. **7**:127–146.

GUTELL, R. R., and G. E. FOX. 1988. A compilation of large subunit RNA sequences presented in a structural format. Nucleic Acids Res. **16**:r175–r269.

GYLLENSTEN, U. B., and H. A. ERLICH. 1988. Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc. Natl. Acad. Sci. USA **85**:7652–7656.

HELM-BYCHOWSKI, K., and J. CRACRAFT. 1993. Recovering phylogenetic signal from DNA sequences: relationships within the corvine assemblage (Class Aves) as inferred from complete sequences of the mitochondrial DNA cytochrome-b gene. Mol. Biol. Evol. **10**:1196–1214.

HIGGINS, D. G., and P. M. SHARP. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. Comp. Appl. Biol. Sci. **5**:151–153.

HILLIS, D. M., and J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. **42**:182–192.

HILLIS, D. M., and J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. **83**:189–195.

KIM, J. 1993. Improving the accuracy of phylogenetic estimation by combining different methods. Syst. Biol. **42**:331–340.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. **38**:7–25.

KLUGE, A. G., and A. J. WOLF. 1993. Cladistics: what's in a word? Cladistics **9**:193–200.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetic analysis. Pennsylvania State University, University Village.

LESAGE, L. 1986. A taxonomic monograph of the Nearctic galerucine genus *Ophraella* Wilcox (Coleoptera: Chrysomelidae). Mem. Entomol. Soc. Canada No. **133**:1–75. Ottawa, Ontario.

LI, W.-H., C.-I. WU, and C.-C. LUO. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J. Mol. Evol. **21**:58–71.

LIU, H., and A. T. BECKENBACH. 1992. Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of insects. Mol. Phylogenet. Evol. **1**:41–52.

MCCAULEY, D. E. 1991. The effect of host plant patch size variation on the population structure of a specialist herbivore insect, *Tetraopes tetraophthalmus*. Evolution **45**:1675–1684.

MCCAULEY, D. E., M. J. WADE, F. J. BREDEN, and M. WOHLTMAN. 1988. Spatial and temporal variation in group relatedness: evidence from the imported willow leaf beetle. Evolution **42**:184–192.

MEYER, A. 1994. DNA technology and phylogeny of fish: molecular phylogenetic studies of fish. Pp. 219–249 *in* A. R. BEAUMONT, ed. Genetics and evolution of aquatic organisms. Chapman & Hall, London.

NEIGEL, J. E., and J. C. AVISE. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515–534 *in* E. NEVO and S. KARLIN, eds. Evolutionary processes and theory. Academic Press, New York.

PALUMBI, S. R., A. MARTIN, S. ROMANO, W. O. MCMILLAN, L. STICE, and G. GRABOWSKI. 1991. Simple fool's guide to PCR. University of Hawaii, Honolulu.

PASHLEY, D. P., and L. D. KE. 1992. Sequence evolution in mitochondrial ribosomal and ND-1 genes in Lepidoptera: implications for phylogenetic analysis. Mol. Biol. Evol. **9**: 1061–1075.

PASHLEY, D. P., B. A. MCPHERON, and E. A. ZIMMER. 1993. Systematics of holometabolous insect orders based on 18S ribosomal RNA. Mol. Phylogenet. Evol. **2**:132–142.

RANK, N.E. 1992. A hierarchical analysis of genetic differentiation in a montane leaf beetle *Chrysomela aeneicollis* (Coleoptera: Chrysomelidae). Evolution **46**:1097–1111.

SAIKI, R. K., D. H. GELFLAND, S. STOFFEL, S. SCHARF, R. HIGUCHI, R. HORN, K., B. MULLIS, and H. A. ERLICH. 1989. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.

SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

SANGER, F., S. NICKLEN, and A. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

SIMON, C. 1991. Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing of amplified DNA. Pp. 1–33 *in* G. M. HEWITT, A. W. M. JOHNSTON, and J. P. W. YOUNG, eds. Molecular techniques in taxonomy. NATO Advanced Studies Institute, Springer, Berlin.

SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU, and P. FLOOK. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved PCR primers. Ann. Entomol. Soc. Am. **87**:1–51.

STURMBAUER, C., and A. MEYER. 1992. Genetic divergence, speciation, and morphological stasis in a lineage of African cichlid fishes. Nature **358**:578–581.

SWOFFORD, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? Pp. 295–333 *in* M. M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, Oxford.

———. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.

SWOFFORD, D. L., and S. H. BERLOCHER. 1987. Inferring evolutionary trees from gene frequency data under the principle of maximum parsimony. Syst. Zool. **36**:293–325.

SWOFFORD, D. L., and G. J. OLSEN. 1990. Phylogeny Reconstruction. Pp. 411–501 *in* D. M. HILLIS and C. MORITZ, eds. Molecular systematics. Sinauer, Sunderland, Mass.

TAMURA, K. 1992. The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. Mol. Biol. Evol. **9**:814–825.

VAWTER, L., and W. M. BROWN. 1993. Rates and patterns of base change in the small subunit ribosomal RNA gene. Genetics **134**:597–608.

WHEELER, W. C. 1991. Congruence among data sets: a Bayesian approach. Pp. 334–346 *in* M. M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, Oxford.

WILCOX, J. A. 1965. A synopsis of North American Galerucinae (Coleoptera: Chrysomelidae). Bull. N.Y. State Mus. Sci. Serv. **400**:1–226.

WILLIAMS, P. L., and W. M. FITCH. 1989. Finding the minimal change in a given tree. Pp. 453–470 *in* B. FERNHOLM, K. BREMER, and H. JORNVALL, eds. The hierarchy of life. Elsevier, Amsterdam.

JAN KLEIN, reviewing editor

Received October 17, 1994

Accepted February 1, 1995