11-1-2002

Molecular Phylogenetics of Mole Crabs (Hippidae : Emerita)

P. A. Haye

Y. K. Tam

Irv Kornfield
University of Maine - Main, irvk@maine.edu

JOURNAL OF CRUSTACEAN BIOLOGY, 22(4): 903–915, 2002

MOLECULAR PHYLOGENETICS OF MOLE CRABS (HIPPIDAE: *EMERITA*)

Pilar A. Haye, Yan K. Tam, and Irv Kornfield

(PAH) Department of Biological Sciences, University of Maine, Orono, Maine 04469, U.S.A. (pilar@umit.maine.edu); (YKT) GSPH-Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, U.S.A. (yktam@aecom.yu.edu); (IK, correspondence) School of Marine Sciences, University of Maine, Orono, Maine 04469, U.S.A. (irvk@maine.edu)

A B S T R A C T

Mole crabs of the genus *Emerita* (Family Hippidae) inhabit many of the temperate and tropical sandy beaches of the world. The nine described species of this genus are rarely sympatric, and most are endemic to broad biogeographic regions. The phylogenetic relationships among the species have not yet been investigated. Based on presumed morphological synapomorphies, it has been suggested that the species inhabiting the New World constitute a monophyletic group, as do the species inhabiting the Old World. The relationships within the New World species were previously studied using sequence data from Cytochrome Oxidase I and 16S rRNA mitochondrial genes; the results strongly suggested that one of the species, *Emerita analoga*, was very divergent from the other taxa examined. This observation prompted uncertainty about monophyly of the New World species. The goal of the present study was to elucidate the relationships among the species within the genus *Emerita*. Partial sequences for the mitochondrial COI and 16S rRNA genes for all nine species of the genus (and several outgroups) were examined. Phylogenetic analyses suggest that *E. analoga* is closer to the Old World taxa than to the other New World species; thus the New World *Emerita* species do not constitute a monophyletic group.

Mole crabs of the genus *Emerita* Scopoli, 1777, are medium-sized benthic crustaceans of the family Hippidae (Anomura: Hippoidea) that live in intertidal and upper subtidal sandy marine environments. Three genera are included in the family: *Emerita* Scopoli, 1777; *Hippa* Fabricius, 1787; and *Mastigochirus* Miers, 1878. Historically, taxonomy of the genera *Hippa* and *Emerita* has been dynamic, with movement of species from one genus to the other. When initially described, the genus *Emerita* had as type species a taxon that is now assigned to the genus *Hippa*. Similarly, in Miers' (1879) revision of the Hippidae, all of the species referred to as belonging to the genus *Hippa* now belong to the genus *Emerita*. Morphologically, the major difference between *Hippa* and *Emerita* is the shape of the carapace: it is dorsoventrally flattened, oval, and moderately convex in *Hippa* species and not flattened, cylindrical, and very convex in *Emerita* species. In addition, compared to *Hippa* species, *Emerita* species have longer ocular peduncles, longer flagella on the antennae, and smoother lateral margins of the carapace (Calado, 1987). The external morphology of these structures suggests that *Hippa* and *Emerita* are sister genera.

There are no described fossils that provide information regarding the history of *Emerita* species or the origin of the genus. On the basis of a molecular clock, it has been suggested that all species in the genus evolved before the mid- to late-Pliocene (Tam *et al.*, 1996), but no center of origin or biogeographic scenarios were suggested. *Emerita* species are distributed along most marine temperate coasts (Fig. 1) but are absent from East Atlantic and West Pacific coasts. In the regions where *Emerita* species are absent, mole crabs of the genera *Hippa* and *Mastigochirus* may be present; these genera are distributed throughout the temperate and tropical seas of the world (Miers, 1879). Six species of *Emerita* occur along New World coasts: *Emerita analoga* (Stimpson, 1857); *E. rathbunae* Schmitt, 1935; *E. talpoida* (Say, 1817); *E. benedicti* Schmitt, 1935; *E. brasiliensis* Schmitt, 1935; and *E. portoricensis* Schmitt, 1935. Three species are found along Old World coasts: *E. austroafricana* Schmitt, 1937; *E. emeritus* (Linnaeus, 1767); and *E. holthuisi* Sankoli, 1965.

Two conflicting phylogenetic hypotheses have been proposed (Fig. 2). On the basis of morphological traits and distribution of species, Efford (1976) suggested that *Emerita analoga*,

Fig. 1. Geographic distribution of the nine species of the genus *Emerita* (based on Efford (1976) and Tam *et al*. (1996)).

E. brasiliensis, and *E. talpoida* probably last occurred together in the late Pliocene. After the rise of the Isthmus of Panamá, their presumed mutual ancestor separated into different lineages during ocean warming and/or because of competition within the *E. rathbunae-E. portoricensis* species complex. *Emerita talpoida* and *E. brasiliensis* share morphological features that suggest a sister-group relationship between them; both are also similar to *E. analoga*. Although not explicitly considered in Efford's (1976) study, *E. benedicti* would also be part of the *E. rathbunae-E. portoricensis* complex (Fig. 2A). Additional evidence supporting Efford's hypothesis is based on the shape of the dactylus of the first peraeonite: the Old World species of *Emerita* share an acutely pointed dactylus on the first leg (Sankoli, 1962). Efford (1976) suggested that *E. austroafricana* and *E. emeritus* are closer to each other than either is to *E. holthuisi*.

The second hypothesis (Fig. 2B) is suggested by the study of Tam *et al*. (1996) on the divergence and biogeography of the New World mole crabs. Using sequence data from the mitochondrial (mt) genes Cytochrome Oxidase I (COI) and 16S ribosomal RNA (16S rRNA), they showed that *Emerita analoga* is divergent from the other five New World species (Fig. 2B). Their results suggest that *E. rathbunae* is more closely related to the species of *Emerita* found in the West Atlantic than it is to *E. analoga* that inhabits the same coastline. Tam *et al*. (1996) also suggested that the *Emerita* species in the Americas evolved from an ancestral stock that was split into two branches, one leading to *E. analoga* and the other to the five remaining species.

Schmitt (1937) did a comparative study of the joint of the antennal peduncle in mole crabs. He found that in *Emerita analoga*, the shape of the second joint is similar to that of *E. emeritus* and that the degree of ornamentation present in *E. analoga* lies between what can be seen in *E. emeritus* and *E. austroafricana*. The rest of the New World species have a second joint of the antennal peduncle that is similar in shape to the one in *E. analoga* or in *E. emeritus*, but its ornamentation is much different. Based on Schmitt's (1937) analysis, *E. analoga* is similar to the Old World taxa with respect to this character. This suggestion is in agreement with Tam *et al*. (1996) (even though they did not include the Old World taxa in their study): *E. analoga* is external to the rest of the New World taxa.

In contrast with Efford's hypothesis, Tam *et al*. (1996) also showed that *Emerita brasiliensis* does not appear to be close to *E. talpoida* and that *E. rathbunae* is more distantly related to American east coast species, suggesting that presumed morphological homologies may be homoplasious. The hypotheses presented by Efford (1976) and by Tam *et al*. (1996) are in conflict (Fig. 2); thus, phylogenetic resolution of all the species within the

genus will provide insights into questions of monophyly.

The present work uses mtDNA to investigate the level of genetic divergence within the genus *Emerita*. Mitochondrial DNA provides a powerful approach for resolving the relationships among taxa, especially when functional constraints influence morphological variation and/or when convergence confounds relationships. We used mtDNA data from the mt 16S rRNA and COI genes to study the levels of relatedness among species within the genus.

The 16S rRNA gene has been used to gain insights into the phylogenies of numerous crustacean taxa at both high and low taxonomic levels (i.e., orders and families as well as genera and species) (Cunningham *et al*., 1992; Crandall and Fitzpatrick, 1996; Sturmbauer *et al*., 1996; Casanova *et al*., 1998; Kitaura *et al*., 1998; Schubart *et al*., 1998, 2000; Tam and Kornfield, 1998; Crandall *et al*., 2000; Grandjean *et al*., 2000; Remigio and Hebert, 2000; Maggioni *et al*., 2001). Similarly, the COI gene has been successfully used to generate phylogenetic hypotheses at the species level in many crustacean groups (Palumbi and Benzie, 1991; Meyran *et al*., 1997; Badwin *et al*., 1998; Harrison and Crespi, 1999). For a wide diversity of taxa, the COI gene is the most conserved (in terms of sequence variation) of the mitochondrial protein-coding genes (Simon *et al*., 1994). The DNA sequences of 16S rRNA and COI provide two data sets that may be informative at different taxonomic levels because of their different rates of evolutionary change.

Materials and Methods

Collection and DNA Extraction

All nine species of the genus were examined in this study. Samples for New World species were characterized by Tam *et al*. (1996). Specimens of the Old World species were collected, fixed in 95% ethanol, and shipped to the University of Maine. Specimens were dissected, and the telson muscle was used for DNA extraction employing proteinase K digestion followed by phenol/chloroform extraction (Ausubel *et al*., 1989).

Partial sequences for COI and 16S rRNA genes were obtained through PCR (Polymerase Chain Reaction; Saiki *et al*., 1988). A 570-basepair (bp) segment of the 16S rRNA gene was amplified using these primer sequences: 16SA 5′CGCCTGTTTATCAAAAACAT and 16SB 5′CTCCGGTTTGAACTCAGATC (Xiong and Kocher, 1991), and a 658-bp segment of the COI gene was amplified using these primers: COIa 5′AGTATAAGCGTCTGGGTAGTC and COIf 5′CCTGCAGGAGGAGGAGAYCC (Palumbi and Benzie, 1991). Amplification reactions were performed following standard protocols described by Tam *et al*. (1996). Sequences were obtained from purified PCR products (QIAquick-spin columns; Qiagen, Inc., Chatsworth, California) followed by cycle-sequencing and electrophoresis on an Applied Biosystems, Inc., Model 377 Automated Sequencer (Foster City, California), using amplification primers and fluorescent dye terminators (Perkin-Elmer; Foster City, California). All sequences are a consensus of two sequencing reactions (one in each direction).

The DNA sequences of *Hippa pacifica* for both genes were obtained through cloning of PCR products that were purified (QIAquick-spin columns; Qiagen, Inc.) and ligated into the pCR® vector (TA Cloning™ System, Invitrogen Corp., San Diego, California), followed by transformation of INVα cells. Bacterial colonies with inserts were cultured in SOC media overnight, following which plasmids were

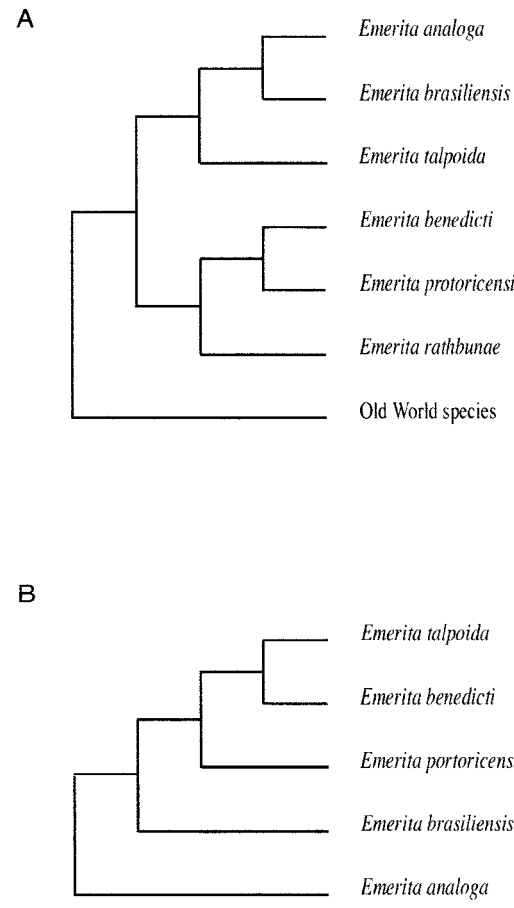

Fig. 2. Phylogenetic hypotheses for species of *Emerita*. The relationships suggested by the molecular analyses of Tam *et al*. (1996) are in conflict with the morphologically based hypothesis of Efford (1976). (A) Efford (1976) suggested that *Emerita analoga*, *E. brasiliensis*, and *E. talpoida* form a clade sister to one containing *E. benedicti*, *E. portoricensis*, and *E. rathbunae*. Together these were hypothesized to form a sister clade to the Old World species (*E. austroafricana*, *E. emeritus*, and *E. holthuisi*). (B) Tam *et al*. (1996) proposed a phylogenetic hypothesis for the New World species of *Emerita* based on mtDNA data. *Emerita analoga* is external to the other New World species of *Emerita*.

purified using the Plasmid Miniprep Kit (Qiagen, Inc.). Presence of the correct-size insert was verified using PCR followed by agarose gel electrophoresis. Plasmid preparations that had the correct-size inserts were sequenced as previously described.

The genus *Hippa* is considered to be the sister taxon to the genus *Emerita* (A. Harvey, personal communication). The Pacific mole crab *Hippa pacifica* (Dana, 1852), widespread in the Indopacific region, was investigated as the sister taxon in the present study. Several anomuran species belonging to the families Porcellanidae and Paguridae were also used as outgroup taxa. Three nonhippid anomuran taxa were available for use as outgroups for the 16s rRNA, and six for the COI data. Table 1 lists the species examined and the associated GenBank accession numbers.

Data Analysis

Sequences were aligned by eye using the program ESEE (Cabot and Beckenbach, 1989) and Sequence Navigator (Applied Biosystems, version 1.0, 1994, Foster City, California). After alignment, the sequences were truncated at the 5′ and 3′ ends to leave no ambiguous sites for any taxon.

Data sets for both genes were examined for base-frequency homogeneity by using a Chi-square test available in PAUP 4.0* version 4.0b8a (Swofford, 2000). Each data set was analyzed with distance by using neighbour-joining (Saitou and Nei, 1987), parsimony (Camin and Sokal, 1965), and maximum-likelihood (Felsenstein, 1981) optimality criteria by using PAUP* (Swofford *et al*., 1996; Steel and Penny, 2000; Swofford, 2000). All analyses were done with random-sequence addition. Gaps were treated as a fifth character for the parsimony analyses.

The program Modeltest version 3.06 (Posada and Crandall, 1998) was used to select the model of molecular evolution for maximum-likelihood analyses. Modeltest performs a hierarchical likelihood-ratio test to compare 56 models of molecular substitution and selects the model that best fits the data. The chosen model was then implemented in PAUP*. Each data set was individually tested using Modeltest (see Results).

The data were evaluated for information content using the skewness statistic (g_1) obtained from the tree-length frequency distribution of an exhaustive parsimony search using PAUP* (Huelsenbeck, 1991; Hillis and Huelsenbeck, 1992). The robustness of tree topologies (nodes of the tree) was evaluated by bootstrap resampling (1,000 replicates; Felsenstein, 1985; Zharkikh and Li, 1992; Hillis and Bull, 1993). We report all bootstrap values over 50%.

The trees obtained from the phylogenetic analyses were further explored using MacClade 4.0 (Maddison and Maddison, 2000). Within MacClade, positions of taxa were changed, and the effect on tree lengths and confidence values were examined. Consistency indices (CI) and retention indices (RI) were obtained from MacClade tree-topology analyses (Farris, 1989). Alternative phylogenetic hypotheses were statistically tested in PAUP* using the corrected nonparametric Shimodaira-Hasegawa test of tree topologies (Shimodaira and Hasegawa, 1999; Goldman *et al*., 2000). Log-likelihood scores for the maximum-likelihood tree were compared to the log-likelihood scores from tree topologies constrained to fit alternative phylogenetic hypotheses. The null hypothesis of the Shimodaira-Hasegawa test is that there is no difference between trees.

The ratio of transversions to transitions (Tv : Ti) was calculated for every pair of taxa for both COI and 16S rRNA genes. The Tv : Ti ratio was plotted *versus* the Jukes-Cantor genetic distance (Jukes and Cantor, 1969) to assess the degree of saturation of the sequences caused by multiple substitutions. Because transitions are much more frequent than transversions, the ratio in which they occur potentially provides information about the level of mutation-saturation of the gene. A Mantel test (Mantel, 1967) was used to test for the degree of association between elements of the distance and Tv : Ti-ratio matrices for each data set.

Gene sequences were analyzed using different weighting schemes for transitions and transversions (Pollock and Goldstein, 1994). Because COI is a protein-coding gene, weights per nucleotide position and the effects of the inclusion or exclusion of third positions within codons were examined. For this gene, phylogenetic analyses at the amino acid level were also performed.

Secondary structure of the partial sequence of the 16S rRNA gene was inferred by correspondence with the secondary structure model for *Drosophila yakuba* (see Gutell and Fox, 1988) and by use of mfold software (version 3.0, 1996, available at http://mfold.burnet.edu.au/), which predicts secondary structure of rRNA. Sites within the sequence were then classified as stems or loops (see also Harris and Mayden, 2001). Phylogenetic analyses were performed by weighting stems and loops both equally and differentially (1–5×) in order to evaluate the effects on topology (Kjer, 1995).

A partition homogeneity (= incongruence length difference) test implemented in PAUP* was performed in order to examine homogeneity among data sets (Farris *et al*., 1994). If data sets are not significantly heterogeneous, then combination of data sets may be appropriate for analysis via a total evidence approach (Kluge, 1989).

Results

Final truncated sequences used for the analyses were 322 bp and 419 bp for the genes 16S rRNA and COI, respectively. Both 16S rRNA and COI were successfully sequenced for seven of the nine *Emerita* species and for *Hippa pacifica*; unfortunately, no amplification products were obtained for *E. austroafricana* and *E. rathbunae* by using 16S primers or for *E. holthuisi* and *E. portoricensis* by using COI primers.

No significant base-frequency heterogeneity was found among taxa for either of the genes; base frequencies were not significantly different among taxa ($P = 0.976$ and $P = 0.995$ for 16S and COI data, respectively).

Modeltest (Posada and Crandall, 1998) revealed that the model of substitution that best fit the 16S rRNA data was the transversion model with a gamma distribution (base frequencies: A = 0.390, C = 0121, G = 0.063, T = 0.425; substitution model rate matrix: [A-C] = 2.028, [A-G] = 19.078, [A-T] = 3.000, [C-G] = 4.060, [C-T] = 19.078, [G-T] = 1.000; proportion of invariable sites = 0; gamma shape parameter = 0.206). The general time reversible model with a proportion of invariable sites

Table 1. List of taxa, collection localities for *Emerita* species, and GenBank accession numbers for all species.

Species	Collection locality	GenBank accession no. 16S	GenBank accession no. COI
Hippoidea: Hippidae			
Emerita analoga	San Diego, California and	AF246153	L43101
	Algarrobo, Chile	AF246154	L43099
E. austroafricana	Durhan, South Africa		AF246160
E. benedicti	Port Aransas, Texas, U.S.A.	AF246155	L43102
E. brasiliensis	Fortaleza de Santa Teresa, Uruguay	L43110	L43151
E. emeritus	Pondichvory, India	AF246156	AF246159
E. holthuisi	Dubai, United Arab Emirates	AF246157	
E. portoricensis	Mayaguez, Puerto Rico	L43111	
E. rathbunae	Golfo de Fonseca, El Salvador		L43103
E. talpoida	West Falmouth, Massachusetts;	AF246151	L43104
	Conway, South Carolina; and	AF246150	L43105
	Panacea, Florida	AF246152	L43106
Hippa pacifica	DNA provided by C. Cunningham	AF246158	AF246161
Galatheoidea: Porcellanidae			
Clatotoechus vanderhorsti			AF222723
Neopisosoma angustifrons			AF222720
Pachycheles serratus			AF296178
Pachycheles chilensis		AF260610	
Petrolisthes galathinus		AF260639	AF222727
Porcellana platycheles			AF222731
Paguroidea: Paguridae			
Pagurus longicarpus		AF150756	AF150756

and with a gamma distribution (GTR + I + G) fit the COI data (base frequencies: A = 0.366, C = 0.149, G = 0.102, T = 0.382; substitution model rate matrix: [A-C] = 65.780, [A-G] = 133.584, [A-T] = 0.000, [C-G] = 48.914, [C-T] = 2,673.434, [G-T] = 1.000; proportion of invariable sites = 0.498; gamma shape parameter = 0.335). For the combined data set, the same GTR + I + G model of substitution provided the best fit (base frequencies: A = 0.348, C = 0.145, G = 0.112, T = 0.395; substitution model rate matrix: [A-C] = 97, 015.945, [A-G] = 270,529.219, [A-T] = 96, 358.883, [C-G] = 57,551.094, [C-T] = 747, 799.375, [G-T] = 1.000; proportion of invariable sites = 0.3969; gamma shape parameter = 0.539).

The 16S rRNA data (102 parsimony-informative characters) generated similar tree topologies for distance, parsimony, and maximum-likelihood analyses (the latter shown in Fig. 3). Under parsimony, the shortest tree had 291 steps ($g_1 = -1.42$ ($P < 0.01$); CI = 0.73; HI = 0.27; RI = 0.60; RC = 0.43; −Ln likelihood = 1,627.14). Differentially weighting transitions and transversions (Pollock and Goldstein, 1994), or stems and loops (Kjer, 1995) did not alter the tree topology (data not shown). In the maximum-likelihood with bootstrap resampling analysis, *Hippa pacifica* and *Emerita analoga* cluster together with 87% bootstrap support (Fig. 3). The *E. analoga-H. pacifica* clade is sister to the *E. holthuisi-E. emeritus* clade (Old World clade) with 75% bootstrap support. All hippid species examined cluster together with 100% bootstrap support when using three nonhippid taxa as outgroups (*H. pacifica* always falls within the ingroup). The maximum-likelihood tree obtained in the PAUP* analysis for 16S rRNA was manipulated in MacClade to test the effects of forcing the topology to make the genus *Emerita* monophyletic (forcing *H. pacifica* to be in the outgroup). The resultant tree has six more steps and slightly lower CI and RI values (CI = 0.71; HI = 0.27; RI = 0.59; RC = 0.43; −Ln likelihood = 1,673.39). Changing the topology so it conforms to Efford's (1976) hypothesis with the New World *Emerita* monophyletic (Fig. 2A) generated a tree with 16 more steps and lower CI and RI values (CI = 0.69; HI = 0.31; RI = 0.52; RC = 0.36; −Ln likelihood = 1,631.36). Shimodaira-Hasegawa tests (Shimodaira and Hasegawa, 1999) were performed to compare alternative tree topologies generated from the 16S rRNA data. The tree where the genus *Emerita* is monophyletic, as in Ef-

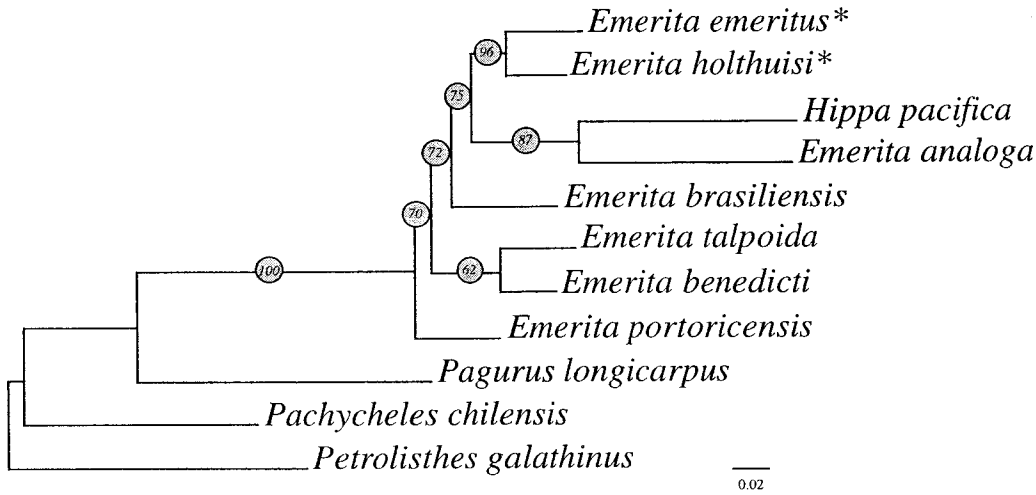

Fig. 3. Phylogram of the 16S rRNA sequence data and values from bootstrap maximum-likelihood analysis (1,000 replicates). Nonhippid anomurans were declared as outgroup taxa. *Hippa pacifica* and all the *Emerita* species cluster together with 100% bootstrap support. *Hippa pacifica* and *E. analoga* form a clade sister to the *E. holthuisi-E. emeritus* clade. Scale bar represents branch length (substitutions/site). (*) Species from the Old World.

ford's (1976) hypothesis (Fig. 2A), is significantly different ($P = 0.030$) from the most parsimonious tree generated with the 16S rRNA data (Fig. 3).

Similar phylogenetic analyses were repeated for the COI gene data set. Compared to 16S rRNA, the COI phylogenetic signal was less robust. Bootstrap likelihood analysis provides 71% support for the *Emerita emeritus-E. austroafricana* clade, and 66% support for the cluster of the Hippidae (Fig. 4). The COI nucleotide data represents a total of 132 parsimony-informative characters. The most parsimonious tree has 588 steps ($g_1 = -0.38$ ($P < 0.01$); CI = 0.47, RI = 0.33, and RC = 0.16) and is topologically consistent with the tree generated under maximum likelihood (−Ln likelihood = 2,512.77) (Fig. 4). Excluding fast-evolving third positions or weighting the codon sites differentially never produced a well-supported tree. Amino acid data only contained three parsimony-informative substitutions and thus was not used for further analyses. Forcing the tree topology to make the genus *Emerita* monophyletic resulted in a parsimony tree five steps longer (CI = 0.46, RI = 0.30, and RC = 0.14; −Ln likelihood = 3,002.10). When the topology was forced to conform to Efford's hypothesis by making New World *Emerita* species monophyletic (Fig. 2A), the tree was 11 steps longer (CI = 0.47, RI = 0.32, RC = 0.15; −Ln likelihood = 3,021.48). As with the 16S rRNA data, Shimodaira-Hasegawa tests (Shimodaira and Hasegawa, 1999) were performed to compare alternative tree topologies of the trees generated from the COI data (trees in Figs. 2A and 4). No tree comparisons were significant ($P > 0.1$).

Tamura-Nei distances (Tamura and Nei, 1993) based on 16S rRNA data indicate that *Emerita analoga* is the most divergent species in the genus (Table 2). *Emerita analoga* has an average distance of 0.165 from its congeners; the average distance among all *Emerita* species is 0.120 (excluding *E. analoga*, the distance decreases to 0.112). *Hippa pacifica* has an average distance of 0.166 from all *Emerita* species. The nonhippid species have an average distance of 0.352 from the *Emerita* species.

As with the 16S rRNA data, Tamura-Nei genetic distance data for the COI gene show that *Emerita analoga* is the most distant of the *Emerita* species, having an average distance of 0.215 from its congeners (Table 2). The average genetic distance among all the *Emerita* species is 0.189, and when excluding *E. analoga*, it is 0.185. *Hippa pacifica* has an average genetic distance of 0.207 (smaller than the average distance for *E. analoga*); the nonhippid outgroup species have an average COI genetic distance of 0.227 from the *Emerita* species.

Partition-homogeneity tests indicated that the 16S rRNA and COI data sets are not signifi-

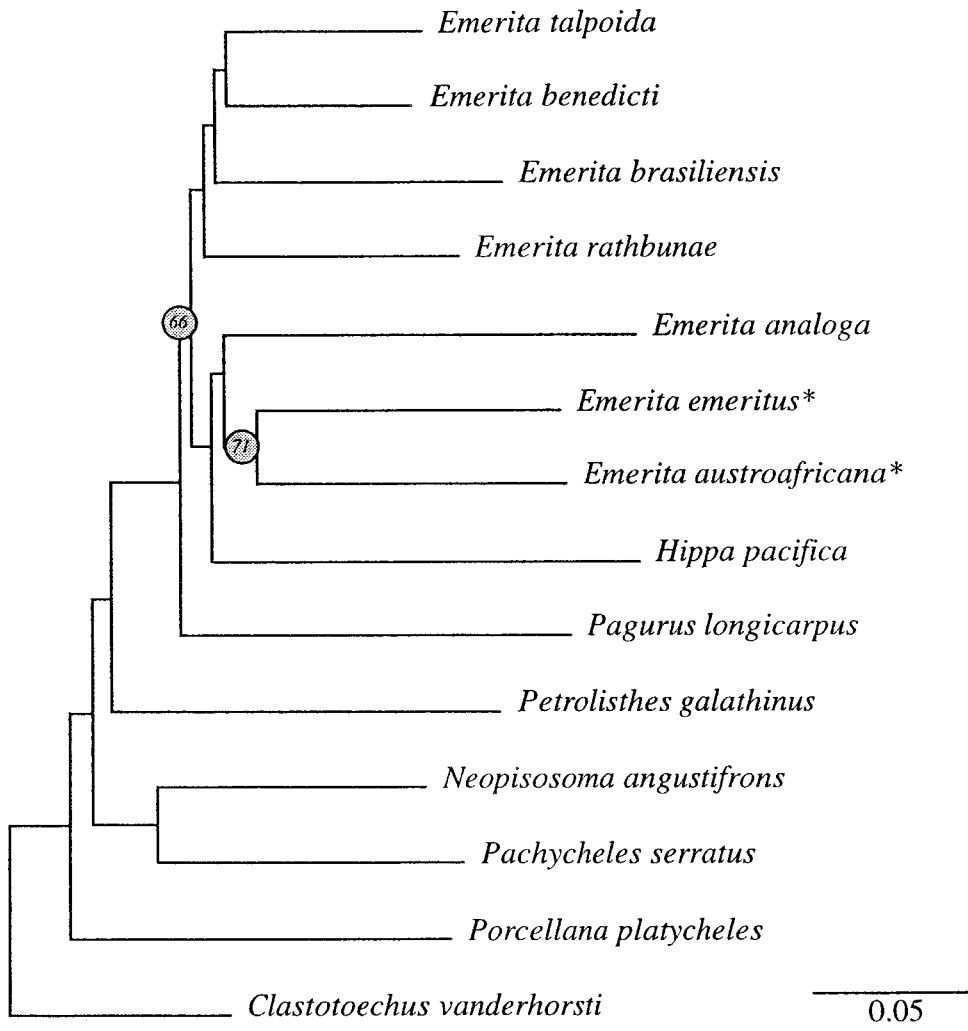

Fig. 4. Phylogram of the COI sequence data and values from bootstrap maximum-likelihood analysis (1,000 replicates). Nonhippid anomurans were declared as outgroup taxa. *Hippa pacifica* and *Emerita analoga* cluster with the Old World species (no bootstrap support over 50% for the formed clade). Scale bar represents branch length (substitutions/site). (*) Species from the Old World.

cantly heterogeneous ($P = 0.176$), which justifies performing phylogenetic analysis using a combined data set (Farris *et al.*, 1994). The combined data set yields results similar to those generated for 16S and COI separately (Fig. 5). *Emerita analoga* and *Hippa pacifica* cluster together with a 60% bootstrap support value. *Emerita analoga* and *H. pacifica* cluster with *E. emeritus* (the only representative of Old World taxa for which there was both 16S rRNA and COI data) with 83% bootstrap support. All *Emerita* species and *H. pacifica* cluster together with 100% bootstrap support.

Gene-saturation analyses provided independent support for utility of the 16S rRNA data but demonstrated potentially less utility for COI gene data. Plots of the transversion : transition ratio (Tv : Ti) *versus* Jukes-Cantor genetic distance (Jukes and Cantor, 1969) show that the genes have different rates of evolution in the genus *Emerita*. For the 16S rRNA gene, the slope of the regression is positive ($P < 0.01$), indicating that the gene is not completely saturated with mutations (Fig. 6A). By contrast, the slope of Tv : Ti ratio *versus* Jukes-Cantor distance for COI (Fig. 6B) is not significant (see Discussion).

Table 2. Genetic distances (Tamura and Nei, 1993) for 16S rRNA (above) and COI (below) sequences for taxa considered in the study.

	1	2	3	4	5	6	7	8	9	10	11
1 *Emerita analoga*	–										
2 *Emerita benedicti*	0.1512	–									
3 *Emerita brasiliensis*	0.1462	0.0826	–								
4 *Emerita emeritus*	0.1758	0.1222	0.1149	–							
5 *Emerita holthuisi*	0.1788	0.1303	0.1143	0.0586	–						
6 *Emerita portoricensis*	0.1662	0.0487	0.0903	0.1423	0.1168	–					
7 *Emerita talpoida*	0.1735	0.0524	0.1033	0.1355	0.1347	0.0827	–				
8 *Hippa pacifica*	0.1581	0.1589	0.1566	0.1674	0.1548	0.1732	0.1906	–			
9 *Pachycheles chilensis*	0.4118	0.3565	0.3527	0.4097	0.3805	0.3322	0.3808	0.3724	–		
10 *Pagurus longicarpus*	0.3624	0.3268	0.3004	0.3366	0.3167	0.2933	0.3278	0.3930	0.2835	–	
11 *Petrolisthes galathinus*	0.4416	0.3874	0.3949	0.4160	0.4094	0.3710	0.3889	0.4411	0.2211	0.3263	–

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 *Emerita analoga*	–													
2 *Emerita austroafricana*	0.2319	–												
3 *Emerita benedicti*	0.2249	0.1983	–											
4 *Emerita brasiliensis*	0.2094	0.2038	0.1463	–										
5 *Emerita emeritus*	0.2036	0.1657	0.1791	0.2115	–									
6 *Emerita rathbunae*	0.2145	0.1998	0.1464	0.1709	0.1997	–								
7 *Emerita talpoida*	0.2064	0.2093	0.1417	0.1445	0.1999	0.1623	–							
8 *Hippa pacifica*	0.2383	0.2239	0.1843	0.2476	0.1968	0.1974	0.1608	–						
9 *Clastotoechus vanderhorsti*	0.2387	0.2320	0.2030	0.2166	0.2660	0.2083	0.2124	0.2493	–					
10 *Neopisosoma angustifrons*	0.2303	0.2598	0.2254	0.2345	0.2399	0.2086	0.2124	0.2558	0.1813	–				
11 *Pachycheles serratus*	0.2618	0.2718	0.2199	0.2111	0.2591	0.2079	0.2115	0.2774	0.1982	0.1613	–			
12 *Pagurus longicarpus*	0.2742	0.2383	0.1612	0.1947	0.2400	0.2004	0.2214	0.2030	0.2343	0.2699	0.2504	–		
13 *Petrolisthes galathinus*	0.2478	0.2713	0.2242	0.2172	0.2722	0.2194	0.2418	0.2510	0.2170	0.2044	0.2152	0.1983	–	
14 *Porcellana platycheles*	0.2733	0.2464	0.2281	0.2366	0.2626	0.2244	0.2396	0.2845	0.1888	0.2124	0.2078	0.2343	0.2326	–

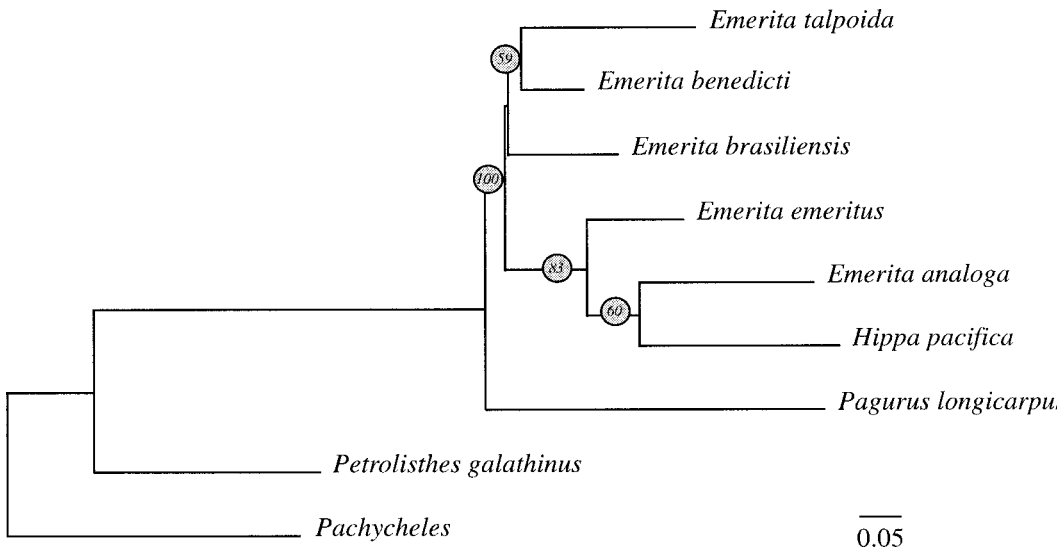

Fig. 5. Phylogram of the combined data set and values from bootstrap maximum likelihood analysis (1,000 replicates). *Pachycheles* (16S rRNA sequence corresponds to *Pachycheles serratus* and the COI sequence to *P. chilensis*), *Pagurus longicarpus*, and *Petrolisthes galathinus* were declared as outgroups. Scale bar represents branch length (substitutions/site). (*) Species from the Old World.

Discussion

That both separate and combined analyses placed *Emerita analoga* and *Hippa pacifica* within the same clade was unexpected. Parsimony and likelihood analyses of both 16S rRNA and COI data support the external placement of *E. analoga* with respect to the other New World members of the genus (Figs. 3–5). The results obtained from the 16S rRNA and COI genes are consistent with the hypothesis presented by Tam *et al.* (1996) (Fig. 2B) that *E. analoga* falls outside the group that includes all other New World *Emerita* species. The placement of *E. analoga* closer to the Old World than to the rest of the New World species follows Schmitt's (1937) observations that the second joint of the antennal peduncle in *E. analoga* is more similar in shape and ornamentation to the one seen in the Old World taxa.

The difference between the results generated with the 16S rRNA and the COI data is mainly the degree to which the trees are supported; this may be a result of saturation that COI exhibits for this group of species (Fig. 6B). The tree resulting from analyses of the 16S rRNA data is well supported and the data are not saturated; the topology generated thus provides a strong and significant phylogenetic hypothesis. This perspective is further corroborated by the fact that the data for the 16S rRNA gene generates a similar topology regardless of the optimality criteria used to generate the phylogeny.

Tamura-Nei distances of 16S rRNA show that the divergence between *Emerita analoga* and all other species of *Emerita* is greater than that of any one of the other *Emerita* species and its congeners (Table 2). The fact that *E. analoga* is highly divergent could be the result of a higher rate of mutation or a longer divergence time. Long branch attraction (i.e., an artifactual association of taxa that are distantly related because of either high substitution rates or very distant outgroups) (Felsenstein, 1978; Huelsenbeck, 1997) has been a frequent explanation for the grouping of some taxa on the basis of homoplasies. Whereas *E. analoga* exhibits large genetic distances relative to its congeners, those distances are not of a magnitude believed to generate long branch attraction (Table 2; Figs. 3–5).

Any particular DNA sequence is informative over a limited divergence range, and that range is variable among different taxa (Naylor and Brown, 1998). That the COI gene would be informative at the species level was predicted *a priori* (Palumbi and Benzie, 1991; Meyran *et al.*, 1997; Badwin *et al.*, 1998). However, DNA sequence data from that gene for *Emerita* was not sufficiently informative to clearly elucidate relationships within the genus. As shown in Fig. 6B, the regression of the ratio of transversions to transitions against

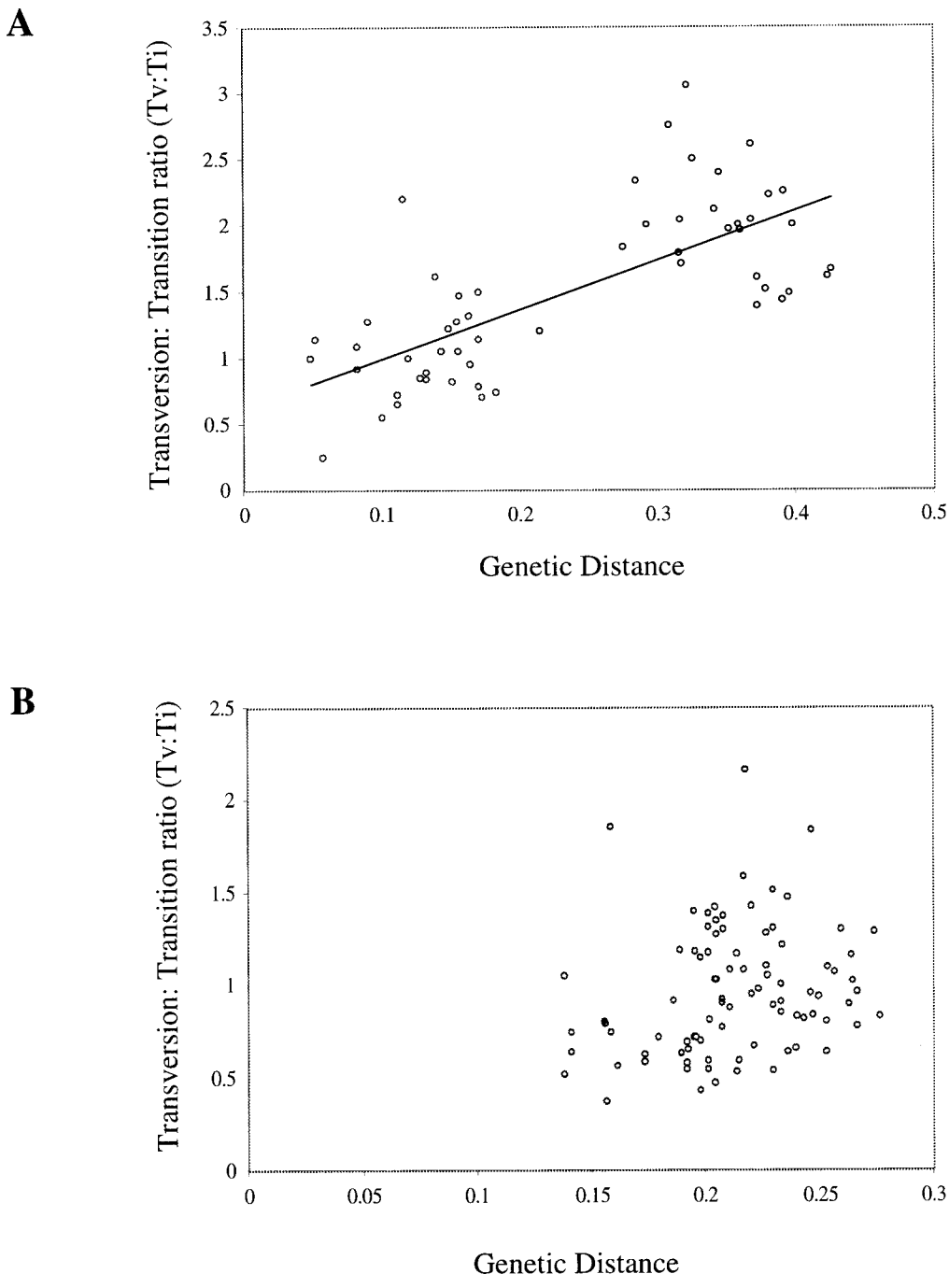

Fig. 6. Substitution patterns for all species considered in the analyses. Transversions vs. transitions (Tv : Ti) are plotted against Jukes-Cantor genetic distance (Jukes and Cantor, 1969) for every pair of taxa. (A) 16S rRNA data. A significant ($P < 0.01$) positive slope is observed ($y = 3.728x + 0.6193$; $r^2 = 0.4996$). (B) COI nucleotide data. Slope is not significant ($P > 0.1$).

Jukes-Cantor distance was not significant for the COI sequences. Two possible but opposite explanations for this finding may be suggested: the gene has accumulated few mutations, or the gene is saturated with mutations. The fact that the Tv : Ti ratio is not significantly different for both closely related and divergent pairs of taxa suggests that the gene is evolving very slowly and has not accumulated enough phylogenetic signal to elucidate relationships (Figs. 4, 6). This is also supported by the observation that there are only three phylogenetically informative amino acid substitutions for *Emerita* within the data set.

If the molecular data presented here provides a valid estimate of the true species tree, then *E. analoga* is closer to the Old World species than to the New World taxa. This result poses a biogeographical challenge for radiation of the remaining taxa. It has been hypothesized that *Emerita* species evolved at least before the Late Neogene (Tam *et al*., 1996). Both vicariance and dispersal undoubtedly played a role in the speciation of *Emerita*; other ecological, physiological, and oceanographic processes probably contributed to the final geographic placement of populations that gave rise to the different *Emerita* species. Range expansion and colonization of new geographic areas with subsequent reduced gene flow were probably the mechanisms by which the majority of these species originated; explicit biogeographic scenarios are difficult to suggest, but we outline two below. Consistent with our results, the genus *Emerita* could have originated in the western side of the Atlantic Ocean. If the center of origin is the Atlantic Ocean, we can hypothesize that species currently distributed in the Pacific (*E. analoga* and *E. rathbunae*) evolved from ancestors in the Atlantic via opening and closing of the Isthmus of Panamá. The Isthmus was closed to circulation of marine surface water approximately 3 million years ago (Late Neogene), and it was closed to deep-water circulation much earlier (Malfait and Dinkelman, 1972; Keigwin, 1982). The placement of *Hippa pacifica*, sister to *E. analoga*, is biogeographically consistent with this idea, in that both species occur in the Pacific. Alternatively, the center of origin of the genus *Emerita* may be the Pacific Ocean. The Atlantic may have been colonized via the Isthmus. Taxa that differentiated in the Atlantic may have been ancestral to species that subsequently recolonized the Pacific.

Morphology may provide misleading information when homoplasy is not taken into consideration. The genera *Emerita* and *Hippa* are differentiated by a few external features (i.e., shape of the carapace, ocular peduncle shape, and length of antennal flagellum) that are highly constrained by functional morphology (and thus, presumably, by the environment). Our data strongly suggest that *Hippa pacifica* and *Emerita analoga* are closely related. The molecular data pose a basic question for understanding relationships within the family Hippidae, but these findings alone are not sufficiently strong to justify taxonomic revision. Further analyses that include more species of the genus *Hippa* and representatives of *Mastigochirus* are necessary to elucidate phylogenetic positions within the family. From the present study, we conclude (1) that the *Emerita* species of the Americas do not constitute a monophyletic unit as previously believed, and (2) that morphological characters traditionally used in the systematics of mole crabs are homoplasious in the family Hippidae, especially within the genus *Emerita*.

Acknowledgements

We thank M. Apel, V. Arul, and B. Riege for kindly providing samples of Old World taxa for our analysis. C. Cunningham generously provided a sample of DNA from *Hippa pacifica*. A. Harvey provided useful information and discussion concerning morphology of mole crabs. IK thanks Maryhaven for encouragement. We also thank anonymous reviewers for their very insightful comments. This work was supported in part by NSF grants DEB97-07532 and OCE98-06712.

Literature Cited

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Siedman, J. A. Smith, and K. Struhl, eds. 1989. Short Protocols in Molecular Biology. John Wiley and Sons, New York.

Badwin, J. D., A. L. Bass, B. W. Bowen, and W. H. Clark. 1998. Molecular phylogeny and biogeography of the marine shrimp *Penaeus*.—Molecular Phylogenetics and Evolution 10: 399–407.

Cabot, E. L., and A. T. Beckenbach. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE.—Cabios 5: 233, 234.

Calado, T. C. 1987. Taxonomia, Biogeografia e Ecologia da Superfamília Hippoidea na costa brasileira (Crustacea, Decapoda).—Master's thesis, Depto. De Oceanografia, UFPE, Brazil.

Camin, J. H., and R. R. Sokal. 1965. A method for deducing branching sequences in phylogeny.—Evolution 19: 311–326.

Casanova, J.-P., L. De Jong, and E. Faure. 1998. Interrelationships of the two families constituting the Lophogastrida (Crustacea: Mysidacea) inferred from

morphological and molecular data.—Marine Biology 132: 59–65.

Crandall, K. A., J. W. Fetzner, Jr., C. G. Jara, and L. Buckup. 2000. On the phylogenetic positioning of the South American freshwater crayfish genera (Decapoda: Parastacidae).—Journal of Crustacean Biology 20: 530–540.

———, and J. F. Fitzpatrick, Jr. 1996. Crayfish molecular systematics: using a combination of procedures to estimate phylogeny.—Systematic Biology 45: 1–26.

Cunningham, C. W., N. W. Blackstone, and L. W. Buss. 1992. Evolution of king crabs from hermit crab ancestors.—Nature 355: 539–542.

Efford, I. E. 1976. Distribution of the sand crabs in the genus *Emerita* (Decapoda, Hippidae).—Crustaceana 30: 169–183.

Farris, J. S. 1989. The retention index and rescaled consistency index.—Cladistics 5: 417–419.

———, M. Kallersjo, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence.—Cladistics 10: 315–320.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading.—Systematic Zoology 27: 401–410.

———. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach.—Journal of Molecular Evolution 17: 368–376.

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

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood based tests of topologies in phylogenetics.—Systematic Biology 49: 652–670.

Grandjean, F., D. J. Harris, C. Souty-Grosset, and K. A. Crandall. 2000. Systematics of the European endangered crayfish species *Autropatamobius pallipes* (Decapoda: Astacidae).—Journal of Crustacean Biology 20: 522–529.

Gutell, R. R., and G. E. Fox. 1988. A compilation of large subunit RNA sequences presented in a structural format.—Nucleic Acid Research 16 (supplement): r175–r269.

Harris, P. M., and R. L. Mayden. 2001. Phylogenetic relationships of major clades of Catostomidae (Teleostei: Cypriniformes) as inferred from mitochondrial SSU and LSU rDNA sequences.—Molecular Phylogenetics and Evolution 20: 225–237.

Harrison, M. K., and B. J. Crespi. 1999. Phylogenetics of *Cancer* crabs (Crustacea: Decapoda: Brachyura).—Molecular Phylogenetics and Evolution 12: 186–199.

Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis.—Systematic Biology 42: 182–192.

———, and J. P. Huelsenbeck. 1992. Signal, noise, and reliability in molecular phylogenetic analyses.—Journal of Heredity 83: 189–195.

Huelsenbeck, J. P. 1991. Tree-length distribution skewness: an indicator of phylogenetic information.—Systematic Zoology 40: 257–270.

———. 1997. Is the Felsenstein Zone a flytrap?—Systematic Biology 46: 69–74.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. Munro, ed. Mammalian Protein Metabolism. Academic Press, New York.

Keigwin, L. D., Jr. 1982. Isotopic paleoceanography of the Caribbean and East Pacific: role of the Panama uplift in Late Neogene time.—Science 217: 350–353.

Kitaura, J., K. Wada, and M. Nishida. 1998. Molecular phylogeny and evolution of unique mud-using territorial behavior in ocypodid crabs (Crustacea: Brachyura: Ocypodidae).—Molecular Biology and Evolution 15: 626–637.

Kjer, K. M. 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs.—Molecular Phylogenetics and Evolution 4: 314–330.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae: Serpentes).—Systematic Zoology 38: 7–25.

Maddison, W. P., and D. R. Maddison. 2000. MacClade 4: Analysis of phylogeny and character evolution. Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Maggioni, R., A. D. Rogers, N. Maclean, and F. D'Incao. 2001. Molecular phylogeny of western Atlantic *Farfantepenaeus* and *Litopenaeus* shrimp based on mitochondrial 16S partial sequences.—Molecular Phylogenetics and Evolution 18: 66–73.

Malfait, B. T., and M. G. Dinkelman. 1972. Circumcaribbean tectonic and igneous activity and the evolution of the Caribbean Plate.—Geological Society of America Bulletin 83: 251–272.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach.—Cancer Research 27: 209–220.

Meyran, J.-C., M. Monnerot, and P. Taberlet. 1997. Taxonomic status and phylogenetic relationships of some species of the genus *Gammarus* (Crustacea, Amphipoda) deduced from mitochondrial DNA sequences.—Molecular Phylogenetics and Evolution 8: 1–10.

Miers, E. J. 1879. Revision of the Hippidea.—Journal of the Linnean Society, Zoology 14: 312–337.

Naylor, G. J., and W. M. Brown. 1998. *Amphioxus* mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences.—Systematic Biology 47: 61–76.

Palumbi, S. R., and J. Benzie. 1991. Large mitochondrial DNA differences between morphologically similar Penaeid shrimp.—Molecular Marine Biology and Biotechnology 1: 27–34.

Pollock, D. D., and D. B. Goldstein. 1994. A comparison of two methods for constructing evolutionary distances from a weighted contribution of transition and transversion differences.—Molecular Biology and Evolution 12: 713–717.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution.—Bioinformatics 14: 817, 818.

Remigio, E. A., and P. D. N. Hebert. 2000. Affinities among anostracan (Crustacea: Branchiopoda) families inferred from phylogenetic analyses of multiple gene sequences.—Molecular Phylogenetics and Evolution 17: 117–128.

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

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees.—Molecular Biology and Evolution 4: 406–425.

Sankoli, K. N. 1962. On a new species of *Emerita* (Decapoda, Anomura) from India, with a note on *Emerita emeritus* (L.).—Crustaceana 8: 48–54.

Schubart, C. D., J. E. Niegel, and D. L. Felder. 2000. Use of the mitochondrial 16S rRNA gene for phylogenetic and

population studies of Crustacea. Pp. 817–830 *in* F. R. Schram, ed. The Biodiversity Crisis and Crustacea. Crustacean Issues 12.

Schubart, D. S., R. Diesel, and S. B. Hedges. 1998. Rapid evolution to terrestrial life in Jamaican crabs.—Nature 393: 363–365.

Schmitt, W. L. 1937. A new species of *Emerita* (Crustacea) from South Africa.—Annals of the South African Museum 32: 25–29 , pl. 3.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference.—Molecular Biology and Evolution 16: 1114–1116.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and compilation of conserved polymerase chain reaction primers.—Annals of the Entomological Society of America 87: 651–701.

Steel, M., and D. Penny. 2000. Parsimony, likelihood and the role models in molecular phylogenetics.—Molecular Biology and Evolution 17: 839–850.

Sturmbauer, C., J. S. Levinton, and J. Christy. 1996. Molecular phylogeny analysis of fiddler crabs: test of the hypothesis of increasing behavior complexity in evolution.—Proceedings of the National Academy of Sciences, USA 93: 10855–108857.

Swofford, D. L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

———, G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. Hillis, C. Mortz, and B. K. Mable, eds. Molecular Systematics. Sinauer Associates, Sunderland, Massachusetts.

Tam, Y. K., and I. Kornfield. 1998. Phylogenetic relationships of clawed lobster genera (Decapoda: Nephropidae) based on mitochondrial 16S rRNA gene sequences.—Journal of Crustacean Biology 18: 138–146.

———, ———, and F. P. Ojeda. 1996. Divergence and zoogeography of mole crabs, *Emerita* spp. (Decapoda: Hippidae), in the Americas.—Marine Biology 125: 489–497.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.—Molecular Biology and Evolution 10: 512–526.

Xiong, B., and T. D. Kocher. 1991. Comparison of mitochondrial DNA sequences of seven morphospecies of black flies (Diptera: Simuliidae).—Genome 34: 306–311.

Zharkikh, A., and W.-H. Li. 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock.—Molecular Biology and Evolution 9: 1119–1147.

RECEIVED: 12 October 2001.
ACCEPTED: 5 March 2002.