

Available online at www.sciencedirect.com



Molecular Phylogenetics and Evolution 39 (2006) 491-502

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

Probing the relationships of the branchiopod crustaceans

Jeremy R. deWaard ^{a,*}, Veronika Sacherova ^b, Melania E.A. Cristescu ^c, Elpidio A. Remigio ^d, Teresa J. Crease ^a, Paul D.N. Hebert ^a

^a Department of Integrative Biology, University of Guelph, Guelph, Ont., Canada N1G2W1

^b Department of Hydrobiology, Charles University, Prague, CZ-128 44, Czech Republic

^c Department of Biology, Indiana University, Bloomington, IN 47405-3700, USA

^d Museum of Zoology, University of Michigan, Ann Arbor, MI 48109-1079, USA

Received 26 August 2005; accepted 1 November 2005 Available online 9 January 2006

Abstract

The Branchiopoda display extraordinary variation in body form, even within the morphologically diverse crustaceans. To fully understand the origin and evolution of these morphological reconfigurations, a robust phylogeny of the group is essential. To infer the affinities among branchiopods, we employed two approaches to taxon and gene sampling, presented new sequence data from three genes, incorporated previously published sequence data from three additional genes, and utilized comprehensive techniques of phylogeny reconstruction. The results provided support for a number of longstanding hypotheses concerning the relationships among the orders. For example, we obtained support for the Cladoceromorpha and Gymnomera, and favoured a unique arrangement of the cladoceran orders. A few affinities remain to be resolved, particularly at the base of the Phyllopoda and within the Anomopoda. However, the results suggest that increased gene sampling is recommended for future investigations of branchiopod systematics. © 2005 Elsevier Inc. All rights reserved.

Keywords: Branchiopoda; Cladocera; Phylogeny; COI; Taxon sampling; Gene sampling

1. Introduction

In contrast to other arthropod lineages, which show limited bauplan diversity, the crustaceans show striking variation in body form. However, among the eight commonly recognized crustacean classes (Martin and Davis, 2001), the class Branchiopoda shows exceptional diversity, especially given the fact that it includes just 800 described species. In contrast to the relatively static body plans of the other classes, these crustaceans display marked variation in their body segmentation patterns, and in the morphology, number, and function of their limbs. While this provides a unique setting to investigate the exploration of phenotypic space, our understanding of the origin and evolution of these morphological reconfigurations is inhibited by our lack of a robust phylogeny for the group.

^{*} Corresponding author. Fax: +1 519 767 1656.

E-mail address: jdewaard@uoguelph.ca (J.R. deWaard).

A great deal of effort has been expended to determine the relationships between the eight orders and 24 families of extant branchiopods (reviewed in Fryer, 1995; Martin and Davis, 2001; Spears and Abele, 2000). Past studies have included examinations of fossil taxa (e.g. Walossek, 1993, 1995), investigations of the embryology, ontogeny, and morphology of key species (e.g. Olesen et al., 1997, 2003; Olesen, 1999), and phylogenetic analyses of morphological (e.g. Negrea et al., 1999; Olesen, 1998, 2000) and molecular characters (Braband et al., 2002; Hanner and Fugate, 1997; Richter et al., 2001; Schwenk et al., 1998; Spears and Abele, 2000; Swain and Taylor, 2003; Taylor et al., 1999). In addition, a few recent studies have established the utility of several complex genetic characters, including rRNA secondary structural motifs (Swain and Taylor, 2003), the distribution patterns of introns (Braband et al., 2002) and rRNA expansion segments (Crease and Taylor, 1998). These studies, a diverse assemblage in themselves, have failed to achieve the holy grail: a consensus on branchiopod relationships.

^{1055-7903/\$ -} see front matter © 2005 Elsevier Inc. All rights reserved. doi:10.1016/j.ympev.2005.11.003



Fig. 1. Currently accepted phylogeny of the class Branchiopoda. (A) Relationships of the large branchiopod orders and the Cladocera. (B) Relationships among the cladoceran orders. Affinities that have not been reliably resolved are drawn as polyphyletic.

Although the phylogeny of the branchiopods has been redrawn on numerous occasions, many details remain incomplete (Fig. 1). The class is generally divided into eight extant orders: the Anostraca (fairy shrimps); the Notostraca (tadpole shrimps); the Laevicaudata and Spinicaudata (collectively known as clam shrimps and previously classified together as the Conchostraca); and the final four orders (Anomopoda, Ctenopoda, Haplopoda, and Onychopoda) which collectively comprise the Cladocera (water fleas) (Fryer, 1995). The affinities between these orders remain unclear, although the placement of Anostraca as the sister group to the remaining branchiopods (Phyllopoda) is well supported (Negrea et al., 1999; Olesen, 1998; Spears and Abele, 2000), as is the monophyly of the Cladocera (Braband et al., 2002; Spears and Abele, 2000; Taylor et al., 1999). Furthermore, Cyclestheria hislopi, formerly placed within the Spinicaudata, unquestionably represents the sister lineage to the Cladocera (= Cladoceromorpha) (e.g. Ax, 1999; Crease and Taylor, 1998; Spears and Abele, 2000). Within the Cladocera, several studies have corroborated the monophyly of the predatory water fleas (Haplopoda+ Onychopoda = Gymnomera) (Richter et al., 2001; Swain and Taylor, 2003), but their relationship to the Anomopoda and Ctenopoda, as well as the affinities of varied cladoceran families, remains under debate.

This study, which expands on previous molecular studies, seeks to obtain a well-supported phylogeny for the branchiopod orders and families, particularly within the speciose Cladocera, and to identify areas of ambiguity that require further study.

2. Materials and methods

2.1. Taxonomic sample and gene selection

This study employed two approaches to data collection. The first, a 'more taxa' (MT) approach (e.g. Pollock et al., 2002), involved the collection of sequence data from a broad diversity of branchiopods for three gene fragments. The second approach, a 'more genes' (MG) approach (e.g. Rosenberg and Kumar, 2001), added sequence data for three additional genes for a subset of the original taxa.

In the MT approach, we examined 56 taxa that included representatives from 22 of the 24 recognized branchiopod families (Martin and Davis, 2001, Table 1). In addition to maximizing taxonomic breadth, an effort was made to include multiple species of key genera (e.g. *Lynceus*) and to provide good representation of speciose taxa (e.g. Chydoridae). Specimens were collected from numerous locations in North America, South America, Europe, and Australia. The complete collection and locality details are available from the authors upon request.

The malacostracan *Anaspides tasmaniae* was employed as the outgroup in the analysis of these data. While prior studies have strongly supported the monophyly of the Branchiopoda (e.g. Sanders, 1963; Spears and Abele, 2000; Walossek, 1993; Wingstrand, 1978), their relation to other Crustacea remains uncertain (reviewed in Martin and Davis, 2001, and references therein). *A. tasmaniae* certainly lies outside the branchiopods, but it may not be their closest relative. However, our preliminary analyses indicated that (a) when multiple outgroups were included, malacostracans were the sister taxa to a monophyletic branchiopod clade and (b) topology was unaffected by the choice of single or multiple outgroups from other crustacean classes (trees not shown).

The MT approach involved the analysis of sequences from two mitochondrial genes; *cytochrome c oxidase subunit 1* (COI) and the *large subunit 16S rRNA* (16S), and a nuclear gene; the *small subunit 18S rRNA* (18S). These three genes were chosen because of their wide use in past studies of arthropod phylogenetics (e.g. Giribet et al., 2001), including studies of branchiopod relationships (Cristescu and Hebert, 2002; Remigio and Hebert, 2000; Sacherova and Hebert, 2003; Spears and Abele, 2000). Furthermore, these three loci provide phylogenetic signal over varying time scales, aiding the resolution of both shallow and deep nodes. This dataset is referred to as the 'MT' dataset.

In the MG approach, we gathered additional sequence information for a subset of 17 taxa, representing all eight Table 1

Taxonomic sample analysed in this study, with GenBank accession numbers

Taxonomy	Species	COI	16S	18S	EF-1α	12S	28S
Class Branchiopoda							
Order Anostraca							
Family Artemiidae	Artemia franciscana	NC_001620	AF209051	X01723	X03349	X69067	^a , AY137143
Family Thamnocephalidae	Thamnocephalus platyurus	AF209066	AF209057	AF144217			_
	Branchinella pinnata	AF308940	DQ310661*	DQ310583*	_	_	_
Family Branchionectidae	Branchinecta paludosa	AF209064	AF209055	AF144206			—
Family Streptocephalidae	Streptocephalus dorothae	AF209065	AF209056	AF144218			—
Family Branchipodidae	Parartemia contracta	AF209059	AF209048	DQ310584*	_	_	_
Family Chirocephalidae	Artemiopsis stefanssoni	AF209062	AF209053	DQ310585*			—
Es miles Dalamatan ii da a	Eubranchipus sp.	AF209061	AF209052	DQ310586*	_	_	_
Family Polyartemildae	Polyartemiella nazeni	AF209063	AF209054	DQ310387	_	_	—
Order Notostraca							
Family Triopsidae	Lepidurus sp.	AF209067	AF209058	AF144212	AF526293	AF494483	AF209047, AY137138
	Lepidurus couessi	DQ310622*	DQ310662*	DQ310588*	—	—	
	Triops sp.	DQ310623*	DQ310663*	DQ310589*	U90058	AF494482	^a , AY137137
	Triops australiensis	DQ310624*	DQ310664*	DQ310590*	_	_	—
	<i>Triops</i> sp. nov.	DQ310625*	DQ310665*	DQ310591*			—
Order Laevicaudata							
Family Lynceidae	Lynceus sp. 1	DQ310626*	DQ310666*	AF144215	AF526294	AF494479	^a , AY137136
	Lynceus sp. 2	DQ310627*	DQ310667*	DQ310592*			_
Order Spinicaudata							
Family Caenestheriidae	Caanastharialla satosa	DO310628*	DO310668*	DO310593*			
I anny Cachestherndae	Caenestheriella sp	DQ310629*	DQ310669*	DQ310594*			
Family Limnadiidae	Limnadia sp.	DQ310630*	DQ310670*	DQ310595*	AF063412	A F494471	AF532886
"Cyclestherida"	Linnuuu sp.	DQ310030	DQ310070	DQ310373	AI 005412		AI 552660
Family Cyclestheriidae "Cladocera"	Cyclestheria hislopi	DQ310631*	DQ310671*	AF144209	AF526292	AF494478	AF532878
Order Anomopoda							
Family Daphniidae	Daphnia pulex	NC 000844	NC 000844	AF014011			_
	Scapholeberis rammneri	DO310632*	DO310672*	DO310596*	AF526282	AF494465	AF532880
	Simocephalus vetulus	DQ310633*	DQ310673*	AF144216	AF526281	AY009492	AF532887
	<i>Ceriodaphnia</i> sp.	DQ310634*	DQ310674*	AF144208	AF526283	AF494466	AF532889
Family Bosminidae	Bosmina sp. 1	DQ310635*	DQ310675*	DQ310597*	AF526284	AF494467	AF482744
	Bosmina sp. 2	DQ310636*	DQ310676*	DQ310598*			
Family Macrothricidae	Ophryoxus gracilis	DQ310637*	DQ310677*	DQ310599*	_	_	_
	Ilyocryptus sp.	DQ310638*	DQ310678*	DQ310600*		_	_
	Acantholeberis curvirostris	DQ310639*	DQ310679*	DQ310601*			_
	Macrothrix sp.	DQ310640*	DQ310680*	DQ310602*			_
	Drepanothrix dentata	DQ310641*	DQ310681*	DQ310603*	_		_
Family Chydoridae							
Subfamily Chydorinae	Chydorus brevilabris	DQ310642 ^b	DQ310682 ^b	DQ310604 ^b	AF526286c	AF494469 ^c	AF532891
	Alonella exigua	DQ310643 ^b	DQ310683 ^b	DQ310605 ^b			—
	Pleuroxus denticulatus	DQ310644 ^b	DQ310684 ^b	DQ310606 ^b			—
	Dunhevedia crassa	DQ310645 ^b	DQ310685 ^b	DQ310607 ^b			—
Subfamily Aloninae	Alona setulosa	DQ310646 ^b	DQ310686 ^b	DQ310608 ^b	_	_	—
	Camptocercus rectirostris	DQ310647 ⁶	DQ310687 ⁶	DQ310609 ^b	_	_	_
	Acroperus harpae	DQ310648 ⁶	DQ310688 ⁶	DQ310610 ⁶			
	Graptoleberis testudinaria	DQ310649 ⁶	DQ310689 ⁶	DQ310611 ^b			—
Subfamily Sayciinae	Saycia cooki	DQ310650°	DQ310690°	DQ310612 ⁶			 a
Subfamily Eurycerinae	Eurycercus longirostris	DQ310651°	DQ310691°	DQ310613°	AF526285	AF494468	u
Equily Mainidaa	Eurycercus giacialis	DQ310652*	DQ310692*	DQ310614*	_	_	_
Family Mollidae	Moina sp. 1 Moina sp. 2	DQ310654*	DQ310693 DQ310694*	DQ310615 DQ310616*	_	_	
Order Onychopoda							
Family Polynhemidae	Polynhemus nediculus	AY075048	AY075066	DO310617*			
Family Cerconagidae	1 orypnemus peuteutus Cerconagis pengoi	AF320013	A Y075067	DO310618*	_	_	_
i anny Cereopagidae	Bythotrephes cederstroemi	DO310655*	DO310695*	AF144207		_	
Family Podonidae	Evadne spinifera	DO310656*	AY075071	AY075085	AF526288	AY009498	AF532906 AY137167
i anny i oaomuue	Podonevadne sp	AY189520	AY075078	AY075092			
	Podon leuckarti	AY075051	AY075073	AY075087	AF526287	AY009496	AF532901. AY137147
	Pleopis polyphemoides	AY075050	AY075072	AY075086		_	
	1 1 11 11 11 11 11						(continued on next page)

Table 1	(continued)
---------	------------	---

Taxonomy	Species	COI	16S	18S	EF-1α	12S	28S
Order Ctenopoda							
Family Sididae	Sida crystallina	DQ310657*	DQ310696*	DQ310619*	AF526280	AY009489	AF532873
	Diaphanosoma sp.	DQ310658*	DQ310697*	AF144210	AF526279	AY009490	AF532910, AY137155
Family Holopedidae	Holopedium gibberum	AF245354	DQ310698*	DQ310620*	_	_	—
Order Haplopoda							
Family Leptodoridae	Leptodora kindtii	DQ310659*	DQ310699*	AF144214	AF526278	AY009488	AF532877
Class Malacostraca							
Order Syncarida	Anaspides tasmaniae	DQ310660*	DQ310700*	DQ310621*		_	—

Asterisk (*) denote sequences novel to this study and dashes (—) indicate missing sequences. For the 28S, 12S, and EF-1 α genes, the same species or a congeneric taxon was acquired from GenBank and/or previous studies.

^a Sequence was not deposited in GenBank; copied from the supplementary materials of Taylor et al. (1999).

^b Sequence from Sacherova and Hebert (2003).

^c The confamilial and closely allied *Pseudochydorus globosus* used to represent *Chydorus sphaericus* for EF-1 α and 12S.

orders of branchiopods. These genes included the mitochondrial *small subunit 12S rRNA* (12S), the nuclear *large subunit 28S rRNA* (28S), and the nuclear protein-coding gene *elongation factor 1 alpha* (EF-1 α). These gene regions have proven informative in the studies from which the sequences were extracted (Braband et al., 2002; Cristescu and Hebert, 2002; Hanner and Fugate, 1997; Sacherova and Hebert, 2003; Swain and Taylor, 2003; Taylor et al., 1999). Due to the lack of a suitable outgroup species represented by all six genes outside the Branchiopoda, we rooted all trees with the anostracan *Artemia franciscana*. We term this the 'MG' dataset.

2.2. Molecular techniques and sequence alignments

Genomic DNA was extracted from whole animals using 25–50 µL aliquots of proteinase K extraction buffer and the method described in Palumbi (1996). The primer pairs LCO1490/HCO2198 (Folmer et al., 1994) and 16Sar/16Sbr (Palumbi, 1996) were used to PCR amplify a 680-base pair (bp) fragment of COI and a 570-bp fragment of 16S, respectively. An approximately 1995-bp fragment of 18S was amplified with the primers 9F (5'-TGG GGA TCA TTG CAG TTC CCA ATC-3'; designed by TJC) and 2004R (Crease and Colbourne, 1998) with about 800 bp near the 5' terminus targeted for sequencing. The $50 \,\mu L$ PCR reactions contained $0.5-2.0 \,\mu\text{L}$ (out of $25-50 \,\mu\text{L}$) of DNA template, $5.0 \,\mu\text{L}$ 10× PCR buffer (Roche), $0.2 \,\mu\text{M}$ of each primer, 2.2 mM MgCl₂, 0.2 mM of each dNTP, and 1 unit of Taq DNA polymerase. The PCR conditions for COI and 16S consisted of 1.5 min at 94 °C, followed by 35 cycles of 45 s at 93 °C, 1 min at 50 °C and 1 min at 72 °C, followed by 1 cycle of 5 min at 72 °C. The PCR conditions for 18S consisted of 1 cycle at 94 °C, 35 cycles of 30 s at 93 °C, 30 s at 50 °C, and 3 min at 72 °C, followed by 1 cycle of 5 min at 72 °C. PCR products were excised from agarose gels and purified using the Qiaex II gel extraction kit (Qiagen) and sequenced using an ABI 377 automated sequencer and the ABI prism BigDye terminator 3 sequencing kit (Applied Biosystems). Gene products were sequenced in both directions or twice in the same direction whenever ambiguous sites were encountered. Some

sequences used in our analysis were obtained from previously published studies (Table 1).

DNA sequences were initially aligned in Sequence Navigator (Applied Biosystems). The alignments for the ribosomal genes required adjustments with reference to proposed secondary structure models (Crease and Colbourne, 1998; De Rijk et al., 2000; Taylor et al., 1998; Van de Peer et al., 2000). Sites within the ribosomal genes that were not easily aligned were excluded from subsequent analyses. The sequence alignments are available for download from the cladoceran website (http://www.cladocera.uoguelph.ca). Sequences obtained for the COI, 16S, and 18S loci that are new to this study have been deposited in GenBank under Accession Nos. AYDQ310583– AYDQ310700 (Table 1).

2.3. Phylogenetic analysis

To reconstruct the phylogenetic relationships of the ingroup taxa, we concatenated the nucleotide sequence alignments for the three and six genes included in the MT and MG datasets, respectively. Tree-building was performed by maximum likelihood (ML), Bayesian inference (BI), and maximum parsimony (MP). These three techniques were used because concordance among different analytical approaches strengthens support for the tree (Cunningham, 1997) and because there are varied opinions on how to best reconstruct phylogenies (Crandall et al., 2000). The best-fit model of sequence evolution was selected by analysing distance-based topologies with hierarchical likelihood ratio tests using ModelTest 3.06 (Posada and Crandall, 1998) following the procedure outlined by Huelsenbeck and Crandall (1997). The ML analysis was performed in PAUP* v4.0b10 (Swofford, 2001) using the model and parameters estimated in ModelTest and the heuristic search option (10 replicates, one tree held per replicate, sequences added at random, branch swapping by nearest neighbour interchanges, starting tree obtained by neighbour-joining).

The BI analysis was computed in the program MrBayes 2.01 and 3.0b4 (Huelsenbeck and Ronquist, 2001). Again, the model and parameters estimated by ModelTest were used for the analysis. Three independent runs, each consisting of four Markov chains, were run for 1,050,000 generations, with the first 50,000 generations discarded as the burn-in. Each run was inspected to ensure that likelihood stationarity was reached during burn-in, and that parameters and posterior probabilities were consistent between runs. MP trees were estimated with PAUP* using the unweighted heuristic search option (1000 replicates, 100 trees held per replicate, sequences added at random, and tree bisection-reconnection branch swapping) on the parsimony-informative sites.

To explore the integrity and dynamics of the results, we performed several tests. First, a goodness-of-fit (χ^2) test, as implemented in PAUP*, was performed to test for stationarity in base composition among taxa. We evaluated the strength of the phylogenetic signal in the datasets by calculating the g_1 skewness statistic (Hillis and Huelsenbeck, 1992) and performing a relative apparent synapomorphy analysis (RASA) (Lyons-Weiler et al., 1996). Despite the inherent limitations associated with these two tests (Källersjö et al., 1992; Simmons et al., 2002), their use in combination should provide a reliable indicator of the presence or absence of phylogenetic signal. To assess confidence in the phylogenies, we performed nonparametric bootstrapping (100 pseudoreplicates for ML, 1000 for MP). Finally, incongruence length difference tests (Farris et al., 1994) were employed to determine the congruence of phylogenetic signal from the different genes. These tests had no bearing on our decision to combine the genes for a totalevidence approach (Kluge, 1989), but instead, to investigate the nature of the phylogenetic signal (Remsen and DeSalle, 1998).

2.4. Comparison of approaches

In addition to performing the tests outlined above on the two datasets, we completed a further evaluation to address the taxon versus gene sampling issue. For this evaluation, we used the 17 taxa for which sequence data from all six genes are available (see Table 1). First, we investigated the effect of increased gene sampling by generating trees for all possible combinations of three, four, five, and six genes. Second, we investigated the effect of increased taxon sampling by evaluating trees reconstructed with the MT dataset. The 17 taxa in the MG dataset were included in each analysis with 12, 24, or 36 additional taxa from the MT dataset. In each case, 10 trees were constructed after randomly choosing the additional taxa. The neighbour-joining algorithm and Kimura's two-parameter model (Kimura, 1980) was employed for tree reconstruction using the program MEGA3 (Kumar et al., 2004). All trees were checked for five commonly accepted and well-supported relationships (see Fig. 1): Anomopoda, Ctenopoda, Gymnomera, Cladocera, and Cladoceromorpha. In addition, mean bootstrap support across all nodes was calculated for each tree.

3. Results

3.1. More taxa approach

The final alignment for the MT dataset was 1546 bp in length, comprising 639, 353, and 554 bp long fragments of COI, 16S, and 18S, respectively. The hierarchical likelihood ratio tests indicated that the best-fit model for subsequent analysis was the general time reversible model with invariable sites and gamma shape parameter (GTR + I + G) with the following parameters selected: unequal base frequencies: A = 0.34, C = 0.14, G = 0.13, T = 0.39; six substitution categories: $A \rightarrow C = 0.46$; $A \rightarrow G = 4.85$, $A \rightarrow T = 0.72$, $C \rightarrow G = 1.05$, $C \rightarrow T = 5.97$, $G \rightarrow T = 1.00$; proportion of invariant sites = 0.48; and gamma distribution shape parameter = 0.40.

The BI and ML analyses produced nearly the same topology (Fig. 2). The node support, assessed with posterior probabilities, was generally high (>80) for nodes at the family level and above. The MP analysis resulted in 26 equally parsimonious trees with a length of 6536 steps (consistency index = 0.18; retention index = 0.40). In contrast to the BI and ML tree, the MP analysis failed to recover several nodes (bootstrap percentages <50; Fig. 3).

3.2. More genes approach

The length of the final alignment for the MG dataset, including the 1546 bp from the MT dataset, was 4096 bp. The sequence data, taken from GenBank and previous studies (Table 1), consisted of 1300 bp of 28S, 250 bp of 12S, and 1000 bp of EF-1 α . The best-fit model selected for this dataset was also the GTR + *I* + *G* model, with the following parameters: unequal base frequencies: A = 0.25, C = 0.23, G = 0.26, T = 0.26; six substitution categories: $A \rightarrow C = 1.09$; $A \rightarrow G = 4.73$, $A \rightarrow T = 4.22$, $C \rightarrow G = 1.42$, $C \rightarrow T = 9.45$, $G \rightarrow T = 1.00$; proportion of invariant sites = 0.45; and gamma distribution shape parameter = 0.69.

All three tree-building approaches used on the MG dataset produced an identical topology (Fig. 4). There was a single most parsimonious tree that is 4950 steps in length (consistency index = 0.48; retention index = 0.35) with only two nodes not recovered with bootstrap support >50. In general, all three trees had modest to strong node support.

3.3. Comparison of approaches

The two approaches to taxon sampling had a significant impact on the specific hypotheses supported by the results, as well as the strength of these nodes (Table 2). To a lesser extent, and mostly limited to the MT dataset, the tree-building approach also impacted the outcome (Table 2). In general, the MT dataset provided good resolution of relationships at the family level and below, but only moderate or poor resolution of deeper divergences. On the other hand, the MG dataset provided good resolution of interordinal divergences, particularly within the Cladoceromorpha. However, there was disagreement between the two approaches on several important



0.1 substitutions/site

Fig. 2. Phylogenetic relationships of the Branchiopoda as determined by maximum likelihood analysis ($-\ln L = 35,100.7$) of the combined COI, 16S, and 18S (MT) dataset for 56 taxa. The Bayesian inference tree was identical in topology except for two nodes that were left unresolved and are marked by an asterisk (*). The tree was rooted with the outgroup *Anaspides tasmaniae*. Branch lengths are proportional to reconstructed distances. Posterior probabilities are given for the nodes at the family level or above.

nodes, for example, the position/monophyly of Cladocera and Gymnomera (Table 2).

It is likely that several factors contribute to the discrepancies in topology obtained with the two datasets. The MG dataset is roughly twice as large as the MT dataset, with respect to both invariant and parsimony-informative characters (Table 3). Our calculation of the g_1 skewness statistics, and the RASA test statistics (Table 3) indicate that there is significant phylogenetic signal in the two datasets, ruling this out as a factor in their incongruence. In contrast, chi-square tests (Table 4) provide evidence for heterogeneous nucleotide composition across the taxa in the MG dataset, which appears to derive from the two protein-coding genes. In addition, the partition homogeneity tests indicated heterogeneity of phylogenetic signal from the various genes. We detected significant heterogeneity ($P \le 0.01$) in all comparisons performed: all three genes (MT dataset), all six genes (MG dataset), mitochondrial genes only, nuclear genes only, protein-coding genes only, and ribosomal genes only.

Our additional evaluation of increased gene and taxon sampling approaches revealed that adding genes, but not



Fig. 3. Phylogenetic relationships of the Branchiopoda as determined by maximum parsimony of the combined COI, 16S, and 18S (MT) dataset for 56 taxa. The majority rule consensus cladogram of the 26 equally parsimonious trees (length = 6536) is shown. The tree was rooted with the outgroup *Anaspides tasmaniae*. MP bootstrap percentages are given for the resolved nodes with values >50.

taxa, had a positive effect on phylogenetic accuracy (Fig. 5). An increase in genes was accompanied by an increase in both average node support and ability to reconstruct accepted clades. Conversely, no trend was apparent with the addition of taxa.

4. Discussion

4.1. Branchiopod interordinal affinities

This study provides support for a number of longstanding hypotheses concerning higher level branchiopod relationships. The monophyletic status of Anostraca, Notostraca, Laevicaudata, and Spinicaudata (excluding *Cyclestheria*) are supported in all analyses and using both datasets, which is consistent with recent studies (e.g. Braband et al., 2002; Spears and Abele, 2000). Cladocera is also found to be a monophyletic group with the MG dataset and in the MP analysis of the MT dataset. Conversely, the clam shrimp *Cyclestheria* groups among the Cladocera in the ML and BI trees constructed from the MT dataset, rendering the latter paraphyletic. The low node support of this placement, and the deep divergences in the MT trees in general, cause us to favour cladoceran monophyly and instead interpret this as support for the Cladoceromorpha concept (Ax, 1999). All other trees support a *Cyclestheria* +



0.1 substitutions/site

Fig. 4. Phylogenetic relationships of the Branchiopoda as determined by Bayesian inference of the combined COI, 16S, 18S, 28S, 12S, and EF-1 α (MG) dataset for 17 taxa. Topologies of the maximum likelihood (ML) tree ($-\ln L = 29,471.7$) and the single most parsimonious tree (length = 4950) are identical. The tree was rooted with the anostracan *Artemia franciscana*. Branch lengths are proportional to reconstructed distances. Branch support values are given for all nodes and are given as BI posterior probabilities/ML bootstrap probabilities/MP bootstrap probabilities. A dash (–) indicates an MP bootstrap probability of <50.

Cladocera sister group relationship, which now seems uncontroversial (e.g. Crease and Taylor, 1998; Spears and Abele, 2000; Swain and Taylor, 2003).

As in Braband et al. (2002) and Spears et al. (2000), we are unable to determine the exact relationships among the large branchiopod orders. The MG dataset suggests that the Laevicaudata may be the sister taxon to the remaining groups of the Phyllopoda, whereas MP analysis of the MT dataset places Laevicaudata as the sister group to the Notostraca. Both of these hypotheses are congruent with the analysis of Braband et al. (2002). Also consistent with Braband et al. (2002), as well as Spears and Abele (2000), is the close affinity between Spinicaudata and Cladoceromorpha suggested by analysis of the MG dataset. This suggestion challenges the traditional "Conchostraca' taxon (Negrea et al., 1999; Schram, 1986; Walossek, 1993), but requires further confirmation.

4.2. Relationships within the Cladocera

Our analysis also provides support for previous hypotheses concerning relationships within the Cladocera. Inferences

499

Table 2

Support	for hype	otheses c	of branchiopod	l relationships.	. Examples o	f previous studie	s that support	the hypotheses	are given
					<u> </u>	*		• •	-

Hypothesis	Examples of previous support ^a	Support from present study							
		More ta	More taxa approach (MT dataset)			More genes approach (MG dataset)			
		MP	ML	BI	MP	ML	BI		
Phyllopoda	1, 2, 3, 4, 5	Ν	Ν	Ν			_		
Diplostraca	1, 2, 3, 5, 6	Ν	Ν	Ν	Ν	Ν	Ν		
Conchostraca	1, 2, 4	Ν	Ν	U	Ν	Ν	Ν		
Cladoceromorpha	4, 5, 6, 7, 8, 9, 10	Y	Y	Y*	Y	Y*	Y*		
Cladocera	2, 3, 4, 5, 6, 11	Y	Ν	Ν	Y*	Y*	Y*		
Spinicaudata + Cladoceromorpha	5, 6	Ν	Ν	Ν	Y	Y	Y*		
Ctenopoda + Anomopoda + Onychopoda (= Eucladocera)	5, 12, 13, 14	Y	Ν	Ν	Ν	Ν	Ν		
Anomopoda + Haplopoda + Onychopoda	11, 15	Ν	Ν	Ν	Ν	Ν	Ν		
Ctenopoda + Anomopoda (= Calyptomera)	4, 5	Ν	Ν	Ν	Y	Y	Y*		
Ctenopoda + Haplopoda + Onychopoda	6	Ν	Ν	Ν	Ν	Ν	Ν		
Gymnomera	3, 10, 11, 15, 16	Ν	Ν	Ν	Y	Y	Y*		
Calyptomera + Gymnomera		Ν	Ν	Ν	Y*	Y*	Y*		
Moininae within Daphniidae	3, 17	Ν	Ν	Ν			_		
Radopoda	18	Ν	Ν	Ν			_		
Chydoridae	3, 17	Ν	Ν	Ν		_	_		
Macrothricidae paraphyly	3, 13	Y	Y	Y*			—		

Support from the different trees estimated in the present study is also given, where Y signifies support for the hypothesis and N signifies no support. Trees with strong support (posterior probabilities or bootstrap percentages >90) from the present study are denoted by an asterisk (*).

^a Sources are as follows: (1) Schram, 1986; (2) Walossek, 1993; (3) Olesen, 1998; (4) Negrea et al., 1999; (5) Spears and Abele, 2000; (6) Braband et al., 2002; (7) Crease and Taylor, 1998; (8) Ax, 1999; (9) Taylor et al., 1999; (10) Swain and Taylor, 2003; (11) Martin and Cash-Clark, 1995; (12) Eriksson, 1934; (13) Wingstrand, 1978; (14) Bowman and Abele, 1982; (15) Schwenk et al., 1998; (16) Richter et al., 2001; (17) Fryer, 1995; (18) Dumont and Silva-Briano, 1998.

Table 3

Sequence statistics for the more taxa ((MT) and more gene	s (MG) datasets
---	--------------------	-----------------

Dataset	bp	Variable sites	Informative sites	g_1 statistic	Р	t _{RASA}	Р
MT dataset	1546	686	588	$-0.52 \\ -0.94$	<0.01	15.9	<0.001
MG dataset	4096	1543	1087		<0.01	12.8	<0.001

Variable and parsimony-informative characters and results of the g_1 skewness test and RASA test for each dataset is given.

Table 4

Base compositions of the six genes used in this study and results of χ^2 tests for base homogeneity

	bp	A	С	G	Т	χ^2	Р
COI	639	0.248	0.194	0.205	0.354	306.5	< 0.01
16S	353	0.288	0.157	0.239	0.316	95.8	1.00
18 S	554	0.255	0.244	0.267	0.233	21.9	1.00
MT dataset	1546	0.260	0.203	0.235	0.302	154.5 (165)	0.71
28S	1300	0.239	0.245	0.337	0.180	27.5	0.99
12S	250	0.318	0.196	0.206	0.281	22.5	1.00
EF-1α	1000	0.237	0.292	0.254	0.216	85.6	< 0.01
MG dataset	4096	0.252	0.237	0.266	0.245	66.5 (48)	0.04

employing both of the datasets and all three tree-building approaches support the monophyly of the orders Onychopoda, Ctenopoda, and Anomopoda. The MG analysis further supports two other hypotheses; the Calyptomera, comprised of the Ctenopoda and the Anomopoda (Negrea et al., 1999; Spears and Abele, 2000), and the Gymnomera (e.g. Richter et al., 2001; Swain and Taylor, 2003). In addition, the sister grouping of Calyptomera and Gymnomera is very well-supported, providing a new hypothesis of cladoceran relationships. The taxon sampling within the Anomopoda in our MT dataset allows inferences about the affinities within this large order. First, it would appear that the Moinidae, recently demoted to subfamily status within the Daphniidae (Fryer, 1995; Olesen, 1998), actually warrant their traditional family status, since they appear most closely related to the Bosminidae (ML and BI), and perhaps some macrothricid lineages (MP). Second, the two speciose families, Chydoridae and Macrothricidae, are paraphyletic in all analyses, suggesting that a revision of these two families is needed before anomopod affinities can be clarified.

4.3. Strategies of taxon and gene sampling

How best to approach data collection for phylogenetic estimation remains a contentious issue, and incomplete taxon sampling is often cited as a major source of error in phylogenetic studies (reviewed in Graybeal, 1998; Poe, 1998; Pollock et al., 2002; Rosenberg and Kumar, 2001). Increasing taxon sampling (e.g. Pollock et al., 2002) or increasing sequence length (e.g. Rosenberg and Kumar, 2001) are two opposing strategies, each with merit and empirical support. In the present study, we naturally do not know the 'true phylogeny' with which to determine with certainty the superior strategy. However, we were able to evaluate them with two metrics that we do possess: node support and recovery of generally accepted and well-supported relationships. Our results suggest that only increasing gene number positively impacts phylogenetic accuracy which is consistent with the lone empirical study that had been done previously (Rokas and Carroll, 2005). Other



Fig. 5. The effect of increasing gene number and taxon number on the phylogenetic accuracy of the branchiopod dataset. Trees were constructed by neighbour-joining and evaluated by calculating mean bootstrap support across all nodes and checking for the presence of five commonly accepted clades: Anomopoda, Ctenopoda, Gymnomera, Cladocera, and Cladoceromorpha. Error bars denote one standard deviation above and below the mean. (A) Results of the 'more genes' (MG) approach. For three, four, five, and six genes, all possible gene combinations (15, 20, 15, and 1, respectively) were used to construct trees. (B) Results of the MG approach. The 17 taxa in the MG analysis were included in each analysis. For 0, 12, 24, and 36 additional taxa, 10 trees were constructed using randomly chosen taxa.

work suggests that, for any given phylogenetic problem, there is a threshold amount of sequence data below which an increase in taxonomic sampling does not improve, or may even decrease, phylogenetic accuracy (Cummings et al., 1995; Mindell et al., 1997). Following this suggestion, it is unclear if the threshold was reached in the present study with six genes, but it was certainly not reached with three genes. For this reason, future work will likely benefit most from increased sequence sampling for the exemplar taxa used in this study, as opposed to more extensive taxon sampling.

5. Conclusions

In summary, we have provided the most comprehensive molecular study of branchiopod relationships to date, both in terms of taxonomic representation and the amount and diversity of sequence data. Our study highlights the importance of sampling strategies for future investigations of branchiopod systematics, as well as phylogenetic analyses in general. A few details of the branchiopod phylogeny remain incomplete, particularly near the root of the Phyllopoda and among families within the Anomopoda, but the lineage relationships have been clarified. In the near future, we may attain a complete and robust phylogeny, finally providing the vantage point needed to interpret the striking morphological reconfigurations of the branchiopods that hampered the creation of a phylogeny in the first place.

Acknowledgments

We sincerely thank S. Adamowicz, A. Cox, J. Green, S. Jarman, C. Rowe, B. Timms, and J. Witt for providing specimens, A. Hollis for assistance with sequencing, and J. Witt for help with the phylogenetic analyses. Funding for this project was provided by Ontario Graduate Scholarships to J.R.D. and M.E.A.C., a NATO Science Fellowship to V.S., a NSERC grant to T.J.C., and grants from NSERC and the Canada Research Chairs program to P.D.N.H.

References

- Ax, P., 1999. Das System der Metazoa II. G. Fischer Verlag, Stuttgart.
- Bowman, T.E., Abele, L.G., 1982. Classification of the recent Crustacea. In: Abele, L.G. (Ed.), Biology of the Crustacea, vol. 1. Academic Press, New York, pp. 1–27.
- Braband, A., Richter, S., Hiesel, R., Scholtz, G., 2002. Phylogenetic relationships within the Phyllopoda (Crustacea, Branchiopoda) based on mitochondrial and nuclear markers. Mol. Phylogenet. Evol. 25, 229–244.
- Crandall, K.A., Harris, D.J., Fetzner, J.W., 2000. The monophyletic origin of freshwater crayfishes estimated from nuclear and mitochondrial DNA sequences. Proc. R. Soc. Lond. Ser. B 267, 1679–1686.
- Crease, T.J., Colbourne, J., 1998. Unusually long small subunit ribosomal RNA of the crustacean, *Daphnia pulex*: sequence and predicted secondary structure. J. Mol. Evol. 46, 307–313.
- Crease, T.J., Taylor, D.J., 1998. The origin and evolution of variable-region helices in V4 and V7 of the small-subunit ribosomal RNA of branchiopod crustaceans. Mol. Biol. Evol. 15, 1430–1446.
- Cristescu, M.E.A., Hebert, P.D.N., 2002. Phylogeny and adaptive radiation in the Onychopoda (Crustacea: Cladocera): evidence from multiple gene sequences. J. Evol. Biol. 15, 838–849.
- Cummings, M.P., Otto, S.P., Wakeley, J., 1995. Sampling properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. 12, 814–822.
- Cunningham, C.W., 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy. Empirically testing an interative procedure for choosing among phylogenetic methods. Syst. Biol. 46, 464–478.
- De Rijk, P., Wuyts, J., Van de Peer, Y., Winkelmans, T., De Wachter, R., 2000. The European large subunit ribosomal RNA database. Nucleic Acids Res. 28, 177–178.
- Dumont, H.J., Silva-Briano, M., 1998. A reclassification of the anomopod families Macrothricidae and Chydoridae, with the creation of a new suborder, the Radopoda (Crustacea: Branchiopoda). Hydrobiology 384, 119–149.
- Eriksson, S., 1934. Studien Über die Fangapparate der Branchiopoden nebst einigen phylogenetischen Bemerkungen. Zool. Bidrag Upps. 15, 23–287.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
- Folmer, O., Black, M., Hoeh, R., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3, 294–299.
- Fryer, G., 1995. Phylogeny and adaptive radiation within the Anomopoda: a preliminary exploration. Hydrobiology 307, 57–68.
- Giribet, G., Edgecombe, G.D., Wheeler, W.C., 2001. Arthropod phylogeny based on eight molecular loci and morphology. Nature 413, 157–161.
- Graybeal, A., 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47, 9–17.
- Hanner, R., Fugate, M., 1997. Branchiopod phylogenetic reconstruction from 12S rDNA sequence data. J. Crust. Biol. 17, 174–183.
- Hillis, D.M., Huelsenbeck, J.P., 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. 83, 189–195.
- Huelsenbeck, J.P., Crandall, K.A., 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. 28, 437–466.
- Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES. Bayesian inference of phylogeny. Bioinformatics 17, 754–755.
- Källersjö, M., Farris, J.S., Kluge, A.G., Bult, C., 1992. Skewness and permutation. Cladistics 8, 275–287.
- Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

- Kluge, A.J., 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38, 7–25.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.
- Lyons-Weiler, J., Hoelzer, G.A., Tausch, R.J., 1996. Relative apparent synapomorphy analysis (RASA) I: the statistical measurement of phylogenetic signal. Mol. Biol. Evol. 13, 749–757.
- Martin, J.W., Davis, G.E., 2001. An updated classification of the recent Crustacea. Natural History Museum of Los Angeles County Science Series vol. 39, 1–164.
- Mindell, D.P., Sorenson, M.D., Huddleston, C.J., Miranda Jr., H.C., Knight, A., Sawchuk, S.J., Yuri, T., 1997. Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. In: Mindell, D.P. (Ed.), Avian Molecular Evolution and Systematics. Academic Press, San Diego, pp. 213–247.
- Negrea, S., Botnariuc, N., Dumont, H.J., 1999. Phylogeny, evolution and classification of the Branchiopoda (Crustacea). Hydrobiology 412, 191–212.
- Olesen, J., 1998. A phylogenetic analysis of the Conchostraca and Cladocera (Crustacea, Branchiopoda, Diplostraca). Zool. J. Linn. Soc. 122, 491–536.
- Olesen, J., 1999. Larval and post-larval development of the branchiopod clam shrimp *Cyclestheria hislopi* (Baird, 1859) (Crustacea, Branchiopoda, Conchostraca, Spinicaudata). Acta Zool. 80, 163–184.
- Olesen, J., 2000. An updated phylogeny of the Conchostraca–Cladocera clade (Branchiopoda, Diplostraca). Crustaceana 73, 869–886.
- Olesen, J., Martin, J.W., Roessler, E.W., 1997. External morphology of the male of *Cyclestheria hislopi* (Baird 1859) (Crustacea, Branchiopoda, Spinicaudata), with comparison of male claspers among the Conchostraca and Cladocera and its bearing on phylogeny of the 'bivalved' Branchiopoda. Zool. Scr. 25, 291–316.
- Olesen, J., Richter, S., Scholtz, G., 2003. On the ontogeny of *Leptodora kindtii* (Crustacea, Branchiopoda, Cladocera), with notes on the phylogeny of the Cladocera. J. Morph. 256, 235–259.
- Palumbi, S.R., 1996. Nucleic acids II: the polymerase chain reaction. In: Hillis, D., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer, Sunderland, MA, pp. 205–247.
- Poe, S., 1998. The effect of taxonomic sampling on accuracy of phylogeny estimation: test case of a known phylogeny. Mol. Biol. Evol. 15, 1086– 1090.
- Pollock, D.D., Zwickl, D.J., McGuire, J.A., Hillis, D.M., 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51, 664–671.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
- Remigio, E.A., Hebert, P.D.N., 2000. Affinities among anostracan (Crustacea: Branchiopoda) families inferred from phylogenetic analyses of multiple gene sequences. Mol. Phylogenet. Evol. 17, 117–128.
- Remsen, J., DeSalle, R., 1998. Character congruence of multiple data partitions and the origin of the Hawaiian Drosophilidae. Mol. Phylogenet. Evol. 9, 225–235.
- Richter, S., Braband, A., Aladin, N., Scholtz, G., 2001. The phylogenetic relationships of predatory water-fleas (Cladocera: Onychopoda, Haplopoda) inferred from 12S rDNA. Mol. Phylogenet. Evol. 19, 105– 113.
- Rokas, A., Carroll, S.B., 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22, 1337–1344.
- Rosenberg, M.S., Kumar, S., 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natl. Acad. Sci. USA 98, 10751–10756.
- Sacherova, V., Hebert, P.D.N., 2003. The evolutionary history of the Chydoridae (Crustacea: Chydoridae). Biol. J. Linn. Soc. 79, 629–643.
- Sanders, H.L., 1963. Significance of the Cephalocarida. In: Whittington, H.B., Rolfe, W.D.I. (Eds.), Phylogeny and Evolution of Crustacea. Museum of Comparative Zoology, Cambridge, MA, pp. 163–175.

Schram, F.R., 1986. Crustacea. Oxford University Press, Oxford.

- Schwenk, K., Sand, A., Boersma, M., Brehm, M., Mader, E., Offerhaus, D., Spaak, P., 1998. Genetic markers, genealogies and biogeographic patterns in the Cladocera. Aquat. Ecol. 32, 37–51.
- Simmons, M.P., Randle, C.P., Freudenstein, J.V., Wenzel, J.W., 2002. Limitations of relative apparent synapomorphy analysis (RASA) for measuring phylogenetic signal. Mol. Biol. Evol. 19, 14–23.
- Spears, T., Abele, L.G., 2000. Branchiopod monophyly and interordinal phylogeny inferred from 18S ribosomal DNA. J. Crust. Biol. 20, 1–24.
- Swain, T.D., Taylor, D.J., 2003. Structural rRNA characters support monophyly of raptorial limbs and paraphyly of limb specialization in water fleas. Proc. R. Soc. Lond. B Biol. Sci. 270, 887–896.
- Swofford, D.L., 2001. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), version 4. Sinauer Associates, Sunderland, MA.
- Taylor, D.J., Finston, T.L., Hebert, P.D.N., 1998. Biogeography of a widespread freshwater crustacean: pseudocongruence and cryptic

endemism in the North American *Daphnia laevis* complex. Evolution 52, 1648–1670.

- Taylor, D.J., Crease, T.J., Brown, W.M., 1999. Phylogenetic evidence for a single long-lived clade of crustacean cyclic parthenogens and its implications for the evolution of sex. Proc. R. Soc. Lond. B Biol. Sci. 266, 791–797.
- Van de Peer, Y., De Rijk, P., Wuyts, J., Winkelmans, T., De Wachter, R., 2000. The European small subunit ribosomal RNA database. Nucleic Acids Res. 28, 175–176.
- Walossek, D., 1993. The Upper Cambrian *Rehbachiella* and the phylogeny of Branchiopoda and Crustacea. Fossils Strata 32, 1–202.
- Walossek, D., 1995. The Upper Cambrian *Rehbachiella*, its larval development, morphology and significance for the phylogeny of Branchiopoda and Crustacea. Hydrobiology 298, 1–13.
- Wingstrand, K.G., 1978. Comparative spermatology of the Crustacea Entomostraca. I. Subclass Branchiopoda. Kolgl. Danske. Vid Selskab Biol. Skr. 22, 1–66.