BRILL

JOURNAL OF CRUSTACEAN BIOLOGY, 35(5), 577-592, 2015

Journal of Crustacean Biology

THE FOREST FOR THE TREES: EVALUATING MOLECULAR PHYLOGENIES WITH AN EMPHASIS ON HIGHER-LEVEL DECAPODA

Laura Timm * and Heather D. Bracken-Grissom

Florida International University, 3000 NE 151st Street, North Miami, FL 33181, USA

ABSTRACT

Since the late 1800s, several infraordinal relationships have been proposed for Decapoda; however, reaching a consensus among higher-level relationships is proving difficult. Molecular methods were first applied to higher-level decapod phylogenetics in the 1990s and have significantly contributed to our understanding of the group: sampling is becoming more thorough, a greater number of phylogenetically informative characters are being sequenced, and analysis procedures are becoming more consistent between studies. However, relationships among the deep lineages of Decapoda remain unclear. Several phylogenetic hypotheses have been suggested, and while there is some agreement among studies, an ultimate consensus among higher-level relationships has yet to be reached. This is largely the result of differences in sampling effort, marker selection, data-recycling, and analysis. Because most studies have generated conflicting phylogenetic hypotheses, the foundation on which the trees were built (data and analysis procedures) must be considered and evaluated. In this review, we summarize the early morphological decapod studies, address common problems that are causing a lack of consensus in molecular studies, present a means of evaluating molecular trees, offer suggestions for good phylogenetic practice, review the previous molecular studies of infraordinal decapod phylogeny, and discuss the future directions of the field, with special attention paid to next-generation sequencing (NGS) techniques.

KEY WORDS: data-recycling, Decapoda, infraorder, insufficient sampling, marker selection, molecular phylogeny, next-generation sequencing, out-group selection

DOI: 10.1163/1937240X-00002371

INTRODUCTION: ENTER, THE DECAPODA

Decapoda is an immense order, containing about 15,000 extant and about 3000 extinct species, including crabs, lobsters, hermit crabs, crayfish, and shrimp (De Grave et al., 2009). The order contains a morphologically diverse group of organisms inferred to have originated in the early Cambrian, possibly earlier (Martin and Davis, 2001; Bracken-Grissom et al., 2013, 2014), and evolving over approximately 400 million years to colonize and exploit almost every aquatic habitat on Earth. This evolutionary experimentation has resulted in perhaps the greatest diversity in body plan, size, and habitat preference (Monod and Laubier, 1996) present in any group of crustaceans (Martin and Davis, 2001; Bracken-Grissom et al., 2013). Because of this diversity, the "propinquity of descent" (Darwin, 1859) within Decapoda is obscured.

Carcinologists continue their search for what Schram (2001) described as the "Holy Grail": To arrive at a phylogeny that recapitulates the classification system and vice versa. Many approaches have been used to determine the origin and evolution of decapod infraorders, and morphological methods based on similarity and cladistics have generated a variety of trees (Calman, 1904; Siewing, 1963; Schram, 1986; Scholtz and Richter, 1995; Dixon et al., 2003). The 1990s saw the dawn of molecular phylogenetics for Decapoda; researchers began to use genetic sequence data to infer evolutionary relationships among major lineages (Kim and Abele, 1990; Abele, 1991). Molecular studies have advanced our understanding of Decapoda, but have not yet led to a consensus. Marker selection, realized sampling effort, data-recycling, and analysis ambiguities have contributed to a lack of resolution and confusion over what constitutes a reliable phylogeny. Next-Generation Sequencing (NGS) methods, such as Targeted Amplicon Sequencing (TAS) and Anchored Hybrid Enrichment (AHE), have the potential to provide new, genome-wide perspectives on the evolution of decapods (Bybee et al., 2011b; Qian et al., 2011; Lemmon et al., 2012; Shen et al., 2013). The current NGS methods are bringing us closer to an inclusively hierarchical phylogeny, which will provide evolutionary insight into decapod biogeography, biodiversity, ecology, character evolution, reproduction, and development.

The aims of this review are to: 1) briefly summarize the morphological studies of decapods; 2) identify common analysis problems that can cause a lack of consensus; 3) present a means of evaluating the strengths and weaknesses of molecular phylogenies; 4) offer suggestions for good phylogenetic practice; 5) review the literature on higher-level decapod molecular phylogenies while evaluating them as described; and 6) discuss the future directions of decapod phylogeny with specific focus on next-generation sequencing methods.

* Corresponding author; e-mail: ltimm004@fiu.edu

Part I: a Brief History of Decapod Classification and Phylogeny as Determined by Morphology

Efforts to classify decapods began in the 1800s and resulted in two schemes of division. Milne Edwards (1834) and Boas (1880) proposed a phenetic division based on primary mode of locomotion: the benthic Reptantia and the swimming Natantia. Huxley (1878) divided the lobster and lobster-like taxa (presently recognized as Achelata, Astacidea, Axiidea, Gebiidea, and Polychelida) into two groups based on gill and branchiostegite morphology: Trichobranchiata and Phyllobranchiata. At the turn of the century, Boas' system was still recognized. In a much-cited publication, Borradaile (1907) retained the Reptantia-Natantia subgroups, but revised the taxa comprising each. However, neither the Reptantia-Natantia classification system nor the Trichobranchiata-Phyllobranchiata classification system had been devised to include many fossil representatives. A study by Beurlen and Glaessner (1930), which included data from taxa represented only in the fossil record, proposed a new system. To accommodate the fossilized taxa, Trichelida and Heterochelida were introduced as the suborders within Decapoda. For the next three decades, studies focused primarily on elucidating the lower-level divisions of families and genera.

Burkenroad (1963) proposed a major restructuring of the higher-level taxonomy of Decapoda. Investigating the gill morphology evident in the eumalacostracan fossil record, he concluded that all previously proposed classification systems exhibited some degree of polyphyly. Noting "peneids" (a name used by Burkenroad to refer to non-brooding shrimp) as one of two major branches within Decapoda, he proposed Dendrobranchiata as a suborder to include this group. The second major group Burkenroad proposed, which contained the majority of decapod infraorders, was Pleocyemata. These two groups were divided primarily by gill morphology and brooding behavior. Carcinologists have long accepted the Dendrobranchiata-Pleocyemata division, and while Natantia is no longer recognized, Reptantia still serves as an unranked group containing crawling/walking lineages (see Boas, 1880 for full definition) (Table A1 in the online edition of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1937240x).

Several approaches have attempted to further elucidate relationships within Pleocyemata: adult morphology (Martin and Davis, 2001), larval morphology (Clark, 2009), spermiocladistics (Martin and Davis, 2001), eye morphology (Porter and Cronin, 2009), ontogeny (Martin and Davis, 2001), and parasite proxies (Boyko and Williams, 2009), to name a few. As early as the 1970s, molecular methods made thousands of characters available for analysis. Since then, molecular phylogenetic analyses have proven informative at many levels of decapod phylogeny, while also uncovering new areas of investigation.

Part II: the Dawn of Molecular Methods and Evaluating the Forest of Trees

As molecular methods were adapted to elucidate decapod phylogeny, many studies proposed different evolutionary hypotheses (Fig. 1). This conflict requires standards by which phylogenies can be evaluated. The field of decapod phylogenetics, along with many other groups, is frequently subject to several potential pitfalls in study design and analysis. These pitfalls, resulting from variability or ambiguity in procedure or analysis, are often overlooked, but are very important to the strength and reliability of results. Here, we identify four such ambiguities: marker selection, realized sampling effort, data-recycling, and analysis ambiguity; and offer suggestions to navigate them.

Markers: Inappropriate or Insufficient

The traditional molecular approaches, and some NGS methods, require the selection of genetic markers targeted and sequenced from representative taxa. Markers can originate from the mitochondrial genome or from the nuclear genome (mtDNA and nDNA, respectively; see Table 1). Both mtDNA and nDNA have advantages and disadvantages that are nontrivial.

Advantages of mtDNA.—Mitochondrial DNA generally mutates faster than nDNA (Brown et al., 1979), making mtDNA markers most informative at lower taxonomic levels, e.g., genus and species (Moore, 1995). These markers are relatively easy to amplify, as universal primers are available for many taxa (Simon et al., 1994) and encoded genes are strictly orthologous (Qian et al., 2011). Because mtDNA is haploid, recombination is rare (Birky, 2001; Elson and Lightowlers, 2006). Using full mitochondrial genomes (mt-genomes) has gained some popularity in studies of deep-level phylogeny (Fenn et al., 2008), such as in Insecta (Talavera and Vila, 2011), because nucleotide sequence, gene order (Boore and Brown, 1998), gene insertion and deletion (Rokas and Holland, 2000), and length variability (Schneider and Ebert, 2004) can provide phylogenetic information. Some argue these properties make the mt-genome one of the most information-rich markers in phylogeny (Fenn et al., 2008). However these approaches have been subject to criticism (Ballard and Whitlock, 2004; Ballard and Rand, 2005; Hurst and Jiggins, 2005; Galtier et al., 2009).

Disadvantages of mtDNA.—Mitochondrial markers are not suited to every study and there are important characteristics that must be considered. First, the increased mutation rate in mtDNA decreases time to saturation (Blouin et al., 1998), limiting the phylogenetic signal at higher taxonomic levels. Second, mtDNA is subject to mitochondrial capture, meaning introgression events in the recent past can obscure true phylogenetic relationships among close relatives (Shaw, 2002; Ballard and Whitlock, 2004; Spinks and Shaffer, 2009). Third, mtDNA is highly subject to site linkage as it does not undergo recombination (Birky, 1995; Avise, 2000; Ballard and Whitlock, 2004). The final characteristic, and perhaps the most contentious, is that mtDNA markers may violate the assumption of marker neutrality: the non-recombining maternal inheritance mechanism can be prone to genetic hitchhiking, fixing new alleles faster than nDNA (Brown et al., 1979; Bazin et al., 2006; Meiklejohn et al., 2007). Additionally, several studies have indicated that mitochondria can be subject to direct and indirect selection, further confounding the assumption of neutral evolution (Ballard and Whitlock, 2004; Ballard and Rand, 2005; Hurst and Jiggins, 2005; Galtier et al., 2009). Due to the

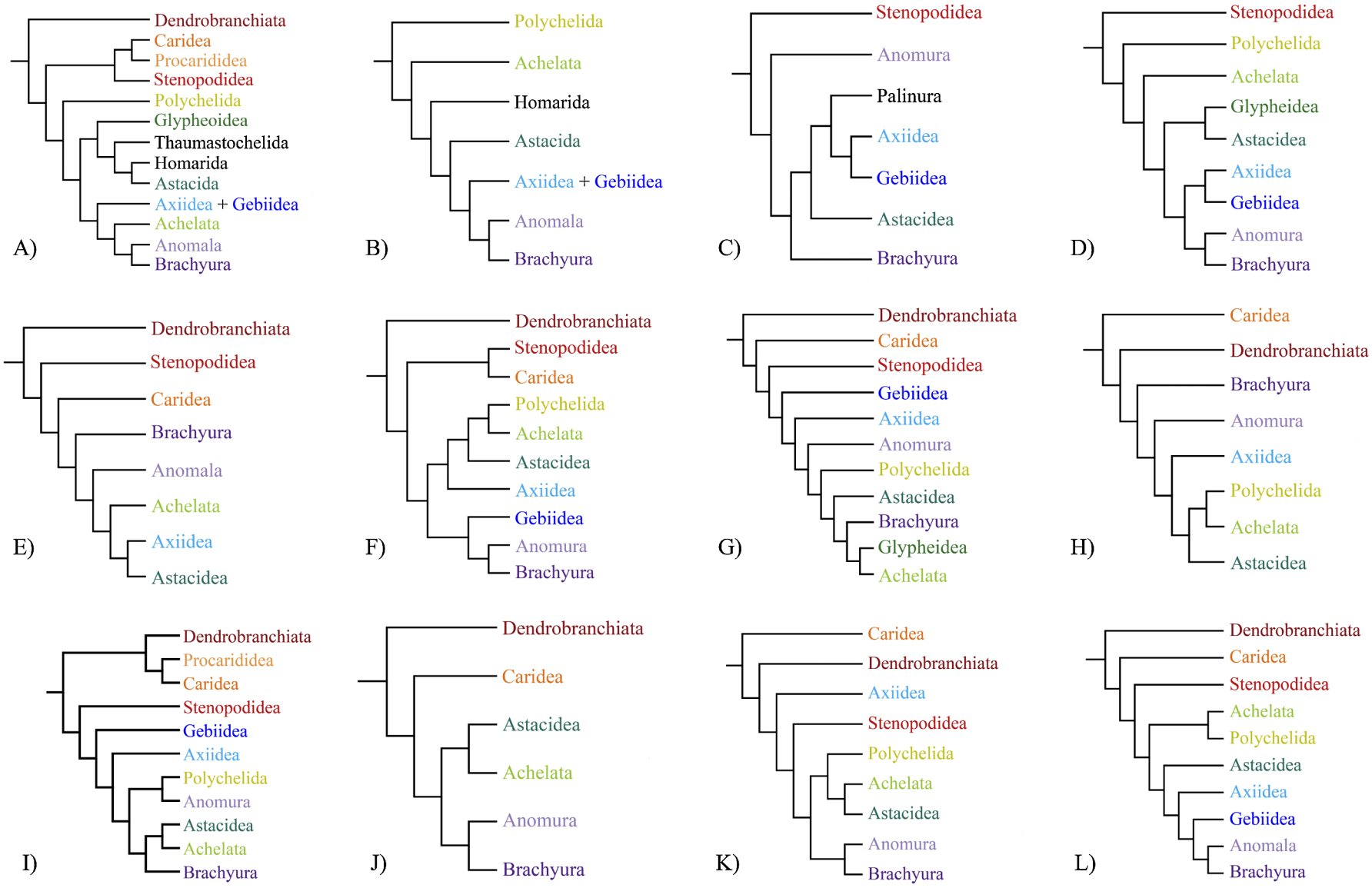

Fig. 1. Phylogenetic trees of infraordinal decapod phylogeny including: A, Dixon et al. (2003) morphological analysis; B, Scholtz and Richter (1995) meta-analysis. Major molecular studies include: C, Crandall et al. (2000) analysis of 18S; D, Ahyong and O'Meally (2004) analysis of 16S, 18S, 28S, and morphological characters; E, Porter et al. (2005) analysis of 16S, 18S, 28S, and H3; F, Tsang et al. (2008b) analysis of PEPCK and NaK; G, Bracken et al. (2009a) analysis of 16S, 18S, 28S, and H3; H, Toon et al. (2009) analysis of 12S, 16S, 18S, 28S, H3, EF-2, EPRS, and TM9sf4; I, Bracken et al. (2010) analysis of 16S, 18S, 28S, and H3; J, Qian et al. (2011) analysis of whole mt-genome; K, Bybee et al. (2011b) analysis of 12S, 16S, 18S, 28S, H3, and COI; L, Shen et al. (2013) analysis of whole mt-genome. This figure is published in colour in the online edition of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1937240x.

Table 1. The gene markers and out-group(s) used in higher-level (Infraorder) decapod phylogeny studies to date.

Publication	Genes used	Gene origin	Gene function	Infraorders not included
Kim and Abele (1990)	18S	nDNA	Ribosomal subunit	Achelata, Anomura, Axiidea, Gebiidea, Glypheidea, Polychelida
Crandall et al. (2000)	16S	mtDNA	Ribosomal subunit	Caridea, Glypheidea, Polychelida, Procarididea
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
Ahyong and O'Meally (2004)	16S	mtDNA	Ribosomal subunit	Caridea, Procarididea
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
Porter et al. (2005)	16S	mtDNA	Ribosomal subunit	Gebiidea, Glypheidea, Polychelida, Procarididea
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
	H3	nDNA	Protein-coding	
Tsang et al. (2008b)	PEPCK	nDNA	Protein-coding	Glypheidea, Procarididea
	NaK	nDNA	Protein-coding	
Chu et al. (2009)	PEPCK	nDNA	Protein-coding	Glypheidea, Procarididea
	NaK	nDNA	Protein-coding	
Toon et al. (2009)	12S	mtDNA	Ribosomal subunit	Gebiidea, Glypheidea, Procarididea, Stenopodidea
	16S	mtDNA	Ribosomal subunit	
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal sununit	
	H3	nDNA	Protein-coding	
	EF-2	nDNA	Protein-coding	
	EPRS	nDNA	Protein-coding	
	TM9sf4	nDNA	Protein-coding	
Bracken et al. (2009a)	16S	mtDNA	Ribosomal subunit	Procarididea
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
	H3	nDNA	Protein-coding	
Bracken et al. (2010)	16S	mtDNA	Ribosomal subunit	Glypheidea
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
	H3	nDNA	Protein-coding	
Bybee et al. (2011a)	12S	mtDNA	Ribosomal subunit	Gebiidea, Glypheidea, Procarididea
	16S	mtDNA	Ribosomal subunit	
	18S	nDNA	Ribosomal subunit	
	28S	nDNA	Ribosomal subunit	
	H3	nDNA	Protein-coding	
	COI	mtDNA	Protein-coding	
Qian et al. (2011) (whole mt-genome)	cox1	mtDNA	Protein-coding	Axiidea, Gebiidea, Glypheidea, Polychelida, Procarididea, Stenopodidea
	cox2	mtDNA	Protein-coding	
	cox3	mtDNA	Protein-coding	
	nad1	mtDNA	Protein-coding	
	nad2	mtDNA	Protein-coding	
	nad3	mtDNA	Protein-coding	
	nad4	mtDNA	Protein-coding	
	nad4L	mtDNA	Protein-coding	
	nad5	mtDNA	Protein-coding	
	nad6	mtDNA	Protein-coding	
	atp6	mtDNA	Protein-coding	
	atp8	mtDNA	Protein-coding	
	cob	mtDNA	Protein-coding	
	rrnS	mtDNA	Ribosomal subunit	
	rrnL	mtDNA	Ribosomal subunit	
	A	mtDNA	tRNA	

Table 1. (Continued.)

Publication	Genes used	Gene origin	Gene function	Infraorders not included
	R	mtDNA	tRNA	
	N	mtDNA	tRNA	
	D	mtDNA	tRNA	
	C	mtDNA	tRNA	
	E	mtDNA	tRNA	
	Q	mtDNA	tRNA	
	G	mtDNA	tRNA	
	H	mtDNA	tRNA	
	I	mtDNA	tRNA	
	L1	mtDNA	tRNA	
	L2	mtDNA	tRNA	
	K	mtDNA	tRNA	
	M	mtDNA	tRNA	
	F	mtDNA	tRNA	
	P	mtDNA	tRNA	
	S1	mtDNA	tRNA	
	S2	mtDNA	tRNA	
	T	mtDNA	tRNA	
	Y	mtDNA	tRNA	
	W	mtDNA	tRNA	
	V	mtDNA	tRNA	
	nCR	mtDNA	intron	
Shen et al. (2013) (whole mt-genome)	cox1	mtDNA	Protein-coding	Glypheidea, Procarididea
	cox2	mtDNA	Protein-coding	
	cox3	mtDNA	Protein-coding	
	nad1	mtDNA	Protein-coding	
	nad2	mtDNA	Protein-coding	
	nad3	mtDNA	Protein-coding	
	nad4	mtDNA	Protein-coding	
	nad4L	mtDNA	Protein-coding	
	nad5	mtDNA	Protein-coding	
	nad6	mtDNA	Protein-coding	
	atp6	mtDNA	Protein-coding	
	atp8	mtDNA	Protein-coding	
	cob	mtDNA	Protein-coding	
	rrnS	mtDNA	Ribosomal subunit	
	rrnL	mtDNA	Ribosomal subunit	
	A	mtDNA	tRNA	
	R	mtDNA	tRNA	
	N	mtDNA	tRNA	
	D	mtDNA	tRNA	
	C	mtDNA	tRNA	
	E	mtDNA	tRNA	
	Q	mtDNA	tRNA	
	G	mtDNA	tRNA	
	H	mtDNA	tRNA	
	I	mtDNA	tRNA	
	L1	mtDNA	tRNA	
	L2	mtDNA	tRNA	
	K	mtDNA	tRNA	
	M	mtDNA	tRNA	
	F	mtDNA	tRNA	
	P	mtDNA	tRNA	
	S1	mtDNA	tRNA	
	S2	mtDNA	tRNA	
	T	mtDNA	tRNA	
	Y	mtDNA	tRNA	
	W	mtDNA	tRNA	
	V	mtDNA	tRNA	
	nCR	mtDNA	Intron	

inheritance mechanism and lack of recombination, it has been argued that the mt-genome should be considered a single marker (Fenn et al., 2008). Moreover, the presence of nuclear copies of mtDNA (NUMTs) can confound analyses based on mtDNA (Zhang and Hewitt, 1996). Used by themselves, mtDNA markers, even mt-genomes, can be inappropriate for studies of deeper relationships, such as those among families and infraorders.

Advantages of nDNA.—Nuclear markers can provide information on taxonomic relationships from species to order, although they are often used to resolve higher-level divergences (Baldwin et al., 1995; Friedrich and Tautz, 1995; Rokas et al., 2003; Chu et al., 2009; Robles et al., 2009). This is due to variable rates of evolution in nDNA, especially among protein-coding genes, ribosomal DNA, and introns. Protein-coding genes tend to be more conserved than other nDNA, as mutations that result in loss of protein function are subject to strong negative selection (Opperdoes, 2009). Ribosomal DNA (rDNA) tends to have highly conserved enzymatic regions and highly variable regions of expansion (Kim and Abele, 1990). Introns tend to be less conserved as they are unconstrained by protein production (Bell et al., 1998; Yeo et al., 2005; Kim and Kim, 2007).

Disadvantages of nDNA.—Aligning nDNA may be complicated by heterozygosity, multiple insertions and deletions, or by the presence of introns (Gatesy et al., 1993; Sota and Vogler, 2003; Tsang et al., 2008b; Chu et al., 2009). Also, nDNA can be more difficult to amplify, as it is typically present in fewer copies in each cell, relative to mtDNA (Zhang and Hewitt, 2003; Chu et al., 2009). This is especially true for protein-coding genes. Due to the relatively slower mutation rate characteristic of nDNA markers, nDNA is often inappropriate for studies of lower-level relationships, such as at the species- and genus-level. A final concern, which has gained appreciation over the past twenty years (Koonin, 2005), is the potential presence and effects of paralogs. Paralogous genes are versions of a gene that arose from a gene duplication event (Fitch, 1970). These copies may be under different selection pressures because they are present as more than one copy within an individual (Kondrashov et al., 2002), although recent studies argue that this is not always the case (Studer and Robinson-Rechavi, 2009). Phylogenies are traditionally constructed using orthologous genes; that is, gene variants that arise from an ancestral gene that has undergone a speciation event (Fitch, 1970). These gene copies are believed to share important properties, such as function, that result in identical evolutionary rates (Baldauf, 2003); though this assumption is also being debated (Gabaldón and Koonin, 2013). As NGS methods have become more widely used, the ability to identify paralogs and estimate their effects is becoming increasingly important (Koonin, 2005).

Suggestions.—Because of the innate properties associated with mtDNA and nDNA, markers used to elucidate phylogenetic relationships must be chosen with the goal of the study in mind: targeted markers must be able to resolve at the taxonomic level of interest. Choice of marker can be a trade-off: low copy-number nDNA (protein-coding genes) markers may be difficult to amplify, but more easily amplified mtDNA markers are not always informative at the necessary taxonomic levels. Thus, phylogenetic studies can be strengthened by including multiple informative markers, including protein-coding genes, mtDNA, and rDNA to inform at several levels. In the decapod literature, this is implemented by Palero and Crandall (2009), Bybee et al. (2011b), Bracken-Grissom et al. (2013, 2014), and Wong et al. (2015). Currently, NGS phylogenomics methods are enabling the discovery and utilization of an unprecedented number of markers (more than 500 in a single study), informative across a range of taxonomic levels (Lemmon et al., 2012).

Species Trees vs. Gene Trees.—The goal of phylogenetic studies is a species tree. That is, a tree that reflects the evolutionary history of species. This is accomplished by reconciling the evolutionary histories of individual genetic markers to arrive at a tree that recapitulates relationships between species (Page and Charleston, 1997). Building trees with multiple, informative markers prevents the recapitulation of single-gene trees (Fig. 2), which are often inappropriate for phylogenetic studies. Individual genes can have their own unique evolutionary histories that differ from the evolutionary histories of the species and other genes (Page and Charleston, 1997). Gene trees can differ from species trees in two ways: 1) the divergence of two alleles may have occurred before the divergence of the species, and 2) the gene tree and species tree may present different topologies (Graur and Li, 2000). Thus, analysis of a single gene recapitulates the gene's evolutionary history, and often cannot reliably inform the true species tree (Pamilo and Nei, 1988; Doyle, 1992; Page and Charleston, 1997; Degnan and Rosenberg, 2009). Indeed, a simulation study by Gadagkar et al. (2005) found that adding one gene to a single-gene analysis increases accuracy of phylogenetic inference by approximately 10%, even when the added gene is less phylogenetically informative than the first. Individually, single-gene markers are insufficient, so a variety of markers should be used to inform at the level of interest (Pamilo and Nei, 1988; Doyle, 1992; Maddison, 1997).

Insufficient Sampling and Out-Group Selection

Adequate sampling is key to reliably recapitulating phylogeny (Wiens, 2003; Maddison and Knowles, 2006). Insufficient sampling can result in long-branch attraction, false results of monophyly, and incorrect out-group rooting. All of the shortcomings associated with insufficient sampling can be curtailed by tailoring sampling effort to the purpose of the study.

Monophyly, Paraphyly, and Polyphyly.—Without adequate representation within the taxonomic level of interest, monophyly can be incorrectly inferred, resulting in subsequent discovery of paraphyly or polyphyly. This was the case for the decapod infraorder Thalassinidea, which was long perceived as monophyletic (Crandall et al., 2000; Ahyong and O'Meally, 2004; Porter et al., 2005) but only later found to be polyphyletic with additional sampling (Tsang et al., 2008b; Bracken et al., 2009a). Thalassinidea has since been divided into Axiidea and Gebiidea – relatively distant infraorders. To best ensure reliable results, every group at

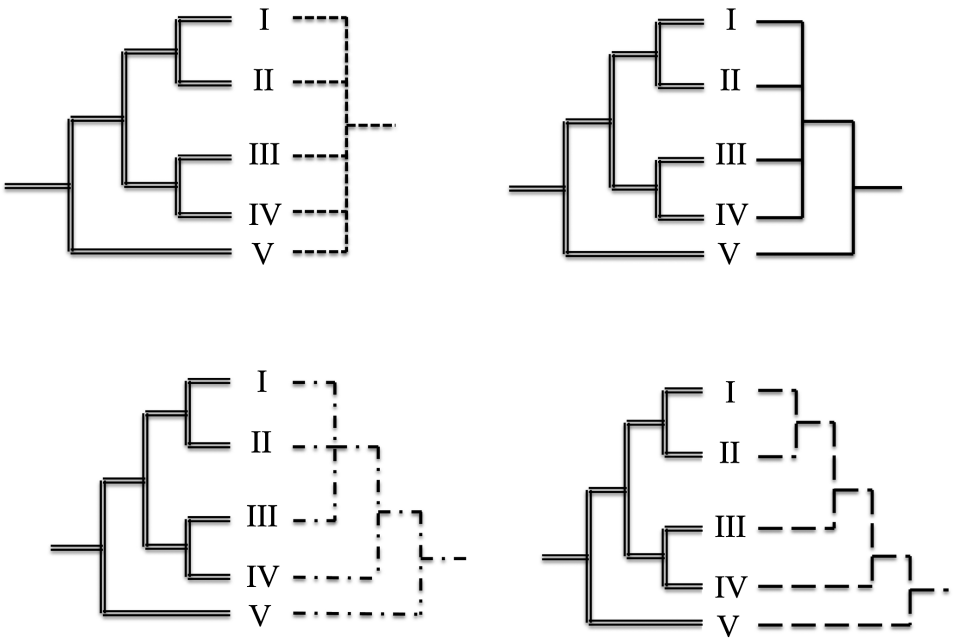

Fig. 2. An illustration of a species tree (depicted with double-lines) compared to four arbitrary single-gene trees. While the true species tree is always the same, the gene trees recapitulate different relationships.

the level of interest should be sampled as broadly as possible. For instance, if one is inferring infraordinal relationships, multiple species within each infraorder should be represented across diverse and divergent lineages. A good example of this is Ahyong et al. (2007) which reconstructs brachyuran phylogeny, and indicates paraphyly of podotremes (also supported by Tsang et al., 2014), by thoroughly sampling sections and families within the infraorder. Frequently this is not possible due to a number of factors. If this is the case, authors should address this in the publication and provide justification for the missing lineages (Valentine et al., 2006).

Long-Branch Attraction.—One of the most confounding results of insufficient sampling is the increased likelihood of long-branch attraction (Bergsten, 2005; Fig. 3), especially in maximum parsimony analysis (Vandamme, 2009). Long-branch attraction (LBA) occurs when taxa are so divergent that mutations begin to be shared due to convergence rather than homology (Felsenstein, 1978). This convergence results in highly dissimilar taxa, which would normally be grouped on separate long branches, being "attracted" onto a single long branch. This problem should be fairly easy to identify, given some background knowledge of the lineage. Sampling more basal representatives from each clade can prevent long-branch attraction by breaking up these groups (Felsenstein, 1978; Zwickl and Hillis, 2002; Yang and Rannala, 2012).

Out-Group Selection.—The final problem of insufficient sampling is improper out-group selection. This subject can, and has, occupied several papers, exclusively. We will discuss it briefly here. Without an accepted common ancestor, polarity assignment of traits is confounded (Throckmorton, 1968; Farris et al., 1970; Lundberg, 1972; de Queiroz and Gauthier, 1990; Wiley et al., 1991) and the selection of an out-group is obscured (Wheeler, 1990). Choosing an out-group that is too distantly related may lead to spurious rooting owing to loss of phylogenetic signal resulting from saturation. However, choosing an out-group that is too closely related can also skew analyses by aligning too closely with the taxon of interest, that is, by not serving as a "true" rooting group (Vandamme, 2009). However, to investigate the ancient relationships within Decapoda, out-group rooting is optimized by rooting with the sister group.

In instances where the "best" out-group is difficult to identify, it is advisable to choose several: study the literature

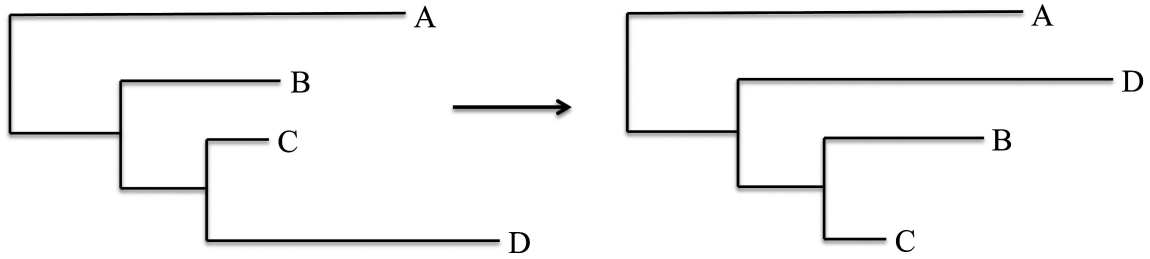

Fig. 3. An illustration of long-branch attraction (LBA), in which distantly related taxa have accrued so many differences that they cluster together. In this figure, species A and D are truly distantly related (left tree), but cluster together due to LBA (right tree). Figure adapted from Forterre and Philippe (1999).

of the group of interest and find what taxa have been used in the previous studies. Since the study of decapod phylogeny began, several taxa have been proposed as the sister group: Siewing (1963), Schram (1986), Wills (1998), and Schram and Hof (1998; tree unresolved) found Euphausiacea to be sister to Decapoda. Schram (1981, 1984) made a case for a polyphyletic group containing both Amphionidacea Williamson, 1973 and Euphausiacea as the sister group. And a study by Richter and Scholtz (2001) identified the subclass Hoplocarida as the sister taxon. More recently, a study by Meland and Willassen (2007) resulted in polyphyly of Decapoda, indicating several sister groups. To overcome this problem, most phylogenetic analyses must include several out-groups when rooting the resulting trees. Most molecular studies use Euphausiacea and Hoplocarida Calman, 1904 as out-groups (Bracken et al., 2009a; Qian et al., 2011; Shen et al., 2013) but to date have not included Amphionidacea due to the lack of molecular-grade tissue for this group.

Data-Recycling

One practice meant to alleviate incomplete sampling is data-recycling, which includes previously published data in a new data matrix. In phylogenetic studies, both taxa and characters are recycled to add robustness to the study. Although data-recycling can have positive impacts on the resulting tree, the pitfalls of data-recycling must be addressed.

Advantages of Recycling.—In general, researchers use previously published sequence data to circumvent the need to resample groups or to bolster taxa with few newly collected representatives. This practice can be beneficial to many phylogenetic studies by allowing them to build upon previously published datasets, which can conserve time and resources. However, using data from several sources and several authors can introduce artifacts of sampling idiosyncrasies, resulting in confounded analyses (Jenner, 2001).

Disadvantages of Recycling.—Phenotypic data matrices compiled in previous studies are reused in derivative analyses, recycling taxa and characters, potentially resulting in the dissemination of flaws in an original matrix through several subsequent studies (Jenner and Schram, 1999; Poe and Wiens, 2000; Jenner, 2001; Dayrat and Tillier, 2002). Molecular studies can analyze the same markers that have been analyzed in previous studies, neglecting to sequence new markers. Or, new markers may be sequenced, but from previously sampled species. All of these practices can serve to reinforce prior assumptions.

Suggestions.—Phylogenetic studies that rely too heavily on recycled data typically generate the same topology, a potentially misleading result. Overall, data-recycling best serves studies when it supplements a study that generates and analyzes new characters in new representatives (Hillis et al., 2003; Bracken-Grissom et al., 2013, 2014). Also, it is a good practice to announce which data were recycled, either taxa or markers.

Inconsistent Analysis Procedures

In all phylogenetic analyses, the researcher is faced with dozens of parameter options and algorithms that could be used to estimate a phylogeny. Previous studies have shown that inputting the same dataset, but altering the model of evolution, the subsampling procedure, or the parameters can result in different trees (Buckley, 2002; Buckley and Cunningham, 2002; Lemmon and Moriarty, 2004). Careful thought must be given to these decisions. Below, we discuss four areas of concern and solutions gleaned from the literature.

Algorithm Selection.—Four algorithms commonly used in phylogenetic analysis are Neighbor Joining (NJ: Saitou and Nei, 1987), Maximum Parsimony (MP: Fitch, 1971), Maximum Likelihood (ML: Felsenstein, 1981), and Bayesian Inference (BI: Huelsenbeck and Ronquist, 2001). The robustness of results from ML and MP algorithms can be evaluated by the designation of a subsampling procedure, such as bootstrapping or jackknifing (Van de Peer, 2009). These subsampling procedures are used to generate branch support values by analyzing pseudo-replicates and calculating the percent of resulting trees containing each branch (Schmidt and von Haeseler, 2009). Bayesian Inference does not rely on subsampling, but rather calculates the posterior probability of every tree sampled from a distribution of all possible trees. Support values, then, are calculated as the percent of sampled trees that contain the nodes seen on the presented tree. For BI, branches with support values $\geqslant$95% are considered statistically well-supported. For ML and MP, $\geqslant$70% are considered statistically well-supported.

Nucleotide Substitution Model Selection.—One can only be as confident in a tree as one is in the model that built it (Goldman, 1993). BI, ML, and NJ require the specification of an evolutionary model. Models can be divided simply into those that assume all nucleotides occur with equal frequency (Jukes and Cantor, 1969; Kimura, 1980) and those that allow all nucleotides to occur at different frequencies (Felsenstein, 1981; Hasegawa et al., 1985; Tavaré, 1986). Some software programs, such as Random Axelerated Maximum Likelihood (RAxML: Stamatakis, 2006; Stamatakis et al., 2005, 2007, 2008), have the model set to GTR, which nests several models (Stamatakis, 2006). Programs, such as MODELTEST (Posada and Crandall, 1998) and jModelTest (Posada, 2008), are available to determine the optimal model based on the likelihood ratio and Akaike Information Criterion calculated over nested hierarchical analyses (Posada and Crandall, 1998). Currently, there is much research effort in model selection (Reid et al., 2013; Brown, 2014a, b; Lewis et al., 2014 are the most recent examples) and in determining whether current models appropriately fit the data.

Data Partitioning.—When analyzing data from multiple markers, it is often necessary to partition the data by substitution rate (Nishihara et al., 2007) or codon position (Yang, 1996). In total evidence studies, partitioning is crucial for datasets that include molecular markers and morphological characters, as seen in the phylogenetic reconstruction of lobsters and anomurans (Schnabel et al., 2011; Bracken-Grissom et al., 2013, 2014). Data that is not partitioned is subject to "mixture models," in which each marker is analyzed under multiple substitution models and every marker is assumed to have evolved under similar processes (Le et al., 2008). This can negatively impact tree topology (Buckley et al., 2001; Telford and Copley, 2011). Data can be

partitioned on the basis of codon position, gene (e.g., 16S, 12S, COI), gene origin (nuclear vs. mitochondrial), or gene function (protein coding vs. intron). By partitioning data, researchers can group markers that are likely to have experienced similar evolutionary processes, and then analyze each group independently. This allows for the reconstruction of a phylogeny that takes into account heterogeneous evolutionary histories (Lanfear et al., 2012). As with model selection, researchers can use programs such as PartitionFinder (Lanfear et al., 2012) to statistically explore and support partitioning schemes.

Application of Coalescent Theory.—In non-coalescent approaches, genes are concatenated into a 'supergene' alignment and traditional tree-building algorithms are applied to generate a phylogeny in a single step (often called "concatenation phylogenies;" Gadagkar et al., 2005; Edwards, 2009). This method has been criticized for failing to resolve the evolutionary history at the species level (Edwards, 2009). Rather, non-coalescent approaches estimate the genealogical history of individuals across a multilocus dataset, which is problematic when individual gene trees are in conflict due to mechanisms such as horizontal gene transfer, gene duplication, deep coalescence, and branch length heterogeneity (Edwards, 2009; Liu et al., 2009b). This approach has also been criticized for over-simplifying evolution and frequently ignoring gene tree heterogeneity by including too few markers (McVay and Carstens, 2013). Coalescent approaches use genetic data to calculate population parameters in an effort to better reflect the history of a taxon (Kingman, 2000; Edwards, 2009). This allows for gene tree heterogeneity, which enables correct species tree estimation, even in the anomaly zone where the most common gene tree does not match the species tree (Degnan and Rosenberg, 2009; Liu et al., 2009a). These analyses can be computationally demanding, and have been described as too complex, especially for long-diverged clades (McVay and Carstens, 2013). However, including variation in gene analysis has been found to be advantageous in theoretical multi-locus analyses (Kubatko and Degnan, 2007). In general, it is a good practice to analyze data using both approaches and present both trees in the publication.

PART III: A REVIEW OF HIGHER-LEVEL DECAPOD MOLECULAR PHYLOGENIES

From the first studies in the 1990s (Kim and Abele, 1990) to the next-generation studies of the 2010s (Bybee et al., 2011a, b; Qian et al., 2011; Shen et al., 2013), much progress has been made in resolving decapod phylogeny. Early studies identified informative markers, both molecular (Kim and Abele, 1990; Crandall et al., 2000; Porter et al., 2005; Tsang et al., 2008b; Chu et al., 2009) and morphological (Ahyong and O'Meally, 2004). These studies helped uncover polyphyly in Palinura (Ahyong and O'Meally, 2004) and Thalassinidea (Tsang et al., 2008b; Bracken et al., 2009a), informing phylogeneticists of which groups required more thorough and targeted sampling for phylogenetic reconstruction. The markers from these studies also served as the starting point for using NGS platforms such as targeted amplicon sequencing (Bybee et al., 2011b).

The first study by Kim and Abele (1990) sampled nine specimens spanning the suborder Dendrobranchiata and five infraorders: Astacidea, Brachyura, Caridea, Procarididea, and Stenopodidea (though the study did not recognize the now accepted division between Procarididea and Caridea and analyzed *Palaemonetes* and *Procaris* as part of Caridea). This study sought to determine whether the 18S ribosomal subunit could and/or would infer a phylogeny that accorded with morphology-based phylogenies. The MP analysis resulted in a significantly supported tree with sufficient variation between infraorders to conclude that 18S was phylogenetically informative at the infraordinal level. The first molecular study to propose a relationship between major decapod lineages, Kim and Abele identified a marker that is frequently used in higher-level decapod phylogenetic studies today. However, interpretation of these results is limited due to the incomplete sampling at the infraordinal level and insufficient marker selection. Nonetheless, this study was based entirely on *de novo* sequences.

Crandall et al. (2000) focused on the monophyletic origins of crayfish, but sampled sufficiently to generate a tree including several decapod infraorders. Analyzing 16S mtDNA, 18S, and 28S rDNA markers, this study included species from Achelata (Palinura in the study), Anomura, Astacidea, Axiidea, Brachyura, Gebiidea, and Stenopodidea (Axiidea and Gebiidea were listed as representatives of Thalassinidea in the study). Trees were estimated using NJ, ML, and MP. The resulting tree (Fig. 1C) generated a similar topology to that of Kim and Abele (1990). Despite a lack of data partitioning, this study provided evidence that utilizing multiple gene regions allowed for resolution at several taxonomic levels.

In 2004, the first decapod total evidence study (molecular + morphology) was performed using 16S, 18S, and 28S, as well as 105 morphological characters which included spermatozoan, gill, branchiostegite, rostrum, and carapace characteristics, among many others (Ahyong and O'Meally, 2004). Here, data-recycling was used to supplement newly generated morphological and molecular data. This study represented the most complete sampling of reptant decapod infraorders yet, including representatives from Achelata, Anomura, Astacidea, Axiidea, Brachyura, Gebiidea, Glypheidea, Polychelida, and Stenopodidea (Axiidea and Gebiidea were still recognized as Thalassinidea in this analysis). The study presented three slightly differing MP trees generated from morphological characters, molecular markers, and a combination of the two (total evidence). The total evidence tree (Fig. 1D) more closely resembled the relationships recovered in the molecular phylogeny, and all three analyses were congruent at the infraordinal level. The thorough sampling scheme helped uncover polyphyly within Palinura, resulting in its eventual division into Achelata, as the most basal of the three and sister to the fractosternalian infraorders; Polychelida, as sister to the remaining reptants; and Glypheidea, as sister to Astacidea. It should be noted that, while a partition was made between molecular and morphological data, the molecular data were unpartitioned, which may have negatively impacted the resulting topology.

Porter et al. (2005) included markers used in previous analyses (16S, 18S, 28S) but also included the histone 3 (H3) nDNA sequence for analysis. This study included representatives from Achelata, Anomura, Astacidea, Axiidea (listed as Thalassinidea), Brachyura, Caridea, Dendrobranchiata, and Stenopodidea; and was one of the first to partition data for analysis. Alignments were analyzed using ML and, for the first time in infraordinal decapod phylogenetic analysis, BI. The resulting tree unexpectedly placed Brachyura and Anomura in the middle of the tree (Fig. 1E), though these two groups traditionally fall out as more derived. Instead, Astacidea and Axiidea appeared more derived. The authors found all sampled infraorders to be monophyletic with strong branch support, but nodal support values for the relationships between infraorders were relatively low. This may be the result of insufficient taxon sampling and/or marker selection, that is, the markers were not sufficient in resolving deep relationships.

A study published in 2008 focused solely on protein-coding markers novel to decapod phylogeny: a sodium potassium pump (NaK) and phosphoenolpyruvate carboxykinase (PEPCK), thus all sequences analyzed in the study were generated *de novo* (Tsang et al., 2008b). Despite the absence of previously generated sequence data, the study included representatives from Achelata, Anomura, Astacidea, Axiidea, Brachyura, Caridea, Dendrobranchiata, Gebiidea, Polychelida, and Stenopodidea (Axiidea and Gebiidea were included as specimens of Thalassinidea). Data was analyzed with ML, MP, and BI, resulting in strongly supported monophyly for all infraorders, except Thalassinidea, which exhibited polyphyly (Fig. 1F). The authors suggested returning to the scheme of Gurney (1938), which divided Thalassinidea into the "Homarine Group" (Axiidea) and the "Anomuran Group" (Gebiidea). In 2009, NaK and PEPCK were used again, but sequence number doubled, and an identical tree was produced (Chu et al., 2009). The protein-coding genes used by Tsang et al. (2008b) and Chu et al. (2009) supported many infraordinal to species level relationships, providing evidence that single-copy, slow-evolving, protein-coding genes are good candidates for inferring phylogenies across broad taxonomic ranges.

Toon et al. (2009) sequenced eight markers, two mitochondrial and six nuclear, for representatives of Achelata, Anomura, Astacidea, Axiidea, Brachyura, Caridea, Dendrobranchiata, and Polychelida. While many sequences were recycled from GenBank (including 12S, 16S, 18S, 28S, and H3 for several specimens), three new nDNA markers were introduced: EF-2, EPRS, and TM9sf4. RAxML analysis of the eight markers inferred relationships that did not concur with other studies, primarily by recovering Dendrobranchiata as sister to Caridea, and Caridea as the most basal pleocyemate (Fig. 1H). However, these branches were not strongly supported. While it is not explicitly stated whether data was partitioned or not, a second analysis, which excluded the mtDNA markers, was performed but not presented.

In 2009, Bracken et al. published their work on the Decapod Tree of Life Project (Bracken et al., 2009a; Fig. 1G), combining an increased sampling effort with multiple-marker analysis. Most of the data was recycled from previous analyses (only 24 *de novo* sequences), including every currently recognized infraorder except for Procarididea. The authors used a subset of the markers used by Toon et al. (2009): 16S, 18S, 28S, and H3. Sequences were analyzed with RAxML and BI. The resulting tree provided further support for the division of Thalassinidea (Gurney, 1938; Tsang et al., 2008a, b; Robles et al., 2009). Although monophyly of all infraorders was statistically supported, there was little to no support for relationships among infraorders, due to the lack of appropriate genes to resolve deep relationships.

Another study, aimed at investigating Procarididea evolution, also generated an infraordinal tree (Bracken et al., 2010). The findings of Bracken et al. agreed with those of Felgenhauer and Abele's (1983) comparative morphological study, establishing Procarididea as an infraorder, sister to Caridea (Fig. 1I). Dendrobranchiata was sampled, as well as every currently recognized decapod infraorder, except for Glypheidea. This study analyzed one mitochondrial marker (16S) and three nuclear markers (18S, 28S, and H3), generating new data for taxa in Dendrobranchiata, Procarididea, and Caridea. Data for representatives from the other infraorders was recycled from GenBank. Genes were concatenated and partitioned for analysis. MODELTEST was used to identify the evolutionary model that best fit the data, and data was analyzed using RAxML (Stamatakis et al., 2005, 2007, 2008) and MrBayes (Huelsenbeck and Ronquist, 2001). To calibrate the resulting tree, thirteen fossils were included in the analysis.

Beginning in the 2010s, high-powered NGS techniques began generating huge quantities of data for decapod phylogenetics. Through massively parallel, multiplexed reactions, NGS is capable of generating genomic, transcriptomic, and epigenomic data (Levin et al., 2009; McKenna et al., 2010; Metzker, 2010; Roukos, 2010; Ku et al., 2011; Martin and Wang, 2011; McCormack et al., 2013; Wong et al., 2015). Such sequencing efforts allow analysis of hundreds to thousands of markers across the genomes of hundreds of individuals (Gnirke et al., 2009; Mamanova et al., 2010; Lemmon and Lemmon, 2012; Lemmon et al., 2012), generating unprecedented amounts of data while using fewer resources. Applied to decapod phylogenetics relatively recently, NGS has enabled the targeting of hundreds of new markers across the order.

Targeted Amplicon Sequencing (TAS) (Bybee et al., 2011a) uses an NGS platform to sequence a high number of markers across a large number of specimens. This PCR-based approach generates amplicons optimized for NGS (Bybee et al., 2011a, b). Target genes undergo two PCRs, which barcode sequences by taxon, enabling them to be multiplexed on a NGS platform (Bybee et al., 2011a, b). The PCR amplification allows for the use of a variety of starting materials, e.g., new specimens, alcohol-preserved tissue, museum samples. Sequencing 12S, 16S, COI, 18S, 28S, and H3 *de novo* for 16 specimens, including a museum specimen, Bybee et al. (2011b) demonstrated the potential of TAS across Pancrustacea (including Decapoda; Fig. 1K). The study itself lacked representatives from Gebiidea, Glypheidea, and Procarididea, however the intention of this study was not to generate a robust phylogeny across Decapoda, but rather to exemplify how the method could be applied to

deep phylogenetic inferences. The authors highlight potential problems with TAS, such as the quality of the data (reviewed by Wicker et al., 2006; Huse et al., 2007; Kunin et al., 2010), the removal of primer dimers, and biases among barcodes.

Two recent studies (Qian et al., 2011; Shen et al., 2013) have taken similar approaches to generate and analyze mt-genomes to infer decapod phylogeny. These are the first phylogenomic studies of decapods thus far. Qian et al. (2011) combined 27 previously sequenced mt-genomes with two *de novo* mt-genomes generated for the analysis. Though data was not partitioned, the 13 protein-coding genes were analyzed in separate alignments. The results of Qian et al. (2011) strongly support topologies from other studies (Tsang et al., 2008b; Bybee et al., 2011b; Shen et al., 2013), with Brachyura and Anomura representing derived branches and Dendrobranchiata and Caridea representing early branching groups (Fig. 1J). In addition to Dendrobranchiata, only five infraorders were sampled: Achelata, Anomura, Astacidea, Brachyura, and Caridea. Shen et al. (2013) generated two datasets: an amino acid alignment and a sequence alignment. Both were partitioned by gene. Results were similar to Qian et al., but the data showed some ambiguity as to the position of Polychelida in relation to Achelata and Astacidea: BI of mt-genome nucleotides and ML analysis of mt-genome amino acids upheld Palinura (Polychelida + Achelata), but all other analyses, including the final tree (Fig. 1L) based on the analysis of all datasets, supported Polychelida + Astacidea. Few analyses resulted in high support values suggesting a relationship between Polychelida + Astacidea or Polychelida + Achelata. ML analysis of mt-genome amino acids also resulted in monophyly of Thalassinida (Gebiidea + Achelata), though this result did not carry to the final tree (Fig. 1K), in which Axiidea is basal to Gebiidea. It should be noted that past results have suggested using mitochondrial genomes to infer phylogeny can be problematic, as previously discussed.

Future Directions

Arguably, one of the most promising methods for resolving the decapod tree of life has not yet been applied to decapod phylogeny: Anchored Hybrid Enrichment (AHE) is capable of targeting hundreds of loci informative at multiple taxonomic levels in a single NGS study. AHE targets many (>500) highly conserved anchored regions of the genome using probes (Lemmon and Lemmon, 2012; Lemmon et al., 2012). Each streptavidin-tagged, oligonucleotide probe targets a highly conserved sequence region flanked by more variable sequence regions. Probes can be designed to target flanking regions exhibiting different levels of variability. The result is sequence data that is phylogenetically informative at multiple taxonomic levels in a single study. By designing probes to target appropriately variable sequences, relationships can be resolved from the deep phylogenetic level to the level of phylogeography (Carstens et al., 2012; Lemmon and Lemmon, 2012).

As NGS methods lower the cost of phylogenetic studies, allowing the discovery of unprecedented numbers of markers and inclusion of many taxa, it is important to remember the value of morphological data in phylogenetic analyses. Previous studies of decapod phylogeny have demonstrated that including morphological characters to a molecular dataset can improve the phylogeny in terms of support and sampling effort (Ahyong and O'Meally, 2004; Schnabel et al., 2011; Bracken-Grissom et al., 2013, 2014). Specifically, the inclusion of fossils in a phylogenetic study can incorporate data that cannot be generated from any other source (Novacek and Norell, 1982). Most notably, fossils can allow extinct taxa to be included in phylogenies (Beurlen and Glaessner, 1930). A rich fossil record allows researchers to estimate the age of clades (Novacek and Norell, 1982; Reid et al., 1996) and explore the origins of diversity within major lineages (Gauthier et al., 1988; Huelsenbeck, 1991; Weishampel, 1996; Bracken-Grissom et al., 2014). Using fossils to date a phylogenetic tree can add directionality to major morphological and/or behavioral transitions and uncover historical patterns in organismal biogeography (Porter et al., 2005; Bracken-Grissom et al., 2014). According to recent studies, even including just one fossil for every ten included taxa can reliably date a phylogeny (Erwin et al., 2011; Bracken-Grissom et al., 2014). In summary, as phylogenetics moves toward NGS approaches, it is important to remember the inimitable role fossils can play in recapitulating a robust, dated phylogenetic tree for Decapoda.

Conclusions

From the earliest classifications of decapods, to the superpowered molecular methods of NGS, morphological and molecular phylogenies have generated a suite of evolutionary hypotheses for deep relationships. From these varied hypotheses, some accord has been seen. Early studies consistently recovered three or four major lineages: Dendrobranchiata, Caridea, Stenopodidea, and Reptantia, with Dendrobranchiata generally considered to be the earliest branching lineage. Reptant infraorders (Achelata, Anomura, Astacidea, Axiidea, Brachyura, Gebiidea, Glypheidea, Polychelida) are typically recovered as derived lineages. Caridea and Stenopodidea frequently cluster together, either as sister groups or as close relatives. Generally, Caridea and Dendrobranchiata represent early branching lineages, while Anomura and Brachyura fall as sister clades in a more derived position on the Decapod Tree of Life. The lobster-like lineages Polychelida, Glypheidea, Achelata, and Astacidea show conflicting relationships as either a monophyletic (Tsang et al., 2008b; Chu et al., 2009; Toon et al., 2009; Bybee et al., 2011b; Qian et al., 2011) or non-monophyletic clade (Ahyong and O'Meally, 2004; Porter et al., 2005; Bracken et al., 2009a, 2010; Shen et al., 2013; Bracken-Grissom et al., 2014). The ghost shrimp infraorders, Axiidea and Gebiidea, are consistently recovered as non-monophyletic (Porter et al., 2005; Tsang et al., 2008b; Bracken et al., 2009a; Chu et al., 2009; Shen et al., 2013). Further contributing to our understanding of decapod phylogeny, many recent molecular phylogenies have focused on family-level relationships within one or more infraorders (Anomura: Ahyong et al., 2009; Tsang et al., 2011; Bracken-Grissom et al., 2013; Axiidea/Gebiidea: Tsang et al., 2008a; Robles et al., 2009; Brachyura: Tsang et al., 2014; Caridea: Bracken et al., 2009b; Li et al., 2011; Dendrobranchiata: Ma et al., 2009; Lobster-like lineages: Bracken-Grissom et al., 2012,

2014). Past studies have undoubtedly enhanced our understanding of the Decapod Tree of Life; however, several infraordinal relationships remain unclear. In pursuit of strong infraordinal-level support across Decapoda, analysis methods have become more standardized and taxon sampling has improved, while a lack of appropriate markers has remained a primary hindrance. Since the introduction of NGS, techniques have advanced and optimized to meet the challenge of deep phylogenetic questions. Excitingly, these advancements now provide researchers with hundreds to thousands of phylogenetically informative markers, enabling unprecedented insight into the evolutionary history of Decapoda.

ACKNOWLEDGEMENTS

The research was made possible in part by a grant from BP/The Gulf of Mexico Research Initiative and in part by the Florida International University Presidential Fellowship. We thank L. Timm's Ph. D. committee: Dr. Eirin-Lopez, Dr. Rodriguez-Lanetty, Dr. von Wettberg, and Dr. Wensong Wu, for additional comments and suggestions. Finally, we thank Mr. Joseph Ahrens, Drs. Phil Stoddard and Brent Thoma, and the anonymous reviewers for feedback on earlier versions of this manuscript.

REFERENCES

Abele, L. G. 1991. Comparison of morphological and molecular phylogeny of the Decapoda. Memoirs of the Queensland Museum 31: 101-108.

Ahyong, S. T., and D. O'Meally. 2004. Phylogeny of the Decapoda Reptantia: resolution using three molecular loci and morphology. The Raffles Bulletin of Zoology 52: 673-693.

———, J. C. Lai, D. Sharkey, D. J. Colgan, and P. K. Ng. 2007. Phylogenetics of the brachyuran crabs (Crustacea: Decapoda): the status of Podotremata based on small subunit nuclear ribosomal RNA. Molecular Phylogenetics and Evolution 45: 576-586.

———, K. E. Schnabel, and E. W. Maas. 2009. Anomuran phylogeny: new insights from molecular data, pp. 399-414. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

Avise, J. C. 2000. Phylogeography: the History and Formation of Species. Harvard University Press.

Baldauf, S. L. 2003. Phylogeny for the faint of heart: a tutorial. Trends in Genetics 19: 345-351.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell, and M. J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82: 247-277.

Ballard, J. W. O., and D. M. Rand. 2005. The population biology of mitochondrial DNA and its phylogenetic implications. Annual Review of Ecology, Evolution, and Systematics 36: 621-642.

———, and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729-744.

Bate, C. S. 1888. Report on the Crustacea Macrura collected by H.M.S. Challenger during the Years 1873-76. In: Murray, J., Zoology. Wyville Thomson, C. and J. Murray, Report on the Scientific Results of the Voyage of H.M.S. Challenger During the Years 1873-76 Under the Command of Captain George S. Nares, R.N., F.R.S. and the Late Captain Frank Tourle Thomson, R.N., Vol. 24. Neill, Edinburgh.

Bazin, E., S. Glémin, and N. Galtier. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312: 570-572.

Bell, M. V., A. E. Cowper, M. P. Lefranc, J. I. Bell, and G. R. Screaton. 1998. Influence of intron length on alternative splicing of CD44. Molecular and Cellular Biology 18: 5930-5941.

Bergsten, J. 2005. A review of long-branch attraction. Cladistics 21: 163-193.

Beurlen, K., and M. F. Glaessner. 1930. Systematik der Crustacea Decapoda auf stammesgeschichtlicher grundlage. Zoolgische Jahrbücher Abteilung Für Systematik, Öekologie Und Geographie Der Tiere 60: 49-84.

Birky, C. W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proceedings of the National Academy of Sciences 92: 11331-11338.

———. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annual Review of Genetics 35: 125-148.

Blouin, M. S., C. A. Yowell, C. H. Courtney, and J. B. Dame. 1998. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. Molecular Biology and Evolution 15: 1719-1727.

Boas, J. E. V. 1880. Studier over Decapodernes Slægtskabsforhold. Danske Videnskabernes Seksjeab, Copenhagen, Skrifter, Naturvidenskabelig og matematisek Afdeling 1: 23-210.

Boore, J. L., and W. M. Brown. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Current Opinion in Genetics and Development 8: 668-674.

Borradaile, L. A. 1907. On the classification of the decapod crustaceans. Annals and Magazine of Natural History 19: 457-486.

Boyko, C. B., and J. D. Williams. 2009. Crustacean parasites as phylogenetic indicators in decapod evolution, pp. 197-220. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

Bracken, H. D., A. Toon, D. L. Felder, J. W. Martin, M. Finley, J. Rasmussen, F. Palero, and K. A. Crandall. 2009a. The decapod tree of life: compiling the data and moving toward a consensus of decapod evolution. Arthropod Systematics and Phylogeny 67: 99-116.

———, S. De Grave, and D. L. Felder. 2009b. Phylogeny of the infraorder Caridea based on mitochondrial and nuclear genes (Crustacea: Decapoda), pp. 281-305. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

———, ———, A. Toon, D. L. Felder, and K. A. Crandall. 2010. Phylogenetic position, systematic status, and divergence time of the Procarididea (Crustacea: Decapoda). Zoologica Scripta 39: 198-212.

Bracken-Grissom, H. D., S. T. Ahyong, R. D. Wilkinson, R. M. Feldmann, C. E. Schweitzer, J. W. Breinholt, M. Bendall, F. Palero, T.-Y. Chan, D. L. Felder, R. Robles, K.-H. Chu, L.-M. Tsang, D. Kim, J. W. Martin, and K. A. Crandall. 2014. The emergence of the lobsters: phylogenetic relationships, morphological evolution and divergence time comparisons of an ancient group (Decapoda: Achelata, Glypheidea, Polychelida). Systematic Biology 63: 457-479.

———, M. E. Cannon, P. Cabezas, R. M. Feldmann, C. E. Schweitzer, S. T. Ahyong, D. L. Felder, R. Lemaitre, and K. A. Crandall. 2013. A comprehensive and integrative reconstruction of evolutionary history for Anomura (Crustacea: Decapoda). BMC Evolutionary Biology 13: 1-128.

———, C. H. Yang, D. Kim, K. A. Crandall, and T. Y. Chan. 2012. Phylogenetic relationships, character evolution, and taxonomic implications within the slipper lobsters (Crustacea: Decapoda: Scyllaridae). Molecular Phylogenetics and Evolution 62: 237-250.

Brown, J. M. 2014a. Predictive approaches to assessing the fit of evolutionary models. Systematic Biology 63: 289-292.

———. 2014b. Detection of implausible phylogenetic inferences using posterior predictive assessment of model fit. Systematic Biology 63: 334-348.

Brown, W. M., M. George, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America 76: 1967-1971.

Buckley, T. R. 2002. Model misspecification and probabilistic tests of topology: evidence from empirical data sets. Systematic Biology 51: 509-523.

———, and C. W. Cunningham. 2002. The effects of nucleotide substitution model assumptions on estimates of nonparametric bootstrap support. Molecular Biology and Evolution 19: 394-405.

———, C. Simon, and G. K. Chambers. 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Systematic Biology 50: 67-86.

Burkenroad, M. D. 1963. The evolution of the Eucarida (Crustacea, Eumalacostraca), in the relation of the fossil record. Tulane Studies in Geology 2: 1-17.

Bybee, S. M., H. D. Bracken-Grissom, B. D. Haynes, R. A. Hermansen, R. L. Byers, M. J. Clement, J. A. Udall, E. R. Wilcox, and K. A. Crandall. 2011a. Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. Genome Biology and Evolution 3: 1312-1323.

———, ———, R. A. Hermansen, M. J. Clement, K. A. Crandall, and D. L. Felder. 2011b. Directed next generation sequencing for

phylogenetics: an example using Decapoda (Crustacea). Zoologischer Anzeiger 250: 497-506.
Calman, W. D. 1904. On the classification of the Crustacea Malacostraca. Annals and Magazine of Natural History 13: 144-158.
Carstens, B., A. R. Lemmon, and E. M. Lemmon. 2012. The promises and pitfalls of next-generation sequencing data in phlyogeography. Systematic Biology 61: 713-715.
Chu, K. H., L.-M. Tsang, K. Y. Ma, T.-Y. Chan, and P. K. L. Ng. 2009. Decapod phylogeny: what can protein-coding genes tell us?, pp. 89-99. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.
Clark, P. F. 2009. The bearing of larval morphology on brachyuran phylogeny, pp. 221-241. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.
Claus, C. 1872. Gründzuge der Zoologie. Zum gebrauche an Universitäten und Höheren Lehranstalten sowie zum Seblststudium (zweite vermehrte Auflage). N. G. Elwert'sche Universitäts-Buchhandlung, Marburg.
Crandall, K. A., D. J. Harris, and J. W. Fetzner. 2000. The monophyletic origin of freshwater crayfish estimated from nuclear and mitochondrial DNA sequences. Proceedings of the Royal Society of London Series B: Biological Sciences 267: 1679-1686.
Dana, J. D. 1852. Crustacea. Part I. United States Exploring Expedition. During the Years 1838, 1839, 1840, 1841, 1842. Under the Command of Charles Wilkes, U.S.N. Vol. 13. C. Sherman, Philadelphia, PA.
Darwin, C. 1859. The Origin of Species by Means of Natural Selection. Murray, London.
Dayrat, B., and S. Tillier. 2002. Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaluation of morphological characters. Zoological Journal of the Linnean Society 135: 403-470.
De Grave, S., N. D. Pentcheff, S. T. Ahyong, T.-Y. Chan, K. A. Crandall, P. C. Dworschak, D. L. Felder, R. M. Feldmann, C. H. J. M. Fransen, L. Y. D. Goulding, R. Lemaitre, M. E. Y. Low, J. W. Martin, P. K. L. Ng, C. E. Schweitzer, S. H. Tan, D. Tshudy, and R. Wetzer. 2009. A classification of living and fossil genera of decapod crustaceans. Raffles Bulletin of Zoology 21: 1-109.
de Queiroz, K., and J. Gauthier. 1990. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. Systematic Zoology 39: 307-322.
de Saint Laurent, M. 1979. Vers une nouvelle classification des Crustacés Décapodes Reptantia. Bulletin de l'Office National des Pêches République Tunisienne, Ministere de l'Agriculture 3: 15-31.
Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology and Evolution 24: 332-340.
Dixon, C. J., S. T. Ahyong, and F. R. Schram. 2003. A new hypothesis of decapod phylogeny. Crustaceana 76: 935-975.
Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17: 144-163.
Edwards, S. V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63: 1-19.
Elson, J. L., and R. N. Lightowlers. 2006. Mitochondrial DNA clonality in the dock: can surveillance swing the case? Trends in Genetics 22: 603-607.
Erwin, D. H., M. Laflamme, S. M. Tweedt, E. A. Sperling, D. Pisani, and K. J. Peterson. 2011. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science 334: 1091-1097.
Farris, J. S., A. G. Kluge, and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. Systematic Biology 19: 172-189.
Felgenhauer, B. E., and L. G. Abele. 1983. Phylogenetic relationships among shrimp-like decapods, pp. 291-311. In, F. R. Schram (ed.), Crustacean Phylogeny. Crustacean Issues 1. A. A. Balkema, Rotterdam.
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.
Fenn, J. D., H. Song, S. L. Cameron, and M. F. Whiting. 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Molecular Phylogenetics and Evolution 49: 59-68.
Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. Systematic Biology 19: 99-113.
———. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Systematic Biology 20: 406-416.
Forterre, P., and H. Philippe. 1999. Where is the root of the universal tree of life?. Bioessays 21: 871-879.
Friedrich, M., and D. Tautz. 1995. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. Nature 376: 165-167.
Gabaldón, T., and E. V. Koonin. 2013. Functional and evolutionary implications of gene orthology. Nature Reviews Genetics 14: 360-366.
Gadagkar, S. R., M. S. Rosenberg, and S. Kumar. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 304: 64-74.
Galtier, N., B. Nabholz, S. Glémin, and G. D. D. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology 18: 4541-4550.
Gatesy, J., R. DeSalle, and W. Wheeler. 1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Molecular Phylogenetics and Evolution 2: 152-157.
Gauthier, J., A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4: 105-209.
Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E. M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher, C. Russ, S. Gabriel, D. B. Jaffe, E. S. Lander, and C. Nusbaum. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nature Biotechnology 27: 182-189.
Goldman, N. 1993. Simple diagnostic statistical tests of models for DNA substitution. Journal of Molecular Evolution 37: 650-661.
Graur, D., and W. H. Li. 2000. Fundamentals of Molecular Evolution. 2nd Edition. Sinauer Associates, Sunderland, MA.
Gurney, R. 1938. Larvae of the decapod Crustacea. V. Nephropsidea and Thalassinidea. Discovery Reports 17: 291-344.
Hasegawa, M., H. Kishino, and T. A. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 222: 160-174.
Hillis, D. M., D. D. Pollock, J. A. McGuire, and D. J. Zwickl. 2003. Is sparse taxon sampling a problem for phylogenetic inference? Systematic Biology 52: 124.
Huelsenbeck, J. P. 1991. When are fossils better than extant taxa in phylogenetic analysis? Systematic Biology 40: 458-469.
———, and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755.
Hurst, G. D., and F. M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proceedings of the Royal Society London Series B: Biological Sciences 272: 1525-1534.
Huse, S. M., J. A. Huber, H. G. Morrison, M. L. Sogin, and D. M. Welch. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology and Evolution 8: R143.
Huxley, T. H. 1878. On the classification and the distribution of the crayfishes. Proceedings of the Zoological Society of London 1878: 752-788.
Jenner, R. A. 2001. Bilaterian phylogeny and uncritical recycling of morphological data sets. Systematic Biology 50: 730-742.
———, and F. R. Schram. 1999. The grand game of metazoan phylogeny: rules and strategies. Biological Reviews 74: 121-142.
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, NY.
Kim, W., and L. G. Abele. 1990. Molecular phylogeny of selected decapod crustaceans based on 18S rRNA nucleotide sequences. Journal of Crustacean Biology 10: 1-13.
Kim, Y. K., and V. N. Kim. 2007. Processing of intronic microRNAs. The EMBO Journal 26: 775-783.
Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111-120.
Kingman, J. F. 2000. Origins of the coalescent: 1974-1982. Genetics 156: 1461-1463.
Kondrashov, F. A., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin. 2002. Selection in the evolution of gene duplications. Genome Biology 3: 1-9.
Koonin, E. V. 2005. Orthologs, paralogs, and evolutionary genomics. Annual Review of Genetics 39: 309-338.

Ku, C. S., N. Naidoo, M. Wu, and R. Soong. 2011. Studying the epigenome using next generation sequencing. Journal of Medical Genetics 48: 721-730.

Kubatko, L. S., and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56: 17-24.

Kunin, V., A. Engelbrektson, H. Ochman, and P. Hugenholtz. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflations of diversity estimates. Environmental Microbiology 12: 118-123.

Lanfear, R., B. Calcott, S. Y. Ho, and S. Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29: 1695-1701.

Latreille, P. A. 1802. Histoire naturelle, générale et particulière des Crustacés et des Insectes. Ouvrage faisant suite à l'histoire naturelle générale et particulière, composée par Leclerc de Buffon, et rédigée par C.S. Sonnini, membre de plusieurs sociétés savantes. Familles naturelles des genres, Vol. 3. F. DuFart, Paris.

———. 1817. Nouveau dictionnaire d'histoire naturelle, appliquée aux arts, à l'agriculture, à l'economie rurale et domestique, à la médecine, etc. Vol. 10. Déterville, Paris.

———. 1831. Cours d'Entomologie, ou de l'Histoire Naturelle des Crustacés, des Arachnides, des Myriapodes et des Insects; a l'Usage des Élèves de l'École du Muséum d'Histoire Naturelle. Librairie Encyclopédique de Roret, Paris.

Le, S. Q., N. Lartillot, and O. Gascuel. 2008. Phylogenetic mixture models for proteins. Philosophical Transactions of the Royal Society B: Biological Sciences 363: 3965-3976.

Lemmon, A. R., S. A. Emme, and E. M. Lemmon. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Systematic Biology 61: 727-744.

———, and E. M. Lemmon. 2012. High-throughput identification of informative nuclear loci for shallow-scale phylogenetics and phylogeography. Systematic Biology 61: 745-761.

———, and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Systematic Biology 53: 265-277.

Levin, J. Z., M. F. Berger, X. Adiconis, P. Rogov, A. Melnikov, T. Fennell, C. Nussbaum, L. A. Garraway, and A. Gnirke. 2009. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. Genome Biology 10: R115.

Lewis, P. O., W. Xie, M. H. Chen, Y. Fan, and L. Kuo. 2014. Posterior predictive Bayesian phylogenetic model selection. Systematic Biology 63: 309-321.

Li, C. P., S. De Grave, T.-Y. Chan, H. C. Lei, and K. H. Chu. 2011. Molecular systematics of caridean shrimps based on five nuclear genes: implications for superfamily classification. Zoologischer Anzeiger 250: 270-279.

Liu, L., L. Yu, L. Kubatko, D. K. Pearl, and S. V. Edwards. 2009a. Coalescent methods for estimating phylogenetic trees. Molecular Phylogenetics and Evolution 53: 320-328.

———, ———, ———, and ———. 2009b. Estimating species phylogenies using coalescence times among sequences. Systematic Biology 58: 468-477.

Lundberg, J. G. 1972. Wagner networks and ancestors. Systematic Biology 21: 398-413.

Ma, K. Y., T.-Y. Chan, and K. H. Chu. 2009. Phylogeny of penaeoid shrimps (Decapoda: Penaeoidea) inferred from nuclear protein-coding genes. Molecular Phylogenetics and Evolution 53: 45-55.

MacLeay, W. S. 1838. On the brachyurous decapod Crustacea brought from the Cape by Dr. Smith, pp. 53-71. In, A. Smith, Illustrations of the Annulosa of South Africa; Being a Portion of the Objects of Natural History Chiefly Collected During an Expedition into the Interior of South Africa, Under the Direction of Dr. Andrew Smith, in the Years 1834, 1835, and 1836; Fitted out by "The Cape of Good Hope Association for Exploring Central Africa". Smith, Elder, and Co., London.

Maddison, W. P. 1997. Gene trees in species trees. Systematic Biology 46: 523-536.

———, and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55: 21-30.

Mamanova, L., A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shendure, and D. J. Turner. 2010. Target-enrichment strategies for next-generation sequencing. Nature Methods 7: 111-118.

Martin, J. A., and Z. Wang. 2011. Next-generation transcriptome assembly. Nature Reviews Genetics 12: 671-682.

Martin, J. W., and G. E. Davis. 2001. An updated classification of the recent Crustacea. Natural History Museum of Los Angeles County Science Series 39: 1-124.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66: 526-538.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20: 1297-1303.

McVay, J. D., and B. C. Carstens. 2013. Phylogenetic model choice: justifying a species tree or concatenation analysis. Journal of Phylogenetics and Evolutionary Biology 1: 1-8.

Meiklejohn, C. D., K. L. Montooth, and D. M. Rand. 2007. Positive and negative selection on the mitochondrial genome. Trends in Genetics 23: 259-263.

Meland, K., and E. Willassen. 2007. The disunity of "Mysidacea" Crustacea. Molecular Phylogenetics and Evolution 44: 1083-1104.

Metzker, M. L. 2010. Sequencing technologies – the next generation. Nature Reviews Genetics 11: 31-46.

Milne Edwards, H. 1834. Histoire naturelle des Crustacés, comprenant l'anatomie, la physiologie et la classification de ces animaux. Librairie Encyclopédique de Roret, Vols 1-3. Roret, Paris. 1468, 2532, 3638 pp., Atlas 1-32, Plates I-XLII.

Monod, T., and L. Laubier. 1996. Les Crustaces dans la Biosphere, pp. 91-166. In, J. Forest (ed.), Traite de Zoologie, Anatomie, Systematique, Biologie. Crustaces. Tome VII. Fascicule II. Generalites suite et Systematique, ed. J. Forest. Masson, Paris.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49: 718-726.

Nishihara, H., N. Okada, and M. Hasegawa. 2007. Rooting the eutherian tree: the power and pitfalls of phylogenomics. Genome Biology 8: R199.

Novacek, M. J., and M. A. Norell. 1982. Fossils, phylogeny, and taxonomic rates of evolution. Systematic Biology 31: 366-375.

Opperdoes, F. R. 2009. Phylogenetic analysis using protein sequences, pp. 313-331. In, P. Lemey, M. Salemi, and A.-M. Vandamme (eds.), The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge.

Page, R. D., and M. A. Charleston. 1997. From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. Molecular Phylogenetics and Evolution 7: 231-240.

Palero, F., and K. A. Crandall. 2009. Phylogenetic inference using molecular data, pp. 67-88. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5: 568-583.

Poe, S., and J. J. Wiens. 2000. Character selection and the methodology of morphological phylogenetics, pp. 20-36. In, J. J. Wiens (ed.), Phylogenetic Analysis of Morphological Data. Smithsonian Institution Press, Washington, DC.

Porter, M. L., and T. W. Cronin. 2009. A shrimp's eye view of evolution: how useful are visual characters in decapod phylogenetics?, pp. 183-195. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

———, M. Perez-Losada, and K. A. Crandall. 2005. Model-based multilocus estimation of decapod phylogeny and divergence times. Molecular Phylogenetics and Evolution 37: 355-369.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25: 1253-1256.

———, and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Qian, G., Q. Zhoa, A. Wang, L. Zhu, K. Zhou, and H. Sun. 2011. Two new decapod Crustacea: Malacostraca complete mitochondrial genomes: bearings on the phylogenetic relationships within the Decapoda. Zoological Journal of the Linnean Society 162: 471-481.

Rafinesque, C. S. 1815. Analyse de la Nature, ou Tableau de l'Univers et des Corps Organisés. L'Imprimerie de Jean Barravecchia, Palermo.

Reid, D. G., E. Rumbak, and R. H. Thomas. 1996. DNA, morphology and fossils: phylogeny and evolutionary rates of the gastropod genus *Littorina*. Philosophical Transactions of the Royal Society B: Biological Sciences 351: 877-895.

Reid, N. M., S. M. Hird, J. M. Brown, T. A. Pelletier, J. D. McVay, J. D. Satler, and B. C. Carstens. 2013. Poor fit to the multispecies coalescent is widely detectable in empirical data. Systematic Biology 63: 322-333.

Richter, S., and G. Scholtz. 2001. Phylogenetic analysis of the Malacostraca Crustacea. Journal of Zoological Systematics and Evolutionary Research 39: 113-136.

Robles, R., C. Tudge, P. C. Dworschak, G. C. B. Poore, and D. L. Felder. 2009. Molecular phylogeny of the Thalassinidea based on nuclear and mitochondrial genes, pp. 309-326. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

Rokas, A., and P. W. Holland. 2000. Rare genomic changes as a tool for phylogenetics. Trends in Ecology and Evolution 15: 454-459.

———, B. L. Williams, N. King, and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425: 798-804.

Roukos, D. H. 2010. Next-generation sequencing and epigenome technologies: potential medical applications. Expert Review of Medical Devices 7: 723-726.

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

Schmidt, H. A., and A. von Haeseler. 2009. Phylogenetic inference using maximum likelihood methods, pp. 181-198. In, P. Lemey, M. Salemi, and A.-M. Vandamme (eds.), The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge, MA.

Schnabel, K. E., S. T. Ahyong, and E. W. Maas. 2011. Galatheoidea are not monophyletic – molecular and morphological phylogeny of the squat lobsters (Decapoda: Anomura) with recognition of a new superfamily. Molecular Phylogenetics and Evolution 58: 157-168.

Schneider, A., and D. Ebert. 2004. Covariation of mitochondrial genome size with gene lengths: evidence for gene length reduction during mitochondrial evolution. Journal of Molecular Evolution 59: 90-96.

Scholtz, G., and S. Richter. 1995. Phylogenetic systematics of the reptantian Decapoda Crustacea, Malacostraca. Zoological Journal of the Linnean Society 113: 289-328.

Schram, F. R. 1981. On the classification of the Eumalacostraca. Journal of Crustacean Biology 1: 1-10.

———. 1984. Relationships within eumalacostracan crustaceans. Transactions of the San Diego Society of Natural History 20: 301-312.

———. 1986. Crustacea. Oxford University Press, New York, NY.

———. 2001. Phylogeny of decapods: moving towards a consensus. Hydrobiologia 449: 1-20.

———, and C. H. J. Hof. 1998. Fossil taxa and the relationships of major crustacean groups, pp. 273-302. In, G. Edgecombe (ed.), Arthropod Fossils and Phylogeny. Columbia University Press, New York, NY.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proceedings of the National Academy of Sciences of the United States of America 99: 16122-16127.

Shen, H., A. Braband, and G. Scholtz. 2013. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. Molecular Phylogenetics and Evolution 66: 776-789.

Siewing, R. 1963. Studies in malacostracan morphology: results and problems, pp. 85-103. In, H. B. Whittington and W. D. I. Rolfe (eds.), Phylogeny and Evolution of Crustacea. Museum of Comparative Zoology, Cambridge, MA.

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

Sota, T., and A. P. Vogler. 2003. Reconstructing species phylogeny of the carabid beetles *Ohomopterus* using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses. Molecular Phylogenetics and Evolution 26: 139-154.

Spinks, P. Q., and H. B. Shaffer. 2009. Conflicting mitochondrial and nuclear phylogenies for the widely disjunct *Emys* (Testudines: Emydidae) species complex, and what they tell us about biogeography and hybridization. Systematic Biology 58: 1-20.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688-2690.

———, F. Blagojevic, D. S. Nikolopoulos, and C. D. Antonopoulos. 2007. Exploring new search algorithms and hardware for phylogenetics: RAxML meets the IBM cell. The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology 48: 271-286.

———, P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML web servers. Systematic Biology 57: 758-771.

———, T. Ludwig, and H. Meier. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21: 456-463.

Studer, R. A., and M. Robinson-Rechavi. 2009. How confident can we be that orthologs are similar, but paralogs differ? Trends in Genetics 25: 210-216.

Talavera, G., and R. Vila. 2011. What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. BMC Evolutionary Biology 11: 315-329.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17: 57-86.

Telford, M. J., and R. R. Copley. 2011. Improving animal phylogenies with genomic data. Trends in Genetics 27: 186-195.

Throckmorton, L. H. 1968. Concordance and discordance of taxonomic characters in *Drosophila* classification. Systematic Biology 17: 355-387.

Toon, A., M. Finley, J. Staples, and K. A. Crandall. 2009. Decapod phylogenetics and molecular evolution, pp. 15-29. In, J. W. Martin, K. A. Crandall, and D. L. Felder (eds.), Decapod Crustacean Phylogenetics. Crustacean Issues 18. CRC Press, Boca Raton, FL.

Tsang, L.-M., T. Y. Chan, S. T. Ahyong, and K. H. Chu. 2011. Hermit to king, or hermit to all: multiple transitions to crab-like forms from hermit crab ancestors. Systematic Biology 60: 616-629.

———, F.-J. Lin, K. H. Chu, and T.-Y. Chan. 2008a. Phylogeny of Thalassinidea Crustacea, Decapoda inferred from three rDNA sequences: implications for morphological evolution and superfamily classification. Journal of Zoological Systematics and Evolutionary Research 46: 216-223.

———, K. Y. Ma, S. T. Ahyong, T.-Y. Chan, and K. H. Chu. 2008b. Phylogeny of Decapoda using two nuclear protein-coding genes: origin and evolution of the Reptantia. Molecular Phylogenetics and Evolution 48: 359-368.

———, C. D. Schubart, S. T. Ahyong, J. C. Lai, E. Y. Au, T.-Y. Chan, P. K. L. Ng, and K. H. Chu. 2014. Evolutionary history of true crabs (Crustacea: Decapoda: Brachyura) and the origin of freshwater crabs. Molecular Biology and Evolution 31: 1173-1187.

Valentine, J. W., D. Jablonski, S. Kidwell, and K. Roy. 2006. Assessing the fidelity of the fossil record by using marine bivalves. Proceedings of the National Academy of Sciences of the United States of America 103: 6599-6604.

Van de Peer, Y. 2009. Phylogenetic inference based on distance methods, pp. 142-160. In, P. Lemey, M. Salemi, and A.-M. Vandamme (eds.), The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge.

Vandamme, A.-M. 2009. Basic concepts of molecular evolution, pp. 3-29. In, P. Lemey, M. Salemi, and A.-M. Vandamme (eds.), The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge.

Weishampel, D. B. 1996. Fossils, phylogeny, and discovery: a cladistic study of the history of tree topologies and ghost lineage durations. Journal of Vertebrate Paleontology 16: 191-197.

Wheeler, W. C. 1990. Nucleic acid sequence phylogeny and random outgroups. Cladistics 6: 363-367.

Wicker, T., E. Schlagenhauf, A. Graner, T. J. Close, B. Keller, and S. Nils. 2006. 454 sequencing put to the test using the complex genome barley. BMC Genomics 7: 275.

Wiens, J. J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Systematic Biology 52: 528-538.

Wiley, E. O., D. Siegel-Causey, D. R. Brooks, and V. A. Funk. 1991. The Compleat Cladist: a Primer of Phylogenetic Procedures. University of Kansas Museum of Natural History, Lawrence, KS.

Williamson, D. I. 1973. *Amphionides reynaudii* (H. Milne Edwards) representative of a proposed new order of eucaridan Malacostraca. Crustaceana 25: 35-50.

Wills, M. 1998. A phylogeny of the recent and fossil Crustacea derives from morphological characters, pp. 189-209. In, R. A. Fortey and R. H. Thomas (eds.), Arthropod Relationships. Chapman and Hall, London.

Winckler, T. C. 1882. Carcinological investigation on the genera *Pemphix*, *Glyphea* and *Araeosternus*. Annals and Magazine of Natural History (5) 10: 133-149, 306-317.

Wong, J. M., J. L. Perez-Moreno, T.-Y. Chan, T. M. Frank, and H. D. Bracken-Grissom. 2015. Phylogenetic and transcriptomic analyses reveal the evolution of bioluminescence and light detection in marine deep-sea shrimps of the family Oplophoridae (Crustacea: Decapoda). Molecular Phylogenetics and Evolution 83: 278-292.

Yang, Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends in Ecology and Evolution 11: 367-372.

———, and B. Rannala. 2012. Molecular phylogenetics: principles and practice. Nature Reviews Genetics 13: 303-314.

Yeo, G. W., E. Van Nostrand, D. Holste, T. Poggio, and C. B. Burge. 2005. Identification and analysis of alternative splicing events conserved in human and mouse. Proceedings of the National Academy of Sciences of the United States of America 102: 2850-2855.

Zhang, D. X., and G. M. Hewitt. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247-251.

———, and ———. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12: 563-584.

Zwickl, D. J., and D. M. Hillis. 2002. Increased taxon sampling greatly reduces phylogenetic error. Systematic Biology 51: 588-598.

RECEIVED: 8 January 2015.
ACCEPTED: 5 June 2015.

APPENDIX

Table A1. A list of taxon names mentioned in this paper. 'Accepted' status indicates whether the name is currently accepted and is listed according to De Grave et al. (2009).

Taxon	Accepted	Taxonomic rank	Includes	Authority
Eurysternalia	No	Unranked	Achelata, Anomura, Brachyura	Dixon et al. (2003)
Gebiidea	Yes	Infraorder	Some representatives of the unaccepted Infraorder Thalassinidea	de Saint Laurent (1979)
Glypheidea	Yes	Infraorder	–	Winckler (1882)
Glypheoidea	Yes	Superfamily	–	Winckler (1882)
Heterochelida	No	Suborder	Caridea, Thalassinidea	Beurlen and Glaessner (1930)
Homarida	No	Infraorder	–	Huxley (1878)
Lineata	Yes	Unranked	Anomura, Brachyura, Thalassinidea	Ahyong and O'Meally (2004)
Meiura	Yes	Unranked	Anomura, Brachyura	Dixon et al. (2003)
Natantia	No	Suborder	Caridea, Penaeoidea, Stenopodidea	Boas (1880)
Paguroidea	Yes	Superfamily	–	Latreille (1802)
Palinura	No	Infraorder	Achelata, Polychelida	Latreille (1802)
Penaeidae	Yes	Family	–	Burkenroad (1963)
Penaeoidea	Yes	Superfamily	–	Rafinesque (1815)
Pleocyemata	Yes	Suborder	Achelata, Anomura, Astacidea, Axiidea, Brachyura, Caridea, Gebiidea, Glypheidea, Polychelida, Procarididea, Stenopodidea	Burkenroad (1963)
Polychelida	Yes	Infraorder	–	Scholtz and Richter (1995)
Procarididea	Yes	Infraorder	–	Felgenhauer and Abele (1983)
Reptantia	No	Unranked	Achelata, Anomura, Astacidea, Axiidea, Brachyura, Gebiidea, Glypheidea, Polychelida	Boas (1880)
Stenopodidea	Yes	Infraorder	–	Claus (1872)
Sterropoda	No	Infraorder	Eurysternalia, Thalassinida	Dixon et al. (2003)
Thalassinida	No	Infraorder	Axiidea, Gebiidea	Dixon et al. (2003)
Thalassinidea	No	Infraorder	Axiidea, Gebiidea	Latreille (1831)
Thaumastochelida	No	Infraorder	–	Bate (1888)