Lobsters (Achelata, Astacidea, Glypheidea, Polychelida) are a morphologically diverse and economically important assemblage of decapod crustaceans with a rich fossil record dating back ~360 Ma (Schram and Dixon 2004). These include the clawed, slipper, furry, coral, spiny, and deep-sea blind lobsters, crayfish and "living fossil" glypheids (Fig. 1). Although lobsters all have a superficially similar body plan, the morphological and ecological diversity is astonishing. Species range from clawed to clawless, sighted to blind, and ornamented to undecorated body forms inhabiting deep to shallow, cave to surface, and freshwater to marine ecosystems. A rich fossil record allows us to estimate the origins and diversification of major lineages while exploring the emergence of morphological and ecological diversity. Apart from their diversity, many types of lobsters have significant economic, conservation, and cultural values. Commercial fisheries and aquacultures rely on representatives within Achelata and Astacidea to bring in billions of dollars each year to an international economy for both human consumption and the aquarium trade (Steneck et al. 2011). Overexploitation of marine resources has led to increased conservation and management efforts on commercially important species, whereas others may be vulnerable due to

rarity in nature (e.g., Glyphidea). Lobsters are not only economically important, but also contribute to the cultural livelihood of many communities that celebrate these delicacies through festivals, crayfish boils, and other social activities (Swahn 2004; Jones et al. 2006).

Resolving the phylogenetic relationships among and within lobster-like decapods will aid in fishery and conservation efforts, biodiversity estimates, and an increased evolutionary understanding; however, current phylogenetic hypotheses are dynamic and under continuous debate (Fig. 2) (Scholtz and Richter 1995; Dixon et al. 2003; Ahyong and Meally 2004; Bracken et al. 2008; Bracken et al. 2009; Toon et al. 2009; Karasawa et al. 2013). Recent molecular studies using nuclear protein coding genes recovered lobster-like decapods as a monophyletic group with high support, whereas alternative studies using both molecules and/or morphology found otherwise. Although these studies are progressive in the understanding of lobster-like evolution, many lack significant taxon representation and/or combined evidence (molecules + morphology).

Lobsters have a rich fossil record due to the enhanced preservation of their exoskeleton. Interestingly, current species’ richness does not adequately reflect the fossil
history of lobster-like decapods and until recently, some major lineages were only known as fossil species. The infraorder Glypheidea was thought to be extinct until 1975, when *Neoglyphea inopinata* was described from the Philippines (Forest and de Saint Laurent 1975, 1976). Today, the glypheids are considered “living fossils,” with only two extant species that very closely resemble their extinct relatives. Their fossil record is the most numerous with over 256 fossil species discovered, with their highest diversity during the Jurassic (147 species, Schweitzer et al. 2010). Astacidea, with over 653 extant taxa contains ∼124 fossil species, whereas Polychelida and Achelata contain 38 and 140 extant and 55 and 72 extinct species, respectively. The integration of fossil and/or geological data with molecular and morphological data adds directionality to the major events and transitions that occurred during the evolution of a group.

Divergence time analyses are commonly implemented in modern-day evolutionary studies, and some have evaluated the performance and error associated with the various dating methods (Perez-Losada et al. 2004; Rutschmann et al. 2007; Battistuzzi et al. 2010; Inoue et al. 2010; Lukoschek et al. 2012). Our study is unique due to the large number of fossil priors (28 calibrations) and taxa (195 individuals) included, allowing us to examine how sampling density (number of fossils) and fossil placement (deep, middle, and shallow) influence divergence time estimates across our phylogeny. Here, we subsample our 28-fossil data set ("inclusive" data set) and based on five performance criteria we assess the performance of each subsampled data set (i.e., 1, 4, 8, 12, 16, 20, and 24 fossils). In addition to subsampling, and for the first time, we compare divergence time estimates using drastically different construction of age priors. Priors were constructed and/or included based on: (i) fossil age information (referred herein as “28 fossil” analysis) or (ii) fossil discovery, age, and extant species counts, a novel Bayesian Posterior Branching Process (referred herein as “BPBP”) approach (Wilkinson et al. 2011). The method proposed by Wilkinson et al. (2011) estimates fossil calibration dates using a model of speciation and fossil recovery (Wilkinson and Tavare 2009). The number of fossils discovered throughout history and their geological age distribution, in combination with extant species counts, are then used to update these prior distributions to give posterior distributions for the node ages given the information in the entire fossil record. Both approaches were implemented using Bayesian evolutionary analysis by sampling trees (BEAST) in order to assess congruence of current methods.

Here, we present the most comprehensive sampling of lobster-like decapods (94% of all extant genera), combining nuclear and mitochondrial sequences with morphological characters, to generate the largest phylogeny of the group to date. We examine the evolutionary relationships among infraorders, families, genera, and species and compare our findings with current hypotheses and taxonomic classifications. A rich fossil record allows us to estimate the divergence times and origins of major lineages. By implementing different dating approaches (28 fossil and BPBP), we can compare congruence across methods. In addition, we subsample the fossil data set to examine the influence of calibration density on divergence time estimation.

MATERIALS AND METHODS

**Taxon Sampling**

A total of 173 species (201 specimens) of lobster-like decapods (Astacidea, Achelata, Polychelida, Glypheidea) were included in the analysis (Supplementary
obtained from GenBank (Supplementary Appendix S1, Caridea, Brachyura, Anomura, Axiidea, Gebiidea) were from major decapod lineages (Dendrobranchiata, monophyly of the lobsters is contentious, 3–5 species http://dx.doi.org/10.5061/dryad.4v875). As the overall diversity and problematic groups. New sequences are directly from GenBank (Supplementary Appendix S1, deposited in GenBank, whereas others were obtained obtained directly from GenBank (Supplementary Appendix S1, http://dx.doi.org/10.5061/dryad.4v875). As the overall monophyly of the lobsters is contentious, 3–5 species from major decapod lineages (Dendrobranchiata, Caridea, Brachyura, Anomura, Axiidea, Gebiidea) were selected as outgroups. All outgroup sequences were obtained from GenBank (Supplementary Appendix S1, http://dx.doi.org/10.5061/dryad.4v875).

DNA Extraction, Polymerase Chain Reaction, and Sequencing

Genomic DNA was extracted from the gills, abdomen, pereiopod or pleopod using the Qiagen DNeasy Blood and Tissue Kit (Cat. No. 69582), QIAamp DNA Mini Kit (Cat. No. 51304) or QIAamp DNA Micro Kit (Cat. No. 56304). Gene regions were amplified with the polymerase chain reaction (PCR) using one or more sets of primers. Regions of the following genes were sequenced: 16S large ribosomal subunit (~350 bp, Crandall and Fitzpatrick 1996), 12S small ribosomal subunit (~400 bp, Buhay et al. 2007), cytochrome oxidase 1 (COI) protein coding gene (~650 bp, Folmer et al. 1994), 28S large ribosomal subunit (~2500 bp, Whiting et al. 1997; Whiting 2002; Palero et al. 2008), 18S small ribosomal subunit (~1800 bp, Medlin et al. 1988; Whiting et al. 1997, Apakupakul et al. 1999; Whiting 2012; Bracken et al. 2009), and protein-coding histone 3 (H3) (~350 bp, Colgan et al. 1998).

PCR amplifications were performed in 25–50 µL volumes containing 1 U of Taq polymerase (HotMaster, AccuPrime or REDTaq), PCR Buffer, 2.5 mM of deoxyribonucleotide triphosphate mix (dNTPs), 0.5 µM forward and reverse primer, and 30–100 ng extracted DNA. The thermal profile used an initial denaturation for 1 min at 94°C followed by 30–40 cycles of 30 s to 1 min at 94°C, 45 s to 1 min at 46–58°C (depending on gene region), 1 min at 72°C and a final extension of 7 min at 72°C. PCR products were purified using filters (PrepEase™PCR Purification 96-well Plate Kit, USB Corporation) and sequenced with ABI BigDye® terminator mix (Applied Biosystems, Foster City, CA). Cycle Sequencing reactions were performed in an Applied Biosystems 9800 Fast Thermal Cycler (Applied Biosystems), and sequencing products were run (forward and reverse) on an ABI 3730×l DNA Analyzer 96-capillary automated sequencer in the Brigham Young University (BYU) sequencing center.

Phylogenetic Analyses

Sequences were cleaned, edited, and assembled using Sequencher 4.8 (GeneCodes, Ann Arbor, MI, USA). To
check for pseudogenes, we followed suggestions by Song et al. (2008) by extracting DNA from tissue with high amounts of mitochondria (gill tissue), translating protein-coding sequences (COI, H3) to check for indels and stop codons, comparing sequences to closely related species and building individual gene trees to ensure similar topologies. Individual gene alignments were performed in MAFFT. As our data set contained multiple genes with conserved domains and large gaps (16S, 12S, 18S, and 28S), we used the “E-INS-i” option in MAFFT. For non-coding genes, Gblocks v0.91b was used to exclude regions of questionable positional homology (Castresana 2000; Talavera and Castresana 2007). Individual gene data sets were concatenated in MESQUITE 2.73 (Maddison and Maddison 2010) and partitioned in the final analyses. To help select the optimal partitioning strategy for our data set, Partition Finder (Lanfear et al. 2012) was implemented using AIC criteria. We explored both a 10-partition scheme (by gene and codon) and 6-partition scheme (by gene). In addition to PartitionFinder, we used the Bayesian Information Criterion (BIC) (Schwarz 1978) in MODELLTEST to determine the best model of evolution for each individual molecular data set (Posada and Crandall 1998).

Before concatenation individual gene/morphology trees were generated using randomized accelerated maximum likelihood (RAxML) (Stamatakis et al. 2005, 2007, 2008) and maximum parsimony methods to identify conflicting histories and identify contamination taxa (see below for detailed methods, Supplementary Appendices S4–S8; http://dx.doi.org/10.5061/dryad.4v875). Individual gene sets were concatenated and two final data sets were generated (i) molecular and (ii) molecular + morphology (combined).

The same gene set was used in both analyses. To explore congruence between concatenation and species tree methods, the program, iGTP (Chaudhary et al. 2010) was used to infer a species tree phylogeny using a relaxed molecular clock model, which assumes that there will be an appropriate number of constant character trees present. Three independent runs were performed (each consisting of 20 chains and nswaps =10) on each data set (molecular and combined). Each analysis ran for 50 000 000 iterations, which we thinned to every 1000th iteration. Trace plots were visually inspected to assess convergence, mixing, and stationarity in Tracer v1.4. (Rambaut and Drummond 2007). Once the split frequency in each analysis was 1% (reached after around 8 and 3 million generations for the molecular and combined analysis, respectively), a 50% majority-rule consensus tree was obtained from the remaining saved trees. Posterior probabilities (PPs) for clades were compared for congruence and then combined between individual analyses. Values >0.7 are presented on the BAY phylograms (Fig. 3a,b; Supplementary Appendix S10, http://dx.doi.org/10.5061/dryad.4v875). In iGTP four gene trees (mitochondrial, H3, 28S, and 18S) were used to estimate a species tree using the minimize deep coalescence model with the leaf-adding algorithm for 1000 hill-climbing heuristic searches.

The maximum likelihood (ML) analysis was conducted on the molecular data set(s) using RAxML (Stamatakis et al. 2005, 2007, 2008). We used a total of 400 searches for the best ML tree (200 from random, 200 from bootstrap starting trees). ML estimates were compared to determine the best tree and bootstraps were mapped on the resulting topology. Likelihood settings were consistent with the General Time Reversible model (GTR) with a γ-distribution. RAxML estimated all free parameters following a partitioned data set. Confidence in the resulting topology was assessed using non-parametric bootstrap estimates (Felsenstein 1985) with 1000 pseudoreplicates and values >60% are presented on the resulting phylogeny (Supplementary Appendix S10, http://dx.doi.org/10.5061/dryad.4v875). In addition to implementing a GTR GAMMA model in RAxML, we also explored the use of a site-heterogeneous CAT-based model in Phylobayes (Lartillot and Philippe 2004; Lartillot et al. 2009). Although CAT-based models have been used less frequently in nucleotide data sets when compared with amino acids, they have shown to be more effective at dealing with site heterogeneity in phylogenomic studies examining deep-level divergences. In Phylobayes, constant sites were removed from the alignment and two independent Markov chains were run with a total length of 25 220 cycles. After burn-in, a consensus tree was recovered with posterior probabilities noted on resulting phylogram (Supplementary Appendix S11, http://dx.doi.org/10.5061/dryad.4v875).

The Bayesian (BAY) analysis was conducted in MrBayes v3.1.2b (Huelsenbeck and Ronquist 2001) for the molecular data set and combined data set (molecular + morphological). For the morphological characters, we used the Markov k (Mk, Lewis 2001) model with equal state frequencies, combined with γ-distributed rates across sites. We implemented a “coding = variable” parameter to use the likelihood conditional on character variability, rather than the ordinary likelihood, which assumes that there will be an appropriate number of constant characters present. Three independent runs were performed (each consisting of 20 chains and nswaps =10) on each data set (molecular and combined). Each analysis ran for 50 000 000 iterations, which we thinned to every 1000th iteration. Trace plots were visually inspected to assess convergence, mixing, and stationarity in Tracer v1.4. (Rambaut and Drummond 2007). Once the split frequency in each analysis was 1% (reached after around 8 and 3 million generations for the molecular and combined analysis, respectively), a 50% majority-rule consensus tree was obtained from the remaining saved trees. Posterior probabilities (PPs) for clades were compared for congruence and then combined between individual analyses. Values >0.7 are presented on the BAY phylograms (Fig. 3a,b; Supplementary Appendix S10, http://dx.doi.org/10.5061/dryad.4v875). All ML and BAY analyses were performed on the Marylou5 Dell PowerEdge M610 computing cluster at Brigham Young University and/or the High Performance Computing Cluster (Panther) at Florida International University.

Divergence Time Analyses

We used two approaches to estimate the relative timing of divergence between lobster lineages. Both methods used a relaxed molecular clock model, which allows the rate of evolution to vary among lineages. Fossil observations were used to calibrate the molecular clock, thus separating rate and time and allowing the estimation of divergence times. An MCMC algorithm was used to sample from the joint posterior distribution of divergence times, branch rates, and substitution
Figure 3. (a, b) Combined Bayesian phylogram based on molecular (4982 characters) and morphological (190 characters) data. Thick vertical bars represent lobster-like infraorders or other decapod lineages (infraorders, suborders). Bayesian posterior probabilities \( P > 0.56 \) are noted above or below branches.
Figure 3. Continued
assigned based on original publications for the fossil. In
assignment were identified (Supplementary Appendix S12; http://dx.doi.org/10.5061/dryad.4v875). We
inconsistency with our combined total evidence
addition, an apomorphy-based diagnosis of each
calibration dates are estimated from the fossil record, using a
model of speciation, fossilization, and recovery to
formulate prior distributions for node ages (Wilkinson et al. (2011), differs in the way palaeontological
data are incorporated. Rather than using user-specified
priors based on the age of the fossil, fossil calibration
rates are estimated from the fossil record, using a
model of speciation, fossilization, and recovery to
formulate prior distributions for node ages (Wilkinson
and Tavare 2009). The number of fossils discovered
throughout history and their geological age distribution,
in combination with extant species counts, are then used
to update these prior distributions to give posterior
distributions for the node ages, given the information
in the entire fossil record. The benefit of using this BPBP
model is that it moves the subjective judgments further
from important quantities and enables the utilization of
more of the fossil record, thus letting data play a larger
role in the analysis. These posterior distributions are
then used as prior distributions in the molecular analysis
implemented in BEAST (Method 2 = BPBP model).

**Fossils, Time Calibrations, and Prior Construction**

**Method 1: 28-fossil age model.**—We included a total of
28 calibration points that represent lineages at both
deep, middle, and shallow nodes within our divergence
time analysis (Table 1; Fig. 4; Supplementary Appendix S12; http://dx.doi.org/10.5061/dryad.4v875). We
followed recommendations by Parham et al. (2012)
when justifying fossil placement and selection and
consulted with decapod paleontologists (co-authors Feldmann and Schweitzer) to help with assignments.
When possible, each fossil used for calibration was based
on a single specimen that could be unambiguously
assigned to a clade. In some cases the fossils were
assigned based on fragmentary specimens or direct
observation of conspecifics; nevertheless, diagnostic
characteristics (i.e., chela, pleonites, uropods) for each
assignment were identified (Supplementary Appendix S12; http://dx.doi.org/10.5061/dryad.4v875). In
addition, an apomorphy-based diagnosis of each fossil was performed and any fossil that showed
inconsistency with our combined total evidence topology is discussed (Supplementary Appendix S12; http://dx.doi.org/10.5061/dryad.4v875). Ages were
assigned based on original publications for the fossil. In
all cases, Feldmann and Schweitzer either described the
fossil specimen, directly observed individuals and/or
thoroughly reviewed the literature to justify inclusion
of the fossil in the study (Schweitzer et al. 2010; Table 1). For all fossil calibrations, we chose the oldest known
fossil representative and assigned this age to either
the crown (MRCA, most recent common ancestor) or
stem node (i.e., node preceding the crown node) based on the above mentioned apomorphy diagnosis
(Parham et al. 2012). In BEAST, an exponential prior
was implemented on all fossil calibrations, with the
offset values set to the minimum calibration age. This
distribution is suitable for modeling fossil calibrations,
because the probability decreases with a growing
discrepancy between estimated nodal age and the
age of the calibrating fossil (Ho 2007). The choice of
an exponential prior allows us to ascribe the greatest
probability within the age range associated with fossil
discovery, while avoiding a hard minimum age.

**Method 2: BPBP model.**—Unlike the previous approach,
the fossil calibration dates are estimated rather than
specified. The number of fossilized species (count data)
found during each epoch was recorded for each of
the major lineages (Astacidea, Achelata, Polychelida,
Glypheidea) (Table 2). If a fossil species was found more
than once in different epochs, it was recorded for each
epoch in which it was found. If one specimen of a
single species spanned multiple time intervals (epochs)
(i.e., the time and error intervals associated with its
discovery overlapped multiple epochs), that species
was accounted for randomly in one of the multiple
epochs that it spread. A random assignment of the
specimen allowed us to avoid biasing its placement in
the table. Extant species counts were taken from De
Grave et al. (2009). Time intervals of epochs and ages
followed the 2009 Geological Time Scale provided by the
Geological Society of America (Walker and Geissman
2009).

As interest lies in the lobster-like decapods divergence
dates, we assumed an evolutionary tree (Polychelida, [Achelata, [Astacidea, Glypheidea]]). We assigned \( \tau_1 \)
to be the date the (Achelata, Astacidea, Glypheidea)
split from the Polychelida, \( \tau_2 \) to be the date Achelata
split from the (Astacidea, Glypheidea), and \( \tau_3 \) to be the
date the Astacidea split from the Glypheidea. The two
oldest lobster-like fossils (Mississippian, Late Devonian;
Table 2) were included as stem fossils in the analysis,
because they share morphological characteristics that
allied them to other lobster-like decapods, but we could
not confidently assign them to a single infraorder. We
used a uniform prior distribution for \( \tau_1 \), setting the
minimum bound using the oldest fossil representative
(Palaeopalaemon newberryi), Late Devonian—Table 2),
and used an arbitrary large maximum bound of 685 Ma (i.e.,
\( \tau_1 \sim U(685,385) \) a priori). The posterior distribution
was found to be robust to the choice of maximum bound
(as long as a large value is used). To complete the
prior specification, we specified an ordering of node
Thereafter (see Equation (3) in Wilkinson et al. 2011). Commissuration rapidly and then slows or remains constant (Fig. 3a,b). The tree was assumed based on the phylogenetic analysis (Fig. 3a,b).

A logistic model of expected diversification was used, which is equivalent to assuming diversification commences rapidly and then slows or remains constant thereafter (see Equation (3) in Wilkinson et al. 2011). The three parameters in this equation were given the following priors: \( \rho \sim U[0.01, 0.4] \) so that diversity was expected to reach 90% of the final expected diversity between 10 and 800 Ma after time of origin; parameter \( y \sim U[0.001, 0.02] \) so that the expected modern diversity ranged between 100 and 2000 species; for \( k \) we specified 1/\( k \sim U[1, 20] \), so that the expected period of time for each species existed was between 1 and 20 Ma. Fossil preservation and discovery was modeled using a Poisson process model, with constant discovery rate per unit of time. The posterior mean fossil discovery rate was estimated to be 0.01, which equates to an average of one fossil find, per species, per 100 Ma. To fit the model to the fossil data in Table 2, we used an ABC rejection algorithm, using the metric given by Equation (5) in Wilkinson et al. 2011, with the same tolerance at

Note: Number corresponds to nodal placement as assigned in Figures 4.

<table>
<thead>
<tr>
<th>Taxonomy Species</th>
<th>Geological age (Ma)</th>
<th>Node</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outgroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natantia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suborder Dendrobanchiata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superfamily Penaeoidea</td>
<td>Aciculapoda napei (Feldmann and Schweitzer 2010)</td>
<td>Fammernian (Devonian), 354–364</td>
</tr>
<tr>
<td>Suborder Pleocyemata</td>
<td>Pinnacaris dentata (Garassino and Teruzzi 1993)</td>
<td>Late Triassic, Norian, 280–221</td>
</tr>
<tr>
<td>Infraorder Caridea</td>
<td>Paleaeopalaemon nebraski (Whitfield 1880)</td>
<td>Late Devonian, 354–370</td>
</tr>
<tr>
<td>Infraorder Axidea</td>
<td>Cunninghamia s.l. bonaparte (Eallon 1861)</td>
<td>Early Jurassic, Toarcian, 180–190</td>
</tr>
<tr>
<td>Infraorder Gebidea</td>
<td>Upogebia obscura (Von Meyer 1854)</td>
<td>Early Triassic, 242–248</td>
</tr>
<tr>
<td>Infraorder Anomura</td>
<td>Platypeltida akane (Chablas, Feldmann and Schweitzer 2010)</td>
<td>Late Triassic, Norian/Rhaetian, 206–221MYA</td>
</tr>
<tr>
<td><strong>Ingroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infraorder Glyphideida</td>
<td>Litogaster turnbullensis (Schaum 1971)</td>
<td>Early Triassic, 242–248</td>
</tr>
<tr>
<td>Infraorder Astacidea</td>
<td>Uncina ollerenshawi (Feldmann and Copeland 1988) and Uncina pacifica (Schweiger et al. 2003)</td>
<td>Early Jurassic, Pliensbachian, 180–200</td>
</tr>
<tr>
<td>Superfamily Nephropoidea</td>
<td>Pseudastacus pusillus (Van Straelen 1924 (imprint 1925))</td>
<td>Middle Jurassic, Bajocian, 169–176</td>
</tr>
<tr>
<td>Genus Homarus</td>
<td>Homarus truxensis (Stenzel 1945)</td>
<td>Early Cretaceous, Alban, 99–111</td>
</tr>
<tr>
<td>Genus Metanephrops</td>
<td>Metanephrops rosseus (T)</td>
<td>Late Cretaceous, Campanian, 71.3–83.5</td>
</tr>
<tr>
<td>Family Astaciidae</td>
<td>Paleoechinastacus australianus (Van Straelen 1928c)</td>
<td>Early Cretaceous, 99–144</td>
</tr>
<tr>
<td>Genus Austropotamobius</td>
<td>Austropotamobius lioptis Via 1971</td>
<td>Late Jurassic, Kimmeridgian, 151–154</td>
</tr>
<tr>
<td>Genus Astacus</td>
<td>Astacus multiannulatus (Bell 1863)</td>
<td>Early Cretaceous, Berriasian-Hauterivian, 127–144</td>
</tr>
<tr>
<td>Genus Pacificostacus</td>
<td>Pacificostacus chodermus (Cope 1871)</td>
<td>Miocene, 5.3–23.8</td>
</tr>
<tr>
<td>Family Parastaciidae</td>
<td>Paleoniscinatus owataniensis (Martin et al. 2008)</td>
<td>Early Cretaceous, 99–144</td>
</tr>
<tr>
<td>Genus Astacopsis</td>
<td>Astacopsis franklini (Gray 1845)</td>
<td>Pleistocene, 0.01–1.8</td>
</tr>
<tr>
<td>Genus Parribacops</td>
<td>Parribacops fordei (Feldmann and Pole 1994)</td>
<td>Miocene, 5.3–23.8</td>
</tr>
<tr>
<td>Genus Procambarus</td>
<td>Procambarus primarius (Packard 1889)</td>
<td>Early Cenozoic, 657</td>
</tr>
<tr>
<td>Infraorder Acheleta</td>
<td>Yunnanopalinura schnuriae (Feldmann et al. 2012)</td>
<td>Middle Triassic, Anisian, 241–247</td>
</tr>
<tr>
<td>Genus Barbus</td>
<td>Barbus vitiensis (Dana et al. 1852a)</td>
<td>Pleistocene, 0.01–1.8</td>
</tr>
<tr>
<td>Genus Parriacpaus</td>
<td>Parrhacaps crassatus (Forster 1884)</td>
<td>Eocene, 33.7–54.8</td>
</tr>
<tr>
<td>Genus Scylarides</td>
<td>Scylarides helveticus (De Angeli and Martin 1994)</td>
<td>Eocene, 33.7–54.8</td>
</tr>
<tr>
<td>Genus Scylaros</td>
<td>Scylaros unguum (Bohn 1922)</td>
<td>Miocene, 5.3–23.8</td>
</tr>
<tr>
<td>Family Palinuridae</td>
<td>Archopepalinura (Pina 1974)</td>
<td>Late Triassic, Norian, 280–221</td>
</tr>
<tr>
<td>Genus Jasus</td>
<td>Jasus flabelliger (Glaessner 1960)</td>
<td>Miocene, 5.3–23.8</td>
</tr>
<tr>
<td>Genus Justitia</td>
<td>Justitia vicina (Bosch et al. 2008)</td>
<td>Eocene, 33.7–54.8</td>
</tr>
<tr>
<td>Genus Linuparus</td>
<td>Linuparus carteri (Reed 1911)</td>
<td>Early Cretaceous, Aptian, 112–121</td>
</tr>
</tbody>
</table>
Figure 4. Divergence time chronogram using 28-fossil approach in BEAST. Fossil calibration points are indicated by numbers 1–28 embedded in black circles. Divergence time estimates (Ma) are noted adjacent to their respective nodes and blue nodal bars correspond to the 95% highest posterior density regions. Red vertical line divides the chronogram into three regions for fossil calibration placement; deep, middle, and shallow. Geological periods are superimposed onto the phylogeny and listed as follows: O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; T, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary. Colored taxa correspond to the lobster-like decapod infraorders. Orange = Polychelida, Green = Achelata, Brown = Glyphidea, Blue = Astacidea.
TABLE 2. A summary of fossil count data used in the BPBP model

<table>
<thead>
<tr>
<th>Epoch</th>
<th>k</th>
<th>Time at base of interval (Ma)</th>
<th>Fossil counts: Poly+</th>
<th>Fossil counts: Ach+</th>
<th>Fossil counts: Gly+</th>
<th>Fossil counts: Ast+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extant</td>
<td>0</td>
<td>0</td>
<td>383</td>
<td>795</td>
<td>655</td>
<td></td>
</tr>
<tr>
<td>Pleistocene</td>
<td>1</td>
<td>0.0</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Pliocene</td>
<td>2</td>
<td>0.3</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miocene</td>
<td>3</td>
<td>0.0</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Oligocene</td>
<td>4</td>
<td>3.3</td>
<td>7</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eocene</td>
<td>5</td>
<td>15.5</td>
<td>32</td>
<td>32</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Paleocene</td>
<td>6</td>
<td>65.6</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Late Cretaceous</td>
<td>7</td>
<td>99.6</td>
<td>140</td>
<td>137</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Early Cretaceous</td>
<td>8</td>
<td>143.5</td>
<td>86</td>
<td>85</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Late Jurassic</td>
<td>9</td>
<td>161.0</td>
<td>95</td>
<td>83</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Middle Jurassic</td>
<td>10</td>
<td>176.0</td>
<td>43</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early Jurassic</td>
<td>11</td>
<td>201.6</td>
<td>70</td>
<td>48</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Late Triassic</td>
<td>12</td>
<td>235.0</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Middle Triassic</td>
<td>13</td>
<td>243.0</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Early Triassic</td>
<td>14</td>
<td>251.0</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Late Permian</td>
<td>15</td>
<td>260.0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Middle Permian</td>
<td>16</td>
<td>271.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Early Permian</td>
<td>17</td>
<td>259.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Pennsylvanian</td>
<td>18</td>
<td>318.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Mississippian</td>
<td>19</td>
<td>344.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Late Devonian</td>
<td>20</td>
<td>345.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Pre-mid-Devonian</td>
<td>21</td>
<td>/</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Note: The table is divided into epochs with the time for each epoch given in millions of years (Ma). Poly, Polychelida; Ach, Acheata; Ast, Astacidea; Gly, Glypheidea.

all three nodes: $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.45$. For more details, see Wilkinson et al. (2011).

Analytic approximations for the posterior.—After experimenting with a range of distributions, we found that fitting independent skew-t distributions (Azzalini and Genton 2007) to each posterior gave the most acceptable parametric fit. The skew-t distribution is parameterized by four parameters: location, scale, shape, and the number of degrees of freedom. As we could not implement a skew-t distribution in BEAST, we fit each posterior to a lognormal (we could not implement a skew-t distribution in shape, and the number of degrees of freedom. As is parameterized by four parameters: location, scale, exponential distribution is parameterized by the mean, offset, standard deviation and the exponential distribution is parameterized by the mean and offset, all shown in Supplementary Appendix S13 (http://dx.doi.org/10.5061/dryad.4v875). A recent study has shown this distribution choice in BEAST produces similar posterior estimates and 95% highest posterior density regions (HPDs) when compared with using the skew-t distribution in mcmctree (Warnock et al. 2012). The lognormal distribution is parameterized by the mean, offset, standard deviation and the exponential distribution is parameterized by the mean and offset, all shown in Supplementary Appendix S13 (http://dx.doi.org/10.5061/dryad.4v875).

Bayesian Evolutionary Analysis by Sampling Trees

The data set was partitioned by gene and assigned independent models of evolution as in previous analyses. Branch rates are assumed to be independently drawn from a log normal distribution, and a Yule prior (constant lineage birth rate) is used to describe the rate of speciation. We recognize that there is a variety of models to consider when using relaxed dating methods. Studies that have compared uncorrelated and autocorrelated rates have recovered conflicting results, with some favoring uncorrelated models (Drummond et al. 2006), others favoring autocorrelated models (Lepage et al. 2007) and some favoring both (autocorrelated and uncorrelated) depending on the data set (Ronquist et al. 2012). We settled on using an uncorrelated model due to the taxonomic level of interest (focused at infraorder, family and genus level), age of the group (oldest fossil ~360 Ma), and evidence based on divergence time analysis, which measures rate autocorrelation on neighboring branches (Ho 2009; Drummond et al. 2006). In order to compare results, we used a starting tree (combined Bayesian consensus tree, see phylogenetic analysis, Fig. 3a,b) and removed the tree searching parameter from our BEAUTI xml file. As we need to provide a starting tree that satisfies the topological and temporal constraints associated with integrating fossil calibrations, we made the branch lengths proportional to timing (chronogram) rather than substitutions per site (phylogram), using the non-parametric rate smoothing algorithm in r8s (Sanderson 2003). We ran two independent runs for 350 000 000 iterations, which we thinned to every 10 000th iteration. Trace plots were visually inspected to assess convergence, mixing, and stationarity in Tracer v1.4. Convergence diagnostics were calculated using the CODA package in R (Plummer et al. 2006); the last 20 000 samples from each chain were used for further analysis with the previous samples discarded. Estimates of the mean divergence times with 95% HPD and posterior probabilities (represented as percentages) are noted on the chronogram (Fig. 4). All BEAST analyses were performed on the Marylou6 Dell PowerEdge M610 computing cluster at Brigham Young University or the High Performance Panther Cluster at Florida International University. For all analyses we examined the affect of priors on the posterior distributions when the data set is run without the data (empty alignment generated from BEAUTI).

Fossil Calibration Subsampling

We examined the range of calibrations used in previous analyses by searching the Web of Science using the keywords to “divergence dating,” refined by “fossil.” We narrowed the search to literature published after 2005 to capture studies using the methods herein (i.e., BEAST). We examined across 165 references for (i) number of fossils and (ii) number of terminals included, and (iii) divergence method used. Our literature search revealed that the 28 fossils used in this analysis represent one of the most fossil-rich data sets to date. This data set provides a unique opportunity to test the sensitivity of divergence time estimation to fossil calibration priors. In particular, we were interested in the effect of sampling...
density (number of fossil calibrations) on node age estimates. More specifically, at what point does the addition of fossil calibrations only slightly affect age estimation across a phylogeny? In this analysis, we randomly subsampled from within the set of 28 fossils using several different sample sizes, \( n \). We used fossil subsets of size \( n = 4, 8, 12, 16, 20, \) and 24. For each \( n \), 20 fossil subsets were randomly chosen for divergence time analysis. Analyses were also performed using a single calibration (\( n = 1 \)) for each of the 28 fossils. The same procedure was applied to each subsampled analysis. All subsampling analyses were performed using the method of Drummond and Rambaut (2007) and implemented in BEAST (Drummond and Rambaut 2007).

We compared the node age estimates for each subsampled analysis (148 total runs) against the estimates obtained from the 28-calibration analysis (herein referred to the “inclusive” analysis). For each of the internal nodes (194 total nodes), the posterior samples from the “subsampled” analyses were compared with the posterior from the “inclusive” analysis, resulting in 194 pairwise comparisons. We established five criteria to serve as a basis for our comparison: i) Does the mean of the “subsampled” analysis fall within the 95% HPD interval of the “inclusive” analysis? (ii) Does the mean of the “inclusive” analysis fall within the 95% HPD interval of the “subsampled” analysis? (iii) Are the population means of the “inclusive” and “subsampled” distributions significantly different from each other? (iv) Does the 95% HPD interval of the “inclusive” and “subsampled” distributions overlap? (v) What is the distance between the “inclusive” and “subsampled” distributions? Criteria 1, 2, and 4 are estimated by calculating the mean and 95% HPD interval of the node age distributions. Criterion 3 is calculated using a two-sample difference of means test; the estimates are significantly different from each other if 0 is not found within the 95% HPD interval. The fifth test, the two-sample Kolmogorov–Smirnov (K-S) test, is a non-parametric procedure for testing whether two samples were drawn from the same underlying continuous distribution. The K-S test statistic, \( D \), is the maximum value of the absolute difference between two empirical distribution functions (EDFs). A value \( D = 0 \) indicates that the EDFs of the two samples are equal, while \( D = 1 \) implies the maximum difference between the two distributions. The \( P \) value of the test can be calculated using the Kolmogorov distribution as the null distribution. A schematic diagram of these performance criteria can be found in Figure 6.

Output from the MCMC chains was analyzed for convergence using the CODA package in R (Plummer et al. 2006). Parallel chains are considered to converge when the potential scale reduction factor (Gelman and Rubin 1992) is close to 1; the final 20,000 samples from each parallel chain are combined for further analysis. The effective sample size for each parameter is estimated to evaluate mixing.
2014 H.D. BRACKEN-GRISSOM ET AL.—LOBSTER PHYLOGENETICS AND DIVERGENCE TIMES 469

nodes were those that did not have high support in one or all of the analysis.

Topologies and support values resulting from the combined and molecular analyses were highly congruent at the family, genus, and species levels. However, deep relationships (infra-infraordinal) were less resolved in the molecular phylogeny (Appendix S10; http://dx.doi.org/10.5061/dryad.dv875). Overall support across the tree increased with the addition of morphological evidence in the combined analysis (Fig. 3a,b). We recovered the lobster-like decapods as non-monophyletic (Fig. 3a,b). All families within lobster-like decapods were significantly supported as monophyletic (PP = 1 for Polychelidae, Palinuridae, Scyllaridae, Glypheidae, Enoplometopidae, Nephropidae, Astacidae, and Parastacidae) with the exception of Cambaridae with Cambaroides japonicus nesting within Astacidae (see “Discussion”). We recovered seven genera to be poly- or paraphyletic including Polychelus, Petractus, Scyllarus, Acantharctus, Fallicambarus, Faxonella, and Procambarus.

**Lobster Divergence Time**

All data were tested to assess if the individual data sets conformed to a strict molecular clock; a likelihood ratio test rejected the null hypothesis that the data followed a molecular clock (P < 0.05) for all individual gene data sets. All effective sample size (ESS) values were >200 for each individual run. Results from the empty alignment analysis showed that the fossil priors were not having a strong effect on our posterior divergence estimates.

**Method 1: 28-fossil age analysis.—**Runs were combined (40 000) in the final tree (Fig. 4). The origin of Decapoda was estimated in the Ordovician (∼447 Ma, Fig. 4). Polychelidae were the earliest lobster-like decapods to emerge (∼372 Ma), with the most recent common ancestor of the present-day genera Polychelus, Pontichelus, Willenoeisia, and Stereomastis placed within the Cretaceous (∼99 Ma). Achelata diverged from the remaining Reptantia (=crawling lineages) around 357 Ma, with Palinuridae and Scyllaridae splitting from one another within the Permian (∼254 Ma). The remaining lobster-like groups (Glypheidea and Astacidea) split from the other infraorders within the Carboniferous (∼339 Ma) and quickly radiated into two separate lineages around ∼318 Ma. Extant glypheid genera emerged only 24 Ma (Laurenteglyphea, Neoglyphea). All five extant families within Astacidae arose throughout the Permian (∼261 Ma: Parastacidae), Triassic (∼222 Ma: Enoplometopidae, Nephropidae) and Jurassic (∼161 Ma: Astacidae, Cambaridae). Northern hemisphere crayfish separated from the southern hemisphere crayfish around 261 Ma with subsequent family-level radiations of the northern hemisphere crayfish within the Jurassic.

**Method 2: BPPB model.—**The BPPB model estimates the origins of Decapoda to be 496 Ma with the earliest branching lobster group, Polychelida, estimated to originate 409 Ma (Table 3 and Fig. 5.). Achelata diverged soon thereafter ∼391 Ma and the youngest lobster-like lineages, Glypheidea and Astacidea split from one another ∼353 Ma (Table 3 and Fig. 5.).

**Fossil Calibration Subsampling Studies**

**Fossil subsampling.—**Using a minimum threshold of 100 samples, all parameters of interest converged for all runs (see Methods). We compared all 148 runs from the subsampled analyses with the 28-calibration (fossil age) run (herein referred to as “inclusive” estimate/run). Each comparison involves 194 pairwise comparisons of the posterior distribution at each internal node. This amounts to over 28 000 nodes compared, each utilizing five different criteria (see Methods). Each node was color coded to indicate whether it satisfied each of the four yes/no criteria (Criteria 1–4, see Methods and Fig. 6). Node color corresponds to the number of criteria passed or failed by that particular node; “warmer” colors indicate that the node failed more criteria, whereas “cooler” colors imply that the subsampled estimate is closer to the 28-calibration estimate (“passed” our criteria) (Fig. 7). The overall trend in this table demonstrates that, with the addition of fossils, each run is converging closer to the “inclusive” estimate of divergence time (Fig. 7). In some instances, a single run did well with the addition of a few fossils (e.g., Fig. 7; run 12, 1-calibration point). Most nodes passed the four criteria once we added 12 fossils to the data set and nearly all nodes passed the criteria once 20 fossils were added. Twelve fossils represented ∼72% of all fossils included in our “inclusive” analysis (∼28 fossils), whereas 20 fossils represented ∼71% of fossils included in our “inclusive” analysis. Our results indicate that some nodes (e.g., Fig. 7; 122, 123; 167, 168) performed poorly even with the.
FIGURE 5. Divergence time chronogram using BPBP approach in BEAST. Divergence time estimates (Ma) are noted adjacent to their respective nodes and blue nodal bars correspond to the 95% highest posterior density regions. Geological periods are superimposed onto the phylogeny and listed as follows: O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; TR, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary. Colored taxa correspond to the lobster-like decapod infraorders. Orange = Polychelida, Green = Achelata, Brown = Glypheidea, Blue = Astacidea.
The figure 6 shows a schematic example comparing the five performance criteria across three representative nodes. The table 4 summarizes the statistics for these criteria across all nodes, given the sample size (n = 1, 4, 8, 12, 16, 20, 24 calibration/s). The statistics include:

- Mean within “inclusive” HPD interval
- “Inclusive” mean within HPD interval
- Two-sample difference of means
- Overlapping HPD intervals
- Average K-S distance per node
- Percentage of nodes dated in tree (%)

Note: For the first four statistics, the number indicates the average percentage of nodes that pass each test. The K-S test statistic measures whether the values sampled by each run are drawn from the same underlying distribution. In all cases, a higher K-S number implies that the run is more different from the “inclusive.” The “min” and “max” represent the worst and best run for each n.

Results were further condensed to illustrate the relative performance (i.e., the five criteria described in fossil subsampling studies materials and methods) across all nodes, given the sample size (i.e., n = 1, 4, 8, 12, 16, 20, 24 calibration/s). A single fossil calibration (n = 1) performed the worst when addition of fossils. In many instances, these nodes were clustered close to each other on the phylogeny.
FIGURE 7. Graphical representation (heat matrix) of all fossil subsampling study results. All runs (n = 149) are sorted by number of calibrations (e.g., 1, 4, 8, 12, 24, 28). All nodes (n = 194) are assigned a success/failure score (coded by color) for each run. Each node was assigned a value between 0 and 4 according to whether that node passed (yes) or failed (no) the four performance criteria (1–4) defined in the fossil subsampling materials and methods. Node color (red, bright yellow, pale yellow, and white) corresponds to the number of criteria passed or failed by that particular node (e.g., red = failed all four criteria, bright yellow = failed three, pale yellow = failed two, white = failed none). Boxed cells represent the priors included in the run.
compared to the “inclusive” estimate and had the largest variation (min, max) among runs. Increasing the number of calibrations from 4 to 8 did not seem to markedly change the relative performance of the divergence time estimations; however, 12 calibrations showed noticeable improvement over 8 calibrations. Similarly, an increase from 12 to 16 calibrations did not seem to notably impact the relative performance. The 20-calibration data set was >95% similar to the “inclusive” run with an average 0.207 K-S distance, whereas the 24-calibration was >96% similar to the “inclusive” run with a K-S distance of 0.151 (Table 4).

**Dating Method Comparisons**

Two methods that implemented different approaches for prior calibration (fossil age model vs. BPBP model) in BEAST were compared with test congruence of current methods. Compared to the 28-fossil age run, the BPBP model estimated slightly older divergence ages for all infraorders, but slightly younger ages (or near identical) for all families (Table 3 and Fig. 4, 5). All but two major clades (Astacidae and Cambaridae) had overlapping HPDs (Table 3).

**Discussion**

Choice of Priors

Several studies have shown the choice of prior distributions and parameters have a direct effect on fossil calibration times (Lee and Skinner 2011; Lukoschek 2011; Warnock et al. 2012); however, our study is the first to compare the effect of priors using a commonly used fossil age approach (Method 1: 28-fossil age analysis) and the novel approach of Wilkinson et al. (2011) (Method 2: BPBP approach). The prior distributions implemented in each approach varied (i.e., exponential vs. exponential/lognormal) and parameters were ultimately chosen depending on (i) the fit of the data, and/or (ii) a model of fossil discovery.

Overall, the 28-fossil and BPBP approaches showed increased concordance in divergence time estimates from deep (infraordinal) to shallow (family, genus) splits (Figs. 4 and 5). Estimates for most major lineages within lobster-like decapods are congruent with overlapping HPDs (see results, Table 3; Figs. 5 and 6). This result is particularly interesting considering the two methods utilize the fossil record very differently (see Methods) and increases our confidence in fossil assignment, selection, and final divergence estimates. As done herein, we suggest reporting the combined ages derived from all approaches when estimating divergence times for major lineages.

Estimates for two lobster families, Astacidae and Cambaridae, were not congruent across the fossil age and BPBP approach. The MRCA of Astacidae was estimated to be much older in the 28-fossil age analysis when compared with the BPBP approach (151 Ma vs. 84 Ma). Based on fossil evidence, an astacid-like fossil, Astropotamobius illopsis Via 1971, was discovered in the Kimeridgian (∼151–154 Ma), which was driving the age estimate in the 28-fossil age analysis. As fossil ages are not assigned to specific nodes, but rather incorporated as a model that uses species discovery through time in the BPBP approach, this node was estimated to be much younger (∼84 Ma). The node that represents the MRCA of Cambaridae was also estimated to be younger in the BPBP analysis (80 Ma vs. 57 Ma, Table 3). One explanation for these drastically different estimates in the 28-fossil age analysis may be the interactions among priors on nearby branches to the node of interest (e.g., MRCA of Cambaridae) (see Fossil Confidence and Interactions below).

**Fossil Calibration Subsampling Studies**

In addition to testing prior choice on divergence estimates, we designed a fossil subsampling study to examine the effects of fossil density on divergence time estimates. We developed five criteria (see Methods) that allowed us to track the performance of the different fossil subsampling studies (Fig. 6). Our results suggest overlapping HPDs are the most conservative estimate with the alternate mean within the “inclusive” HPD interval being the most stringent (Table 4). Using the overlapping HPDs criteria, the inclusion of four calibration points resulted in 92% of the nodes appearing similar to the “inclusive” (28-fossil data set) run. In light of other performance tests, it becomes evident that the inclusion of four fossils may perform poorly when subject to more stringent performance criteria. For example, we see a steady decrease in performance (85.9%, 73.8%, 70.6%) when comparing the (i) 2-sample difference in means, (ii) “inclusive” mean within alternate HPD, and (iii) alternate mean with “inclusive” HPD, respectively (criteria 2–4, see Methods, Table 4). Although over 70% of the nodes in the 4-calibration run passed the most stringent performance criteria when compared with the “inclusive” run, an alarming result becomes evident when we take a closer look at the amount of variation associated with the performance criteria. As the performance criteria tests become more stringent, the amount of variation increases within runs (i.e., 20 runs for 1-calibration analysis), indicating that in many cases individual runs did much more poorly or better than our summary statistics report (Table 4). The largest amount of variation can be seen within the 1-fossil calibration runs, with the worst run (i.e., run 1) failing all four tests for all 194 nodes (Table 4; Fig. 7). In other words, the divergence time estimates for each node in this run showed no resemblance to our “inclusive” estimates, even under the most relaxed criteria (criteria 4, HPD overlap). On the contrary, several of the 1-calibration analyses performed very well; over 91% of the nodes in the best-performing 1-calibration (i.e., run 12) passed the most stringent test (Table 4, Fig. 7). Our results suggest

---

**References**


We were particularly interested in the following question: at what point does the addition of fossils only slightly affect our divergence time estimates? With the addition of fossils, accuracy improves (number of nodes that pass the criteria) and variation decreases, as expected (Table 4 and Fig. 7). However, when comparing performance across the table, it becomes clear that 4 and 8 calibrations yield similar results, as do 12 and 16 (Table 4). Although increasing the number of fossils may be a good strategy, these examples indicate some increases may result in similar performance and accuracy outcomes. It is important to note that the sheer number of fossils used in divergence dating analyses may not be as important as the proportion of fossil calibrations to OTUs. Our analysis contains 196 OTUs and with 28 fossil calibrations, allowing us to calibrate 14.4% of the nodes in the tree. Our results suggest only 10% of the nodes need to be calibrated (=20 fossils in our tree) to yield similar results (>95% similarity) to the “inclusive” estimate of divergence time (Table 4).

Interestingly, a recent study by Erwin et al. (2011), which examined the effect of fossil calibration sampling on deep animal divergences, came to similar findings. In that study, they performed a 50% calibration-jack-knife analysis in which they repeatedly added 50% of their fossils at random (total number of fossils in analysis =24, each jack-knife analysis =12). Results indicated that deleting 50% of their calibration points did not significantly affect their divergence time estimates. A 50% reduction in fossils (n = 12) across the 124 OTUs used in their study is equivalent to dating 10% of the nodes in their tree. Based on our study and in accordance with Erwin et al. (2011), we recommend including at least 1 fossil per 10 OTUs in divergence dating analyses that use fossil calibrations. We do recognize that this recommendation will be dependent on quality and availability of the fossil record and degree of rate heterogeneity among taxa, all of which need to be considered in any divergence dating study.

The K-S test statistic has more power to differentiate among analyses by their relative performance. We identified at least two factors that contribute to the discriminating power of this test. First, the K-S test uses the EDF of the posterior distribution, instead of point estimates (mean) and interval estimates (HPD interval) of the parameter value. Using sampled values from the posterior allows for [better] comparison between the underlying distributions. Second, the K-S test not only accounts for the location of the distribution, but is also sensitive to the shape of the distribution. Criteria 1–4 compare the locations of the “subsampled” and “inclusive” distributions while accounting for uncertainty using interval estimates. However, by only considering location, critical differences between estimates may be obscured. In this study, we found that the increased sensitivity of the K-S test provides a better understanding of the differences between and among different calibration sample sizes (Table 4). We witnessed similar trends for the K-S test as we did for the other four performance criteria. The 4- and 8-calibration runs yield similar K-S distances and variation (min/max values) whereas the 12- and 16-calibration runs yield similar trends. The inclusion of 20- and 24-calibrations shows a progressive improvement in K-S distances and lower variation (Table 4).

**Fossil Confidence and Interactions**

Perhaps equally, if not more important, than the number (percentage) of calibrations in the tree, is the selection of priors and fossil confidence. All calibration points used in divergence time analyses should be based on age data (fossil and geological) generated from reliable sources. Fossils that cannot be assigned to a particular clade due to poor preservation or incompleteness of the fossil record should not be included. Fossil calibrations should only be used when they can be confidently assigned to a node on the tree for a fossil age approach. Likewise, for the BPBP approach, priors should only be constructed from fossils that can be assigned to a major lineage and geological age (i.e., epoch in our study) with associated extant species count data. Recently, detailed protocols for justifying fossil calibrations have been published (Patham et al. 2012).

In many instances, calibration points were omitted from our phylogeny due to questionable nodal assignment. The importance of fossil confidence and the selection of priors are evident in our fossil subsampling studies. Our results indicate that fossils interact dynamically with each other, and the removal or addition of any one fossil can have extreme effects on the divergence estimations (Fig. 7 and Table 4). It is apparent that certain regions of the tree performed relatively poorly in the majority of the subsampled runs (Fig. 7). One explanation for these “poor” nodal estimations could be interactions among the fossil calibration priors, because the sampled age for an ancestral node imposes a maximum bound (=upper bound) on the age of the descendant node. Alternatively, missing or absent calibrations on nearby nodes may allow the node to estimate an incorrect age. An example of this is within one of our 1-calibration runs (e.g., run 12, Fig. 7) where a mid-tree fossil, representing the family Cambaridae, estimated 195 of the 196 nodes correctly (when compared to the “inclusive” run). The only node estimated incorrectly was the calibration node itself. In this example, Cambaridae is directly affected by fossils above and below this node, which drive its divergence time estimate.

**Fossil Placement**

A secondary goal that emerged from the fossil subsampling study was to evaluate fossil placement (deep, middle, and shallow) on divergence time estimates. We looked closely at the 4-calibration subsampled runs in an attempt to identify patterns that...
could predict relative performance. For example, if all calibration points are positioned at deep splits, this will have negative impacts (in performance) on shallow-level estimates? We partitioned the chronogram (deep, middle, and shallow; Fig. 4) and looked for correlations between the placement combinations (e.g., 3 shallow/1 middle; 3 middle/1 shallow; 1 deep/1 middle/2 shallow, 2 middle/2 shallow; Table 5). No clear pattern emerged relating placement combination to relative performance. Ironically, the best and worst runs from the 4-calibration study both contained 1 deep/1 middle/2 shallow fossils (Table 5). These results imply that correct age placement of fossil calibrations is currently being explored to further examine this issue.

### Evolutionary Relationships and Origins of Lobster-Like Decapods

Understanding phylogenetic relationships among and within lobster-like decapods is vital for unveiling the key factors responsible for their evolutionary success. For decades, the monophyly of lobster-like decapods has been under continuous debate. Much of the conflict comes from insufficient taxon sampling and alternative methods used to generate phylogenetic hypotheses (morphology vs. molecular data vs. combined). Owing to the rarity and inaccessibility of specimens of some taxa (i.e., Glyphheida), inclusion of all lobster lineages suitable for molecular and morphological analyses has been difficult. Furthermore, a dynamic classification has been under continuous debate. Much of the conflict comes from insufficient taxon sampling and alternative methods used to generate phylogenetic hypotheses (morphology vs. molecular data vs. combined).
monophyletic relationship between Astacidea, Achetata, and Polychelida (Tsang et al. 2008; Toon et al. 2009). Our current estimation of phylogenetic relationships based on six genes (mitochondrial and nuclear) and 190 morphological characters render the lobster-like decapods to be a non-monophyletic assemblage (Fig. 3a,b) in accordance with other studies based on total evidence (molecules and morphology).

**Polychelida**

Polychelids are commonly known as the deep-sea (exceeding 5000 m in Willemoesia, Galil 2000) blind lobsters due to the fact that all extant taxa possess reduced eyes and live in deep oceanic waters (Ahyong 2009). These lobsters are unique when compared with other reptant lineages in that all (five) or most (four) pairs of their legs are chelate. Present-day counts include 1 family, 6 genera, and 38 species with 3 families, 13 genera, and 55 species known as fossil representatives. Polychelids represent the oldest lobster lineage, dating back to the Devonian (~372/409 Ma, all dates reflect 28 fossil/BPBP approaches, Table 3; Figs. 4 and 5). Extinct fossil representatives differ most notably from present-day species in that many possessed well-developed eyes, suggesting a shallow water origin.

**Achelata**

Achelatans were the next to appear (~357/391 Ma), diverging into the two extant families, Palinuridae and Scyllaridae, around 250 Ma (Permian, Wuchiapingian). Achetata share numerous characters, most notably their phyllosoma larvae, that separate Achelata from all other Decapoda. In fact, the appearance of this specialized larval form may be traced back to the origin of achelatans, because fossil phyllosomas from the Upper Jurassic show that this type of larva has changed little in the last 200 million years (Polz 1971). Soon thereafter, Palinuridae split into two morphologically distinct lineages. Linuparidae, Justitia, Nupalirus, Palinustus, Pueraulisa, Palinurus, Polithy bras, and Panulirus form a clad known as the Stridentes, whereas Projasus, Sagmariasus, Jasus and the former synaxid genus Palinurellus form a clad known as the Silentes. The appearance of a stridulating organ, a complex sound-producing structure located on the antennal plate (found in Stridentes), may have provided these lobsters with advantages during predator escape (Lewis and Cane 1990), while expanding their communication abilities (Patek and Oakley 2003). Our phylogenetic analysis is consistent with the hypothesis that the stridulating organ provided Stridentes with an adaptive advantage promoting rapid diversification throughout the Jurassic (~200–160 Ma, Fig. 4). This time frame also coincides with the early breakup of Pangaea about 180 Ma. The ability to colonize new habitats during this event may have also facilitated the radiation and divergence of this group, as suggested in Astacidea (see below). Finally, the main clades found within Scyllaridae are in agreement with current taxonomy based on adult morphology (Holthuis 1985, 1991, 2002) and recent molecular studies (Yang et al. 2011). Interestingly, the close relationship between Parribacus and Evibacus is also supported by larval morphology (Palero, unpublished results).

**Glypheidida**

Glypheid lobsters were presumed extinct since the Eocene (33.9–55.8 Ma). However, in 1908, a chance collection off the Philippines recovered a single male specimen that went unnoticed for almost seven decades until Forest and De Saint Laurent described it as Neoglyphea inopinata in 1975. Since its discovery, additional specimens have been collected from the opposite side of the Equator (Timor Sea) along with a new genus and species, Laurentaglyphea neocaledonica (Richer de Forges 2006), from New Caledonia. Both living glypheid species (Neoglyphea inopinata and Laurentaglyphea neocaledonica) are represented in our phylogeny. These two lineages appear to have evolved relatively recently, splitting from each other only 24/22 Ma (Oligocene, Chattian; Figs. 4 and 5).

**Astacidea**

Astacidean includes the clawed lobsters and freshwater crayfish. They are divided into 4 superfamilies, 5 families, 44 genera, and 653 species. Our analysis includes all superfamilies, families, and 41 of the 44 genera. Higher level relationships have been well-resolved and well-documented within Astacidea (Crandall et al. 2000; Rode and Babcock 2005) and our results corroborate past findings. All families (Astacidea, Parastacidea, Enoplometopidea, and Nephropidea) are monophyletic with the exception of Cambaridae. The finding corroborates numerous morphological and molecular studies (Crandall et al. 2000; Rode and Babcock 2003; Porter et al. 2005; Brabant et al. 2006; Bracken et al. 2009; Breinholt et al. 2009) and we conclude that Cambaroides should be included within Astacidea, as treated herein (Table 1 and Fig. 4). Additionally, the family-level status of Thaumastochelidae (Thaumastochela, Thaumastochelopsis, Dinocetes) has been variously supported (Dixon et al. 2003; Ahyong and Meally 2004; Schram and Dixon 2004) or challenged (Tsang et al. 2008) in previous analyses. Although our analysis supports the monophyly of the thaumastochelid group, they are deeply nested within Nephropidea and are best considered as derived nephropids.

Current hypotheses have speculated that northern hemisphere crayfish (Astacidea) diverged from the southern hemisphere crayfish (Parastacidea) ~268/185 Ma following the breakup of Pangaea (Crandall et al. 2000). Our results suggest this split may have commenced ~68–61 Ma earlier (261/268 Ma), and was later facilitated by the break-up of the supercontinent,
when several speciation events occurred within these lineages (~215–140 Ma). Interestingly, the 28-fossil run and BBPP approach converged on very similar mean values for this split (261/268 Ma). Our divergence time analysis is the first to not place a maximum age limit on this node, and without this constraint, we estimated an older age for this split when compared with previous studies (Porter et al. 2005; Breinholt et al. 2009; Bracken et al. 2010). Our results show that present-day southern hemisphere crayfish represent much older lineages when compared with their northern hemisphere counterparts. These older speciation events may have been facilitated by vicariance, resulting from the break-up of Gondwana in the early Jurassic (~176–201 Ma), although this conclusion has been recently confounded by the discovery of a parastacid crayfish in British Columbia dating to the Eocene (55 Ma) (Feldmann et al. 2011). This hypothesis is supported in previous studies (Toon et al. 2010) and our divergence time analysis that shows the early-branching South American crayfish clade splitting from the remaining southern hemisphere crayfish ~215/205 Ma. Present-day genera found on Madagascar (Astacoidea) and New Zealand + Australia (sister clade to Astacoidea) originated soon thereafter (~209/199 Ma).

**SUPPLEMENTARY MATERIAL**

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.4v875.

**FUNDING**

This work was supported by the US National Science Foundation [EF-0531762, EF-0531603, EF-0531670], Research Fellowship Scheme of the Chinese University of Hong Kong [to L.M.T.]; the Research Grants Council of the Hong Kong Special Administrative Region [CUHK4419/04M to K.H.C.], the National Science Council and Academia Sinica [to T.V.C.], and Brigham Young University [to K.A.C.].

**ACKNOWLEDGMENTS**

We acknowledge A. Crosnier, R. Cleva, and L. Corbari of the Muséum National d’Histoire Naturelle, Paris for providing many samples for the present molecular analysis. We also thank all other colleagues and institutions that have help with specimen collection, accompanied us on research cruises, and or provided advice/input to the preparation of this manuscript.

**REFERENCES**


Dana J.D. 1852a. Parts I and II. Crustacea. U.S. Exploring Expedition During the Years 1838, 1839, 1840, 1841, 1842, under the Command of H.D. Bracken-Grisson et al.—lobster phylogenetics and divergence times 477


and morphological evidence: a comparison of several divergence
(Bergamo). Memorie della Società italiana di Scienze naturali e del
Museo civico di Storia naturale in Milano 21:5–34.
Plummer, M., Best, N., Cowles K. 2006. CODA: Convergence diagnosis
and output analysis for MCMC. R News 6:7–11.
Plattnkalken. Neues Jahrbuch für Geologie und Palaontologie.
Monatsshefte 8:474–488.
locus estimation of decapod phylogeny and divergence times. Mol.
Posada D., Crandall K.A. 1998. MODELEST: testing the model of DNA
Reed F.R. 1911. New Crustacea from the Lower Greensand of the Isle
of Wight. Geol. Mag. 8:115–120.
Richer de Forges B. 2006. Découverte en mer du Corail d’une deuxième
espèce de glyptéide (Crustacea, Decapoda, Glyptephyridae).
Molecular phylogeny of the Thalasinosidae based on nuclear and
mitochondrial genes. In: Martin J.W., Crandall K.A., Felder D.L.,
editors. Decapod crustacean phylogenetics. Boca Raton, FL: CRC
crayfish and some closely related nephropid lobsters. J. Crust.
Ronquist F., Kiepfstein S., Vihelmens L., Schultemeier S., Murray D.L.,
Kausly A.P. 2012. A total-evidence approach to dating with
fossils, applied to the early radiation of the Hymenoptera. Syst.
Biol. 61:973–999.
calibration uncertainty in molecular dating: the assignment of fossils
Sanderson M.J. 2003. r8s: inferring absolute rates of molecular
evolution and divergence times in the absence of a molecular clock.
Sanderson M.J. 2003. r8s: inferring absolute rates of molecular
evolution and divergence times in the absence of a molecular clock.
Sanderson M.J. 2003. r8s: inferring absolute rates of molecular
evolution and divergence times in the absence of a molecular clock.