FRONTIERS IN ZOOLOGY

RESEARCH Open Access

Cross-bracing uncalibrated nodes in molecular dating improves congruence of fossil and molecular age estimates

Prashant P Sharma* and Ward C Wheeler

Abstract

Introduction: The practice of molecular dating is an essential tool for hypothesis testing in evolutionary biology. Vagaries of fossilization and taphonomic bias commonly engender high uncertainty in molecular dating in taxonomic groups wherein few fossils can be unambiguously assigned to phylogenetic nodes. A recent and novel implementation in molecular dating, "cross-bracing", exploits gene duplications by formally linking calibrated node dates throughout the paralogous subtrees through hierarchical Bayesian models. An unexplored refinement of this method is cross-bracing nodes with unknown dates, in addition to calibrated nodes, such that all nodes representing the same cladogenetic events have linked priors. We applied such a refinement to molecular dating in chelicerates, one of the earliest groups of arthropods present in the fossil record, but whose molecular dating has been greatly inconsistent in the literature. We inferred divergence times using hemocyanin paralogs isolated from *de novo* assembled transcriptomic libraries, and multiple fossil calibrations.

Results: We show that extending cross-bracing to uncalibrated nodes greatly reduced variance in estimates of divergence times throughout the phylogeny, particularly for estimated diversification ages of spiders and scorpions, whereas cross-bracing calibrated nodes alone did not affect age estimation for uncalibrated, derived clades. Comparing ages inferred with extended cross-bracing to the fossil record, we observe smaller gaps between diversification and the first appearance of crown group fossils than have previously been inferred, particularly for spiders. Our dating indicates that scorpions have a Silurian origin, but diversification of extant lineages occurred near the Triassic-Jurassic boundary, falsifying previous inference of Permian diversification age based on extant distribution alone.

Conclusion: The significant reduction of variance in divergence time estimates upon extending cross-bracing to uncalibrated nodes makes this approach greatly suited for evolutionary inference in groups with poor fossil records, with particular reference to terrestrial arthropods.

Keywords: Divergence time estimation, Gene duplication, Fossils, Molecular clock, Paralogy, Prior

Introduction

The concept of the molecular evolutionary clock has been one of the most transformative ideas in molecular evolution [1]. Grounded upon the tenet that the amount of time elapsed since the last common ancestor of two homologous sequences is statistically proportional to the number of differences between sequences, molecular dating has become an invaluable tool for hypothesis testing in evolutionary biology [2-4]. Numerous evolutionary processes are informed by inference of molecular divergence times, such as quantifying major historical shifts in cladogenetic rate [5,6], falsifying biogeographic hypotheses [7], or identifying co-diversification events in diverse, symbiotic lineages [8].

In contrast to simple early approaches that relied upon assumptions of a global strict molecular clock or a series of local clocks, current methods in molecular dating deploy an array of sophisticated models and algorithms for inferring evolutionary rates over phylogenetic trees, including relaxed assumptions for rate variation across

* Correspondence: psharma@amnh.org
Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, 10024 New York, NY, USA

molecular phylogenies, use of fossil taxa as terminals in a phylogeny, and analysis of historical molecular sequence data [9-13]. These methodological advances facilitate comprehensive quantification of uncertainty in molecular dating, whose sources include analysis of molecular sequence data (e.g., inter-partition conflict) and the use of fossil taxa as calibration points. Sources of uncertainty engendered by fossil calibrators include estimating the age of each fossil, the accurate assignment of fossils in the phylogeny, the use of appropriate prior distributions for fossil calibrators, and potential conflict between multiple calibration points [14-16]. Consequently, even under relaxed clock methods, molecular dates often have very large variance. This phenomenon is especially acute for lineages with poor fossil records and *ipso facto* few available calibration points; the large size of the ensuing confidence intervals limits these clades' dispositive power in hypothesis testing.

A recent and promising approach to refining inference of divergence times leverages paralogy in gene families for reduction of uncertainty in molecular dating. These refinements consist of two strategies: (a) cross-calibration, wherein a fossil-calibrated node is assigned the same prior distribution at every incidence of that node in each paralog's subtree; and (b) cross-bracing, an extension of cross-calibration wherein priors of fossil calibrated nodes are linked using an additional hierarchical prior for node age equity [17]. These strategies were shown to provide significant gains in precision over dating with a single set of orthologous genes, as inferred from reduction of the 95% highest posterior density intervals (HPD) of surveyed nodes.

An unexplored refinement of the cross-bracing method is linking nodes with dates that are uncalibrated, but correspond to the same divergence events. Such a strategy could provide further gains over cross-bracing calibrated nodes alone. Here, we test this proposed extension of cross-bracing, using as a test case the hemocyanin gene family of chelicerate arthropods. Hemocyanins constitute the oxygen-transporting metalloproteins of various arthropods and mollusks [18-25]. Arthropod hemocyanins are composed of various subunits, each of which contains two copper moieties that reversibly bind oxygen molecules. In chelicerates, the presence of hemocyanins has been biochemically analyzed in horseshoe crabs (Xiphosura), spiders (Araneae), vinegaroons (Uropygi), tailless whip scorpions (Amblypygi), and scorpions (Scorpiones), all of which respire using book gills or book lungs, respiratory organs that are putatively homologous [26-28]. Most arachnids (terrestrial chelicerates) bear an archetypal 4×6 (24-mer) hemocyanin, whereas horseshoe crab hemocyanin consists of an 8×6 (48-mer) configuration [29]. Some variation in the 4×6 hemocyanin macromolecule of arachnids occurs in entelegyne spiders, wherein certain lineages have lost multiple subunits [26,27,30]. Hemocyanins do not occur in various terrestrial chelicerate orders that lack book lungs (e.g., Acariformes [mites]) or in sea spiders (Pycnogonida), although a recent study reported the presence of a single hemocyanin ortholog in the EST library of the pycnogonid *Endeis spinosa* [31]. The function of this ortholog in sea spiders is not presently known.

Hemocyanins and chelicerates together provide an ideal test case for two reasons. First, up to eight paralogs of hemocyanins have been reported in terrestrial chelicerates, proffering multiple targets for cross-bracing [31]. Second, the poor fossil record of terrestrial chelicerates has greatly impeded inference of evolutionary history through molecular dating. Especially problematic is the age of crown-group scorpions, which are difficult to distinguish from stem-group species due to poor preservation and/or disputes over the aquatic habitat of extinct forms [32-34]. The global distribution of scorpions and the common notion that they constitute "living fossils" has engendered the interpretation of Permian or older (>300 Ma) diversification of extant scorpions (i.e., current geographic distribution achieved via the breakup of Pangea).

Here we utilized the hemocyanin gene family to test the effect of cross-bracing uncalibrated nodes. We complemented existing sequence data of chelicerate hemocyanins with orthologs of several chelicerate species drawn from transcriptomic data sets. These species included a mesothele spider (*Liphistius malayanus*), a member of the lineage sister to all remaining spiders, and a buthid scorpion (*Centruroides sculpturatus*), a member of a cluster of families sister to the remaining scorpions (Iurida). These additions uniquely enable age estimation of the most recent common ancestor (MRCA) of both spiders and scorpions. We additionally report the sequence of a hemocyanin ortholog in a cyphophthalmid harvestman (*Metasiro americanus*), an apulmonate chelicerate.

Results and discussion

Phylogenetic placement of novel hemocyanin sequences corroborates consistency of phylogenetic signal

To an existing dataset of chelicerate hemocyanin sequences [31], we added orthologs of hemocyanins from four spiders (*Frontinella communis*, *Leucauge venusta*, *Liphistius malayanus*, and *Neoscona arabesca*); an amblypygid (*Damon variegatus*); a scorpion (*Centruroides sculpturatus*); and two previously unknown orthologs of the Atlantic horseshoe crab (*Limulus polyphemus*); Hc1/A and HcVI. Putative hemocyanins were extracted from transcriptomic assemblies using reciprocal best hits and orthology determined by phylogenetic placement. Together with existing sequences, all three major rami of spider phylogeny (Mesothelae, Mygalomorphae, and Araneomorphae) were represented, and the basal split in scorpion phylogeny between buthoid and non-buthoid scorpion families was

captured by the inclusion of *Centruroides sculpturatus* and *Androctonus australis* (Buthidae), and *Pandinus imperator* (Scorpionidae) [35].

The tree topology resulting from maximum likelihood analysis consistently recovered the monophyly of chelicerate orders represented by multiple specimens, namely, Xiphosura, Araneae, Scorpiones, and Amblypygi (Figure 1). Relationships among the paralog groups were identical to those recovered in a previous study [31]. Within the chelicerate orders, relationships of the three xiphosuran genera could only be assessed by one paralog (Hc1/A is the only hemocyanin sequenced for *Tachypleus tridentatus*), but are consistent with a recent multilocus phylogeny of Xiphosura using a separate set of molecular loci [36]. Among the better-sampled spiders, the expected basal relationship of (Mesothelae, (Mygalomorphae + Araneomorphae)) was obtained in only two out of six subtrees wherein all three lineages were represented. In one of the remaining four cases, a second paralog of HcD in *Neoscona arabesca* was recovered sister to the *Liphistius malayanus* HcD sequence. The placement of two hemocyanins from araneomorph spiders in the HcC

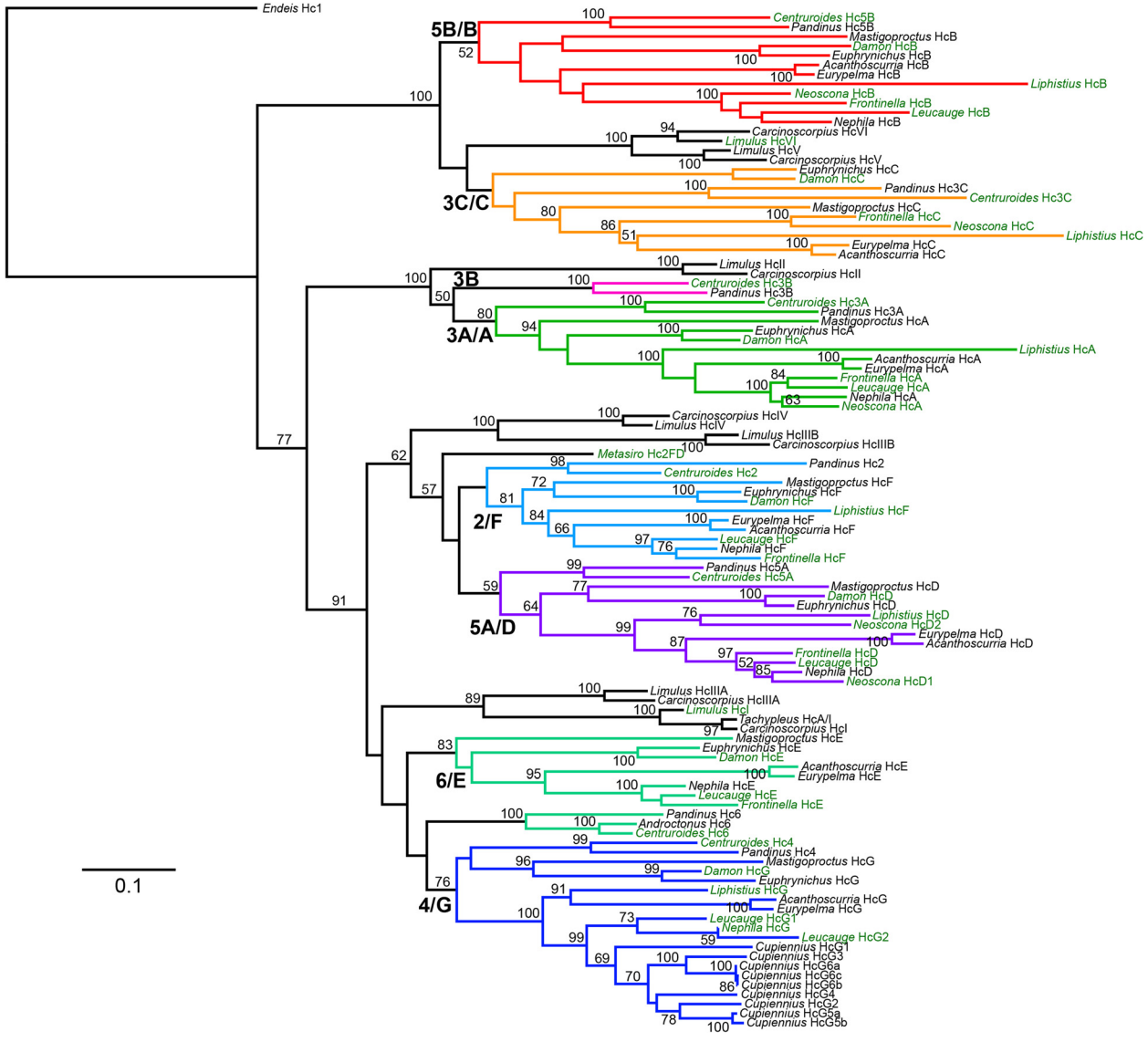

Figure 1 Maximum likelihood tree topology of chelicerate hemocyanin gene family. Tree topology with ln*L* = -49049.39. Colors in tree correspond to paralog groups. Numbers on nodes indicate bootstrap resampling frequencies. Green terminals indicate novel sequences. Pycnogonida: *Endeis spinosa*; Xiphosura: *Carcinoscorpius rotundicauda*, *Limulus polyphemus*, *Tachypleus tridentatus*; Opiliones: *Metasiro americanus*; Scorpiones: *Androctonus australis*; *Centruroides sculpturatus*, *Pandinus imperator*; Uropygi: *Mastigoproctus giganteus*; Amblypygi: *Damon variegatus*, *Euphrynichus bacillifer*; Araneae: *Acanthoscurria gomesiana*, *Cupiennius salei*, *Eurypelma californicum*, *Frontinella communis*, *Leucauge venusta*, *Liphistius malayanus*, *Neoscona arabesca*, *Nephila inaurata*.

subtree rules out loss of the HcC paralog in the common ancestor of Araneomorphae, previously suggested by [37]. The monophyly of Tetrapulmonata (Araneae + Uropygi + Amblypygi) was recovered in five of seven paralog subtrees (all except HcC and HcG). Intriguingly, a single hemocyanin ortholog was recovered from a search among transcriptomes of Opiliones; the cyphophthalmid *Metasiro americanus* harbors a hemocyanin sequence that diverged prior to the split between the Hc5A/D and Hc2/F paralogs. The implication of this discovery is discussed separately below.

Recovery of ordinal monophyly with high fidelity in each paralog's subtree indicates that individual hemocyanin paralogs exhibit a surprisingly consistent degree of phylogenetic resolution at the level of chelicerate orders, with some topological inconsistency in relationships within and between orders, relative to a reference topology based on 62 genes [38]. Many gene families do not retain such consistency in phylogenetic signal, owing to rate heterogeneity among paralogs and/or functional convergence of paralogs, as exemplified by the Hox genes [39-41].

We therefore utilized the hemocyanin gene family tree to infer the effects of linking priors in divergence time estimation, following the methods introduced by [17]. To facilitate precise calibration, we culled two out-paralog sequences of questionable orthology that engendered diphyly of species: *Acanthoscurria gomesiana* HcX and *Neoscona arabesca* HcD2. Other out-paralogs that rendered species paraphyletic (e.g., *Leucauge venusta* HcG1 and HcG2) and all in-paralogs were retained, as they did not affect the calibration of nodes. We began with cross-calibration and cross-bracing approaches wherein only calibrated nodes were constrained, following [17].

Cross-bracing calibrated nodes only does not affect age intervals of some uncalibrated nodes

We used three fossil-based calibrations to constrain (1) the split between Pycnogonida and Euchelicerata (applied once, at the root), (2) the origin of Amblypygi (applied to seven nodes, HcA-HcF), and (3) the origin of spiders (applied to six nodes, due to missing data for *Liphistius malayanus* HcE). In cross-calibration *sensu* [17], calibrations are thus repeatedly applied to nodes corresponding to the same divergences using the same prior distributions, but these priors are not formally linked. Results of this analysis, in accordance with previous efforts in molecular dating using concatenated hemocyanin sequences, retrieved large (≥75 Myr) 95% HPD intervals for several derived nodes corresponding to the divergences of, and within, chelicerate orders [31] (Figure 2A, Additional file 1: File S1). The largest HPD interval is that of basal scorpion divergence, with ages ranging from 50.0 Ma (paralog Hc3B) to 362.9 Ma (paralog Hc2).

We then implemented a cross-bracing strategy *sensu* [17], which incorporates an additional prior linking the ages of the calibrated nodes. Thus, 13 nodes (seven corresponding to origin of Amblypygi and six corresponding to origin of spiders) were linked. As expected, the resulting tree topology revealed reduced confidence intervals for calibrated nodes (Figure 2B, Additional file 2: File S2). No significant differences were observed in estimated median ages for either the calibrated ($p >> 0.05$; Figure 3A) or the uncalibrated ($p >> 0.05$; Figure 3D) nodes. For calibrated nodes, cross-bracing significantly decreased uncertainty (the 95% HPD interval) by 34-51% relative to the cross-calibrated analysis (the F-test was used for regression analysis; $p < 0.0001$) (Figure 3G).

However, cross-bracing did not decrease uncertainty in uncalibrated nodes; a near 1:1 relationship was observed in HPD intervals in cross-calibrated and cross-braced runs ($p = 0.94$) (Figure 3J). These results indicate that cross-bracing calibrated nodes alone can have limited effects in reducing uncertainty for derived nodes unavailable for calibration. This shortcoming is especially pronounced for chelicerates, due to the fragmentary nature of the terrestrial arthropod fossil record [34].

We therefore implemented an extension of cross-bracing to uncalibrated nodes. In this way, ages of nodes that represent the same speciation events were linked using additional priors, even if the ages of those nodes were unknown. We compared the resulting dates from extending cross-bracing (abbreviated "XCB") to counterparts from the cross-calibrated (abbreviated "CC") and original cross-braced (abbreviated "CB") runs.

Extending cross-bracing to uncalibrated nodes enhances precision in molecular dating

Median ages of the XCB analysis did not significantly change for either uncalibrated or calibrated nodes, in comparison to either CC or CB analyses (in all comparisons, $p > 0.95$) (Figures 2C, 3B-C, 3E-F, Additional file 3: File S3). By contrast, as measured from HPD intervals, the XCB analysis dramatically reduced uncertainty for uncalibrated nodes by 8-80% in comparison to CC ($p << 0.0001$; Figure 3K), and by 2-79% in comparison to CB ($p << 0.0001$) (Figure 3L). The variance in reduction of uncertainty stems from seven nodes corresponding to the divergence of Arachnopulmonata (=Scorpiones + Araneae + Uropygi + Amblypygi); while uncalibrated, this node lies between two calibrated nodes (origin of Xiphosura and origin of Araneae), whereas other nodes are not similarly bounded by fossil calibrations (Additional file 4: Table S1). Consequently, whereas XCB reduced uncertainty in the age of Arachnopulmonata by only 8-13% compared to CC, reductions of uncertainty by 27-80% were observed in the remaining nodes. Similarly, XCB achieved reduction of uncertainty by 26-79% in

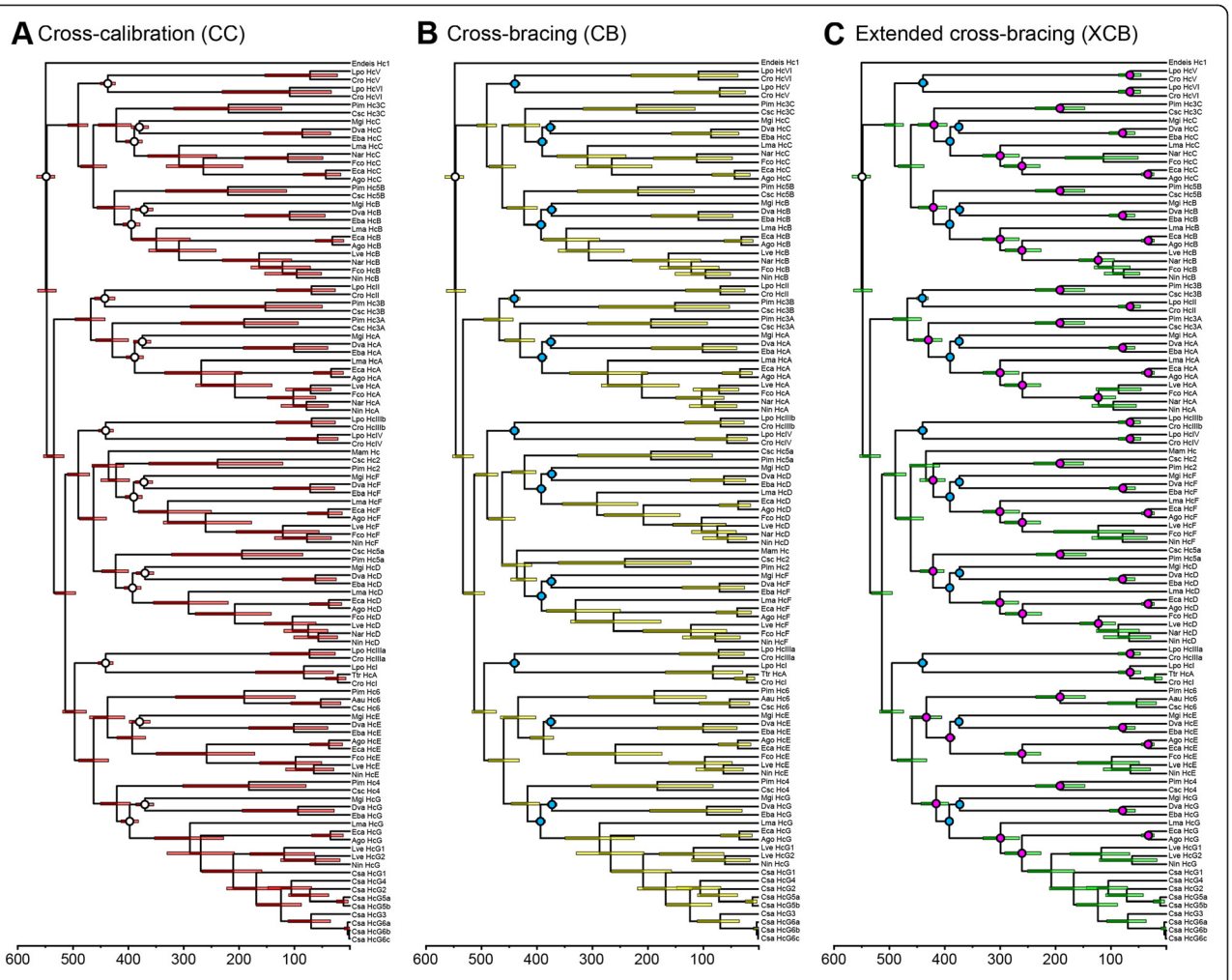

Figure 2 Chronograms of chelicerate hemocyanins. Under **(A)** cross-calibration, **(B)** cross-bracing of calibrated nodes, and **(C)** extended cross-bracing of calibrated and uncalibrated nodes. Horizontal axis indicates time in millions of years. Bars on nodes indicate 95% HPD intervals inferred by BEAST. White icons indicate calibrated nodes, blue icons indicate cross-braced calibrated nodes, and purple icons indicate cross-braced uncalibrated nodes.

uncalibrated nodes that did not correspond to Arachnopulmonata, compared to CB (Additional file 4: Table S1).

These observations suggest that large numbers of fossil calibrations bounding node ages of interest may reduce discrepancies observed between CB and XCB analyses. However, in the absence of numerous fossil calibrations, as in ordinal and intra-ordinal nodes within Chelicerata, comparison of the three dating methods demonstrates the effectiveness of XCB in estimating molecular dates with precision through leveraging replicated signal in gene families, even when calibrations are unavailable.

Cross-bracing closes gaps in fossil and molecular evolutionary age estimates

To gauge the accuracy and plausibility of *a posteriori* node age intervals generated by XCB, we compared divergence time estimates for three key groups—spiders, horseshoe crabs, and scorpions—to both the fossil record and published age estimates from multi-locus phylogenies. We discuss each in turn.

Of several previously published datasets estimating molecular divergence times in spiders, one was based on a single gene (EF-1γ; [42]) and a second was based on a concatenated analysis of spider hemocyanins [37]. Due to the patchiness and amplicon brevity of the latter data set, in addition to the present study's emphasis on basal chelicerate relationships, the sequences reported by Starrett et al. [37] were not included here. Both of these studies ([37,42]) utilized the Middle Devonian fossil *Attercopus fimbriunguis* (382.7 Ma [43])—a putative member of a spider stem-group [44]—to constrain the basal split of spiders, but neither data set included a representative of Mesothelae, the lineage sister to all remaining spiders. Apropos, both Ayoub et al. [42] and Starrett et al. [37]

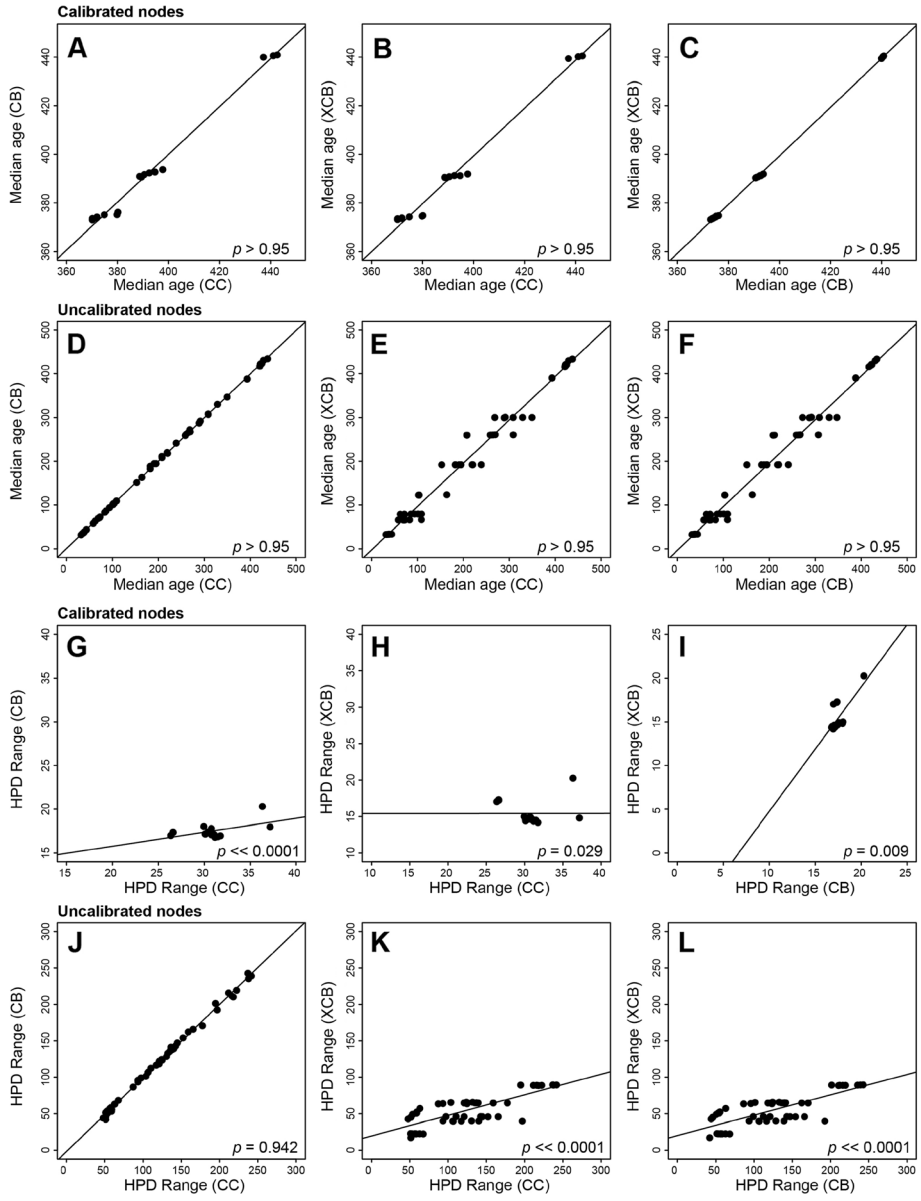

Figure 3 Effects of cross-bracing nodes on median age and variance. Pairwise regression analyses indicate that CC, CB, and XCB strategies do not affect inferred median ages of calibrated **(A-C)** or uncalibrated **(D-F)** nodes. **(G, H)** Compared to CC, CB and XCB analyses reduce 95% HPD intervals for calibrated nodes; less significant reduction in uncertainty is achieved by XCB compared to CB analysis **(I)**. For uncalibrated nodes, CB does not achieve significant reduction in 95% HPD intervals compared to CC **(J)**, but uncertainty decreases markedly for these nodes in XCB analyses **(K-L)**. *p* values indicate results of F tests.

recovered Devonian ages (i.e., the calibration itself) for the split between Araneomorphae and Mygalomorphae, a *de facto* overestimate stemming from the application of the fossil date to a derived (internal) node.

By contrast, a previous analysis also based on concatenated hemocyanin sequences of Chelicerata, and similarly lacking mesothele sequence data, obtained younger ages for the Araneomorphae-Mygalomorphae split [31]. This analysis employed a broad calibration prior for the split (240–382.7 Ma, with the floor chosen for the oldest crown-group member of Mygalomorphae and the ceiling for the age of *Attercopus*) and recovered a younger posterior node age distribution of 271 Ma (HPD: 254–288 Ma). The XCB analysis in the present study obtained a broadly overlapping posterior estimate for this split, 260 Ma (HPD: 227–292 Ma) (Figure 4). These dates strongly accord with the fossil record of opisthothele spiders. The oldest mygalomorph is dated to 240 Ma (*Rosamygale grauvogeli* [45]), and the earliest araneomorphs date to ca. 225 Ma [46]. The appearances of

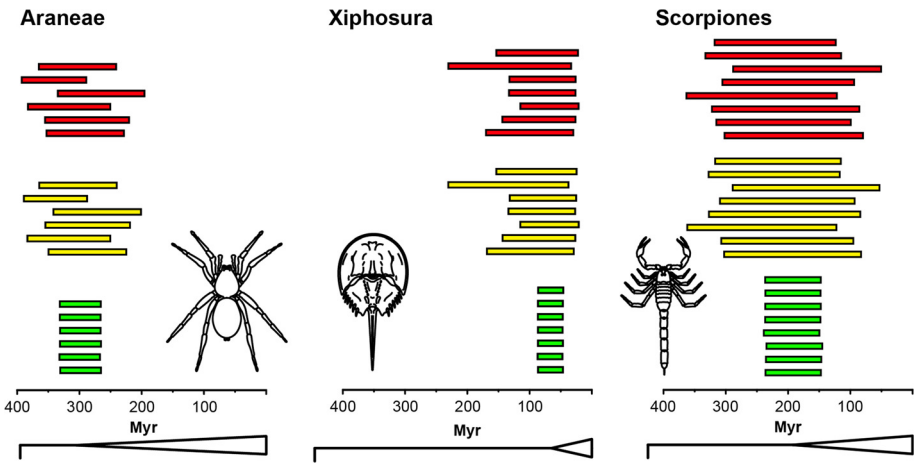

Figure 4 Inferred evolutionary history of key chelicerate groups. Comparison of 95% HPD intervals for diversification of spiders (left), horseshoe crabs (middle), and scorpions (right). 95% HPD intervals are shown for CC (red), CB (yellow), and XCB (green) analyses. Icons below axes indicate estimated temporal gaps between median ages of origin and diversification.

both opisthothele spider lineages shortly after the estimated origin of araneomorph-mygalomorph split (ca. 260 Ma) suggest that gaps between fossil and molecular dates may have been overestimated for spiders. The MRCA of Araneae was estimated in the XCB analysis to be 300 Ma (HPD: 266–331 Ma), a crown-group age consistent with the earliest known appearance of Mesothelae in the Upper Carboniferous (Stephanian [47]), as well as the frequency of Carboniferous diversifications in numerous terrestrial arthropod groups [48-54].

Another two recent studies of arachnid relationships at the transcriptomic scale dated the phylogeny of Araneae and included an exemplar of Mesothelae [55,56]. However, in the first case, Bond et al. [55] dated the tree with 128 loci in the absence of any outgroups (a pseudoscorpion, a tick, and a water flea used in phylogenomic analyses were culled from the dating analysis), and did not include closely related taxa such as amblypygids or uropygids. Their dating resulted in untenably and implausibly large age estimates for nodes at the base of Araneae. For example, a confidence interval of ca. 250–525 Ma was obtained for the MRCA of spiders—an inexplicable result, given that Bond et al. had specified that the ceiling of spider divergence should not exceed the arbitrary value of 400 Ma and the floor of that date should not be below 300 Ma [55]. Similarly, both araneomorphs and mygalomorphs were constrained to have a maximum age of 386 Ma (based on the age of *Attercopus*) in that study, but ceilings of the confidence intervals of both clades exceeded 400 Ma [55]. This outcome is indicative of procedural and/or algorithmic error, but was never discussed by Bond et al. [55].

By contrast, a phylogenomic analysis by Sharma and Giribet [56] focusing on the internal dating of Opiliones (based on 3,644 loci) also included among the outgroups exemplars of all three major spider lineages, Amblypygi, and Uropygi, and employed the age of the earliest Mesothelae as a minimum age constraint for spider divergence (in addition to separate calibrations within Opiliones). In that study, the age of spider diversification was estimated at 325–339 Ma in two Bayesian analyses (95% HPD interval: 305–387 Ma, across both analyses), a result very consistent with the one obtained in the present study. The congruence of results from these very disparate data sets [31,56] reinforces the tenet that proper algorithmic treatment of fossil taxa is far more important for molecular dating than quantity of sequence data [56].

Only a single study has inferred internal divergence dates in Xiphosura [36]. However, in that work, the basal split of the four extant horseshoe crab species was itself calibrated, based on the fossil *Mesolimulus walchi* and the inference of basal divergence driven by the opening of the Atlantic Ocean ca. 130–150 Ma. While dates estimated by CC and CB analyses encompassed the 120–160 Ma interval estimated by Obst et al. [36] for this node, the XCB analysis recovered estimates of 66 Ma (HPD: 47–87 Ma) (Figure 4). These dates indicate a younger Late Cretaceous estimated diversification of extant horseshoe crabs. We add the caveat that the sampling of *Tachypleus* is limited to a single paralog (HcA/1), which may undersample potential rate variation within Xiphosura. In either case, both the results of the present study and those of Obst et al. [36] suggest a prolonged (>300 Myr) gap between origin and diversification of extant Xiphosura, underscoring the characterization of this lineage as an evolutionary relict.

The crown-group age of scorpions is one of the more challenging problems in chelicerate paleontology. In contrast to many other arthropod orders, scorpions have a rich Paleozoic fossil record with over 80 species, and

phylogenetic analyses of the group indicate that extant scorpions (Orthosterni) constitute a small branch of a once-diverse assemblage [32,33]. Many of these early fossils are contended to be aquatic, whereas all extant scorpion species are terrestrial [34]. *Bona fide* crown-group fossils that can be placed within extant superfamilies do not exceed the Cretaceous in age [57], and questionable crown-group species have been described from the Early Triassic [34,58]. Furthermore, in the absence of molecular dates, some workers have inferred a divergence in the Permian or before, based upon the global distribution of extant scorpions and presumed mechanism of variance driven by Pangean breakup. Indeed, many scorpion lineages exemplify temperate Gondwanan distributions (e.g., Bothriuridae [59]), implying a minimum age of these lineages in the Late Jurassic and diversification coincident with supercontinental breakup.

To our knowledge, the only molecular dating available for the basal split of scorpions (between buthids and allies, and the remaining scorpions) was conducted by Rehm et al. [31] and Sharma and Giribet [56]. In the former study, as only a single buthid sequence (*Androctonus australis* Hc6) was available, this split was not as well represented as the spider divergences in that dataset. An age of 221 Ma and HPD of 107–355 Ma were inferred for extant scorpions from that study, too imprecise to be dispositive of hypotheses concerning scorpion biogeographic origins. Sharma and Giribet [56] also included two scorpions (a buthoid and a scorpionoid) as outgroups in a phylogenomic dating of Opiliones, but obtained markedly different dates for the MRCA of scorpions: 182 Ma under one model and 301.4 Ma under another (confidence intervals spanning 61 to 356 Ma across both analyses). In the present study, even with the inclusion of all hemocyanin paralogs of the buthid *Centruroides sculpturatus*, CC and CB analyses recovered similarly large HPD intervals for the age of scorpions, indicative of significant rate heterogeneity in scorpion hemocyanin subunits.

Propitiously, scorpions comprise the most opportune target for refinement by cross-bracing because they bear the greatest number of hemocyanin subunits among chelicerates (eight paralogs occur in scorpions, in contrast to seven in most tetrapulmonate arachnids and horseshoe crabs). The XCB analysis obtained the age of 192 Ma (HPD: 147–236 Ma) for the crown-group age of Scorpiones, one of the most significant reductions in HPD range in our dataset (Figure 4). Intriguingly, these results suggest a 200-Myr gap between origin and diversification of extant scorpions, comparable to, but less extreme than, xiphosuran phylogeny. Surprisingly, XCB dating rejects a Permian age of extant scorpions, but is consistent with Gondwanan vicariance, as Gondwana began to fragment ca. 180–165 Myr. This implies that a significant aspect of the global distribution of scorpions must be attributable to dispersal, not Pangean vicariance. Without additional sampling of scorpion species, it is not presently feasible to infer whether putatively Gondwanan families like Bothriuridae were present by the Late Jurassic or diversified later and dispersed to Gondwanan landmasses. While the HPD interval for scorpion diversification is loosely consistent with the crown-group membership of such Early Triassic (245–251 Ma) fossils as *Protobuthus elegans* and *Gallioscorpio voltzi*, the placement of these species is in dispute and awaits further investigation [34,58].

Taken together, the dates obtained by XCB analysis suggest that large gaps between the fossil record and previous divergence time estimates (either based on molecular dating or inferred from biogeographic patterns) may have been overestimated. Ages of extant chelicerate orders accord more closely with fossil dates that previously presumed, suggesting that the chelicerate terrestrial fossil record may be better reflective of historical divergence times than previously thought.

Incidence of hemocyanins in apulmonate chelicerates

Despite topological instability among basal chelicerate lineages, it is generally accepted that Xiphosura and Arachnida (terrestrial chelicerates) are monophyletic sister taxa [38]. Given this tree topology, occurrence of multiple hemocyanin subunits in horseshoe crabs and Arachnopulmonata (=Scorpiones + Tetrapulmonata) implies that several hemocyanin subunits were present in the common ancestor of the Eucheliceratа and have subsequently been lost independently in apulmonate chelicerate orders, which respire through a tracheal respiratory system. Contingency of subunit loss on physiology is supported by the observation that many derived spider species have lost most hemocyanin subunits and/or undergone duplications of remaining paralogs (e.g., the *g* paralogs of *Cupiennius salei*; absence of hemocyanin in *Dysdera* sp. [30,31]) Accordingly, biochemical assays have not identified hemocyanins in Pycnogonida, Solifugae, or Acariformes ([26,27,60]). We note that hemocyanins are also not observed in the genomes of the mite *Tetranychus urticae* (Acariformes) or the tick *Ixodes scapularis* (Parasitiformes).

Opiliones (harvestmen) constitute a curiosity in this regard. All Opiliones bear a tracheal respiratory system, and should therefore lack hemocyanins. In a review of harvestman functional morphology, Shultz [61] indicated that harvestmen constitute unusual apulmonate arachnids in that they have hemocyanin (citing Markl et al. [27]), and that the respiratory system of harvestmen may therefore constitute a “tracheal lung”, i.e., a system separate from that observed in such lineages as Solifugae or Acari. Oddly, both Rehm et al. [31] and Burmester [62] reported the absence of harvestman hemocyanins, citing the same source (Markl et al. [27]). The source in

question in fact examined a single harvestman species, *Leiobunum limbatum*, and reported dodecameric hemocyanins composed of two subunit types (A and F subunits) based on immunochemical analyses [27]. Rehm et al. [31] later argued that the protein in question may instead be a vitellogenin-like di-tetrameric protein, not a harvestman hemocyanin, though experimental data were not shown in support of this contention. Moreover, Rehm et al. [31] did not recover any hemocyanin sequences from an unpublished transcriptome of the harvestman *Phalangium opilio*.

To resolve this discordance in the literature with new empirical data, we searched for hemocyanin sequences in the transcriptomic libraries of 14 Opiliones species spanning all suborders [56,63,64]. We identified a single copy of hemocyanin in the transcriptome of *Metasiro americanus*, a member of the suborder Cyphophthalmi (the lineage sister to the remaining suborders [64,65]). Phylogenetic placement of this hemocyanin sequence, tentatively named "Hc2FD", indicates that it diverged prior to the split between the paralogs Hc5A/D and Hc2/F. This placement suggests the intriguing possibility that diversification of some hemocyanin paralogs may have occurred uniquely in the ancestor of Arachnopulmonata, not in the common ancestor of all arachnids. This is methodologically significant because previous analyses have assumed that hemocyanin paralogs of Xiphosura and Arachnopulmonata are directly orthologous, which would justify concatenation approaches [31]. Concatenation has proven challenging, however, because clear orthologous relationships are not supported in the hemocyanin phylogeny (e.g., Hc3B in scorpions; clustering of xiphosuran Hc1 + Hc3A), and has required such workarounds as alternating orthology assignments and replicating the same sequence many-fold in the concatenated matrix (see [31]). A notably singular advantage of cross-calibration and cross-bracing techniques, beyond those elucidated by Shih and Matzke [17], is that orthology assignment is not required *a priori*, in contrast to concatenation methods.

The placement of the *Metasiro americanus* hemocyanin sequence corroborates the sister relationship of scorpions to tetrapulmonates [38], and suggests that the 4×6 hemocyanin subunit configuration is synapomorphic for Arachnopulmonata (with secondary losses of some subunits in some entelegyne spiders), not a plesiomorphy retained since the last common ancestor of arachnids. However, we add the caveat that biochemical and functional analysis of the *Metasiro americanus* hemocyanin is required to assess whether it constitutes a true hemocyanin subunit or a runaway gene with a novel function. We further note that the discovery of a hemocyanin in this apulmonate arachnid may have been made uniquely possible by sequencing a large number of developmental stages for this species, as evidenced by numerous sequences in its transcriptome with gene ontogeny pertaining to developmental processes [63], in contrast to larger libraries based on one or two developmental stages [64].

The scenario proposed herein for the evolutionary history of hemocyanins (Figure 5) should be further corroborated by sequencing additional developmental transcriptomes of such apulmonate arachnids as non-cyphophthalmid Opiliones, Ricinulei, Solifugae, and Pseudoscorpiones. A genome of a Cyphophthalmi species such as *Metasiro americanus* would also greatly refine inference of how many hemocyanins have been retained by this lineage (in contrast to such chelicerate genomes as those of *Tetranychus urticae* and *Ixodes scapularis*, which retain no hemocyanins at all), and whether these are still transcriptionally active or have become pseudo-genes.

Conclusion

Extension of the cross-bracing strategy to uncalibrated nodes in molecular dating greatly reduced uncertainty in divergence time estimation for a chelicerate hemocyanin dataset. The dates recovered by this analysis suggest smaller gaps between fossil and molecular age estimates than previously inferred. We showed that crown-group spiders diversified ca. 300 Ma, whereas a young, ca. 200 Ma age was recovered for the basal split of scorpions. Phylogenetic placement of a hemocyanin sequence from an apulmonate arachnid suggests that a 4×6 hemocyanin subunit configuration is synapomorphic of Arachnopulmonata (=Scorpiones + Tetrapulmonata).

Materials and methods

Identification of hemocyanin orthologs

Hemocyanin sequences were identified using reciprocal best BLAST hit searches. Published translated peptide sequences of *Carcinoscorpius rotundicauda*, *Pandinus imperator*, and *Eurypelma californicum* were used simultaneously to identify hemocyanins in transcriptomes of the following species: *Limulus polyphemus* (Xiphosura; Sharma et al. *in press*), *Liphistius malayanus* (Araneae, Mesothelae, Liphistiidae; Sharma et al. *in press*), *Frontinella communis* (Araneae, Opisthothelae, Linyphiidae; Sharma et al. *in press*), *Leucauge venusta* (Araneae, Opisthothelae, Tetragnathidae; Sharma et al. *in press*), *Neoscona arabesca* (Araneae, Opisthothelae, Araneidae; Sharma et al. *in press*), *Damon variegatus* (Amblypygi; Sharma et al. *in press*), and *Centruroides sculpturatus* (Scorpiones; Sharma et al. *in press*). These transcriptomes were accessioned in the NCBI Sequence Read Archive (accession numbers provided in Additional file 5: Table S2). Hemocyanin sequences were added to the chelicerate hemocyanin data set of Rehm et al. [31] with the following modification: the *Acanthoscurria gomesiana* HcX sequence, which is of unknown origin and orthology, and has a highly divergent sequence, was culled from the dataset. Previous sequences of the horseshoe

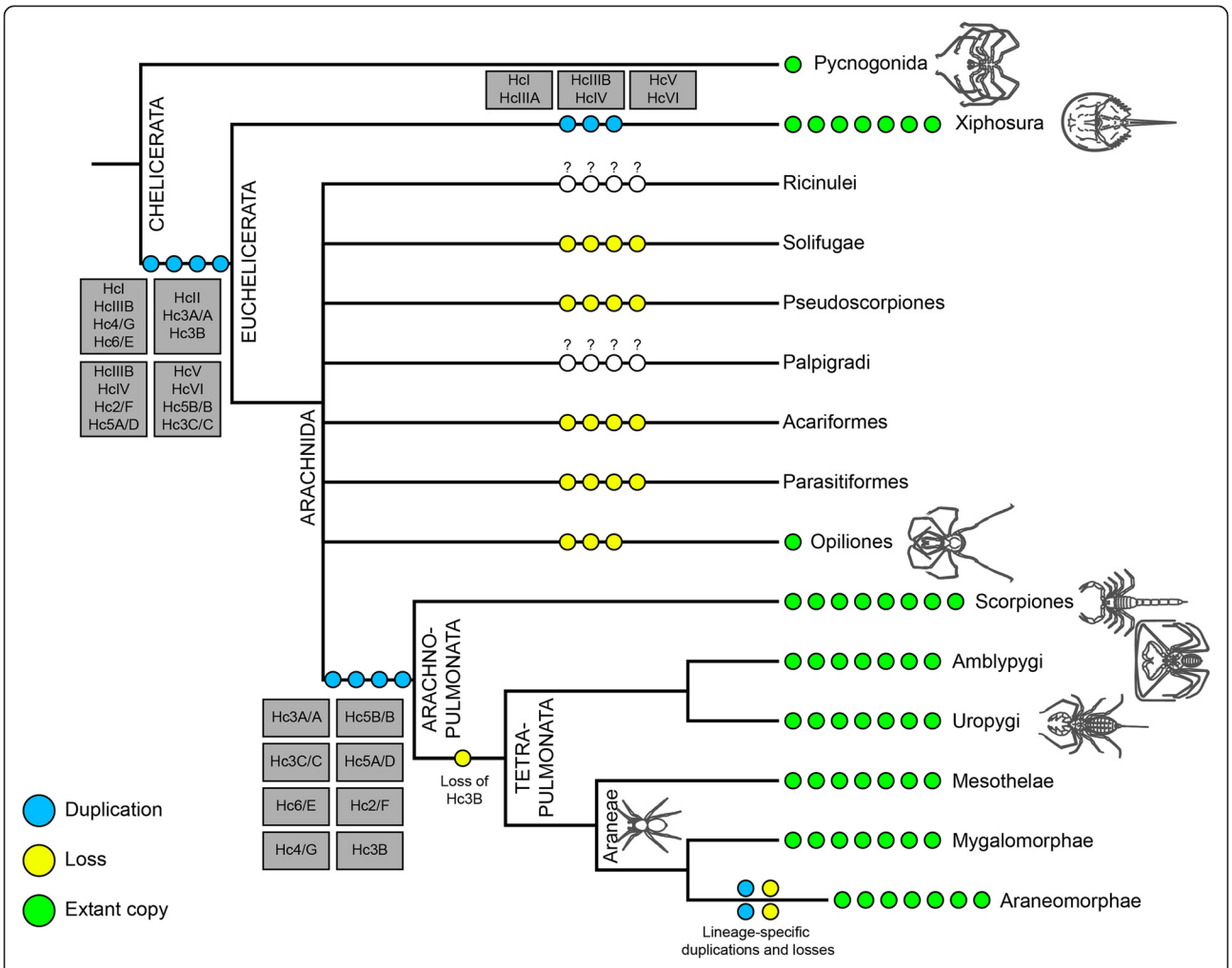

Figure 5 Inferred evolution of chelicerate hemocyanin, assuming minimization of duplication events. Circles indicate gene duplication (blue), gene losse (yellow), observed extant copy (green), or unknown (white). Grey boxes indicate inferred ancestral copies; enclosed text indicates descendent hemocyanin paralogs. Line drawings indicate orders wherein hemocyanin has been discovered. Basal relationships are shown for spiders (Araneae).

crab *Limulus polyphemus* were checked against novel sequences and augmented if novel sequences had greater length. Assembled sequences of hemocyanins are provided as aligned conceptual translations in Additional file 6.

Maximum likelihood analysis of tree topology

Maximum likelihood (ML) inference was conducted on static alignments, which were inferred by removing all indels from the Rehm et al. [31] submatrix, adding translated peptide sequences for new terminals' hemocyanins, and realigning the dataset with MUSCLE v.3.6 [66] with default parameters. The ML tree topology was inferred using RAxML v.7.3.0 [67] on 12 2.4-GHz Intel Xeon CPUs, with 500 independent starts. A WAG [68] model of sequence evolution with corrections for a discrete gamma distribution with four rate categories [69] was specified, following model selection with ProtTest 3 [70]. Nodal support was estimated with the rapid bootstrap algorithm of Stamatakis et al. [71] with 500 replicates.

Estimation of divergence times

Divergence time estimation was conducted using BEAST v.1.7.4 [9]. A WAG model with corrections for a discrete gamma distribution was used in all analyses. Fossil taxa were used to calibrate divergence times as follows. We used a Middle Devonian age (ca. 385–392 Ma) to calibrate the origin of Amblypygi (i.e., the split from Uropygi), based on limb and cuticle fragments that include a patella with trichobothria, a character that occurs uniquely on legs 2–4 of modern amblypygids [34]; we employed a normal prior with a mean of 385 Ma and a standard deviation of 10 Myr. The origin of Xiphosura was calibrated using a normal prior with a mean of 445 Ma and a standard

deviation of 10 Mya, based on the clear morphology of the Ordovician xiphosuran *Lunataspis aurora* [72]. The origin of spiders was calibrated with a normal prior with a mean of 386 Ma and a standard deviation of 10 Myr, based on recent reassessment of spigot morphology in the Middle Devonian fossil *Attercopus fimbriunguis* [43,44]. Finally, the root of the tree was calibrated using a normal prior with a mean of 501 Ma and a standard deviation of 10 Mya, based on the pycnogonid larval fossil *Cambropycnogon klausmuelleri* [73]. We used normal distributions as priors for calibrated nodes because these are more tractable for cross-calibration and cross-bracing analyses [17]; the use of large standard deviations enabled calibrated nodes to overcome underestimates imposed by fossil ages (e.g., the root of the tree).

Cross-calibration (CC) and cross-bracing (CB) analyses followed the implementation of Shih and Matzke [17]. We reused the same prior distribution for all nodes corresponding to the calibrations for CC analysis. For CB analyses, we added to the XML file an additional normally distributed prior whereby the difference in the calibrated node ages had a mean of zero and a standard deviation equal to 1% of the mean age of the calibration, for all calibrated nodes. As indicated by Shih and Matzke [17], this standard deviation was used to confer ease of sampling tree space, as tighter linking of node ages will limit MCMC sampling efficiency and increase computation time required to reach stationarity. Thirteen nodes (seven corresponding to origin of Amblypygi and six corresponding to origin of spiders) were cross-braced.

Extended cross-bracing (XCB) augmented the CB analysis with normally distribution priors linking the mean ages of the following nodes: diversification of Araneae (six nodes, due to the missing HcE paralog of *Liphistius malayanus*), diversification of Amblypygi (seven nodes), diversification of Xiphosura (seven nodes), diversification of Pedipalpi (=Amblypygi + Uropygi) (seven nodes), diversification of opishthothele spiders (seven nodes), diversification of Tetrapulmonata (one node for HcE; other paralogs already calibrated with spider origin), divergence of the two mygalomorph spiders (seven nodes), diversification of four araneomorph spiders (three nodes), and diversification of scorpions (eight nodes). Thus, a total of 53 additional nodes were braced in XCB analyses. For these additional uncalibrated nodes, the standard deviation of the linking normally distributed prior was set to 3, a large value anticipated to enable more efficient sampling of tree space and agnostic of the nodes' median ages.

All three analyses consisted of four runs, each with 5×10^7 generations. Stationarity was assessed using Tracer v.1.5 [74], and ESS values for posterior likelihood were observed to exceed 500 in all runs. 1×10^7 generations were discarded as burnin.

Additional files

Additional file 1: File S1. Chronogram of CC analysis, in nexus format.

Additional file 2: File S2. Chronogram of CB analysis, in nexus format.

Additional file 3: File S3. Chronogram of XCB analysis, in nexus format.

Additional file 4: Table S1. Median ages and bounds of 95% HPD intervals calculated for calibrated and uncalibrated nodes in CC, CB, and XCB analyses.

Additional file 5: Table S2. SRA accession numbers for Illumina transcriptomes used to reconstruct hemocyanin sequences.

Additional file 6: Multiple sequence alignment of chelicerate hemocyanins in fasta format.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

PPS conceived of the study, analyzed data, and wrote the paper. WCW edited the manuscript and supervised the work. Both authors read and approved the final manuscript.

Acknowledgments

We are indebted to Nicholas J. Matzke for assistance with XML formatting for BEAST. Alistair McGregor and two anonymous referees vetted the manuscript. Specimen collection and transcriptome construction for some species were facilitated by Rosa Fernández, Stefan T. Kaluziak, Alicia R. Pérez-Porro, Gustavo Hormiga, and Gonzalo Giribet. This material is based on work supported by the National Science Foundation Postdoctoral Research Fellowship in Biology under Grant No. DBI-1202751 to PPS.

Received: 16 June 2014 Accepted: 31 July 2014
Published: 8 August 2014

References

1. Zuckerkandl E, Pauling L: **Evolutionary divergence and convergence in proteins.** In *Evolving genes and proteins.* Edited by Bryson V, Vogel HJ. New York: Academic Press; 1965:97–166.
2. Sanderson MJ: **A nonparametric approach to estimating divergence times in the absence of rate constancy.** *Mol Biol Evol* 1997, **14:**1218–1231.
3. Renner SS: **Relaxed molecular clocks for dating historical plant dispersal events.** *Trends Plant Sci* 2005, **10:**550–558.
4. Rutschmann F: **Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times.** *Divers Distrib* 2006, **12:**35–48.
5. Rokas A, Krüger D, Carroll SB: **Animal evolution and the molecular signature of radiations compressed in time.** *Science* 2005, **310:**1933–1938.
6. Rota-Stabelli O, Daley AC, Pisani D: **Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution.** *Curr Biol* 2013, **23:**392–398.
7. Crisp MD, Trewick SA, Cook LG: **Hypothesis testing in biogeography.** *Trends Ecol Evol* 2011, **26:**66–72.
8. Cruaud A, Rønsted N, Chantarasuwan B, Chou LS, Clement WL, Couloux A, Cousins B, Genson G, Harrison RD, Hanson PE, Hossaert-Mckey M, Jabbour-Zahab R, Jousselin E, Kerdelhué C, Kjellberg F, Lopez-Vaamonde C, Peebles J, Peng Y-Q, Pereira RAS, Schramm T, Ubaidillah R, van Noort S, Weiblen GD, Yang D-R, Yodpinyanee A, Libeskind-Hadas R, Cook JM, Rasplus J-Y, Savolainen V: **An extreme case of plant–insect codiversification: figs and fig-pollinating wasps.** *Syst Biol* 2012, **61:**1029–1047.
9. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A: **Relaxed phylogenetics and dating with confidence.** *PLoS Biol* 2006, **4:**e88.
10. Pyron RA: **Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia.** *Syst Biol* 2011, **60:**466–481.
11. Ronquist F, Klopfstein S, Vilhelmsen S, Schulmeister S, Murray DL, Rasnitsyn AP: **A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera.** *Syst Biol* 2012, **61:**973–999.
12. Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S: **Estimating divergence times in large molecular phylogenies.** *Proc Natl Acad Sci U S A* 2012, **109:**19333–19338.

13. Stadler T, Yang Z: **Dating phylogenies with sequentially sampled tips.** *Syst Biol* 2013, **62:**674–688.
14. Rutschmann F, Eriksson T, Abu Salim K, Conti E: **Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points.** *Syst Biol* 2007, **56:**591–608.
15. Marshall CR: **A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points.** *Am Nat* 2008, **171:**726–742.
16. Ho SYW, Phillips MJ: **Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times.** *Syst Biol* 2009, **58:**367–380.
17. Shih PM, Matzke NJ: **Primary endosymbiosis events date to the later proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins.** *Proc Natl Acad Sci U S A* 2013, **110:**12355–12360.
18. Mellema JE, Klug A: **Quaternary structure of gastropod haemocyanin.** *Nature* 1972, **239:**146–150.
19. Markl J, Decker H: **Molecular structure of the arthropod hemocyanins.** *Adv Comp Env Physiol* 1992, **13:**325–376.
20. Burmester T: **Molecular evolution of the arthropod hemocyanin superfamily.** *Mol Biol Evol* 2001, **18:**184–195.
21. Burmester T: **Origin and evolution of arthropod hemocyanins and related proteins.** *J Comp Physiol B* 2002, **172:**95–107.
22. Decker H, Hellmann N, Jaenicke E, Lieb B, Meissner U, Markl J: **Minireview: recent progress in hemocyanin research.** *Integr Comp Biol* 2007, **47:**631–644.
23. Lieb B, Gebauer W, Gatsogiannis C, Depoix F, Hellmann N, Harasewych MG, Strong EE, Markl J: **Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550 kDa polypeptide.** *Front Zool* 2010, **7:**14.
24. Thonig A, Oellermann M, Lieb B, Mark FC: **A new haemocyanin in cuttlefish (*Sepia officinalis*) eggs: sequence analysis and relevance during ontogeny.** *EvoDevo* 2014, **5:**6.
25. Markl J: **Evolution of molluscan hemocyanin structures.** *Biochim Biophys Acta* 1834, **2013:**1840–1852.
26. Markl J: **Evolution and function of structurally diverse subunits in the respiratory protein hemocyanin from arthropods.** *Biol Bull* 1986, **171:**90–115.
27. Markl J, Stöcker W, Runzler R, Precht E: **Immunological correspondences between the hemocyanin subunits of 86 arthropods: evolution of a multigene protein family.** In *Invertebrate oxygen carriers.* Edited by Linzen B. Heidelberg: Springer Press; 1986:281–292.
28. Scholtz G, Kamenz C: **The book lungs of scorpiones and tetrapulmonata (Chelicerata, Arachnida): Evidence for homology and a single terrestrialisation event of a common arachnid ancestor.** *Zool* 2006, **109:**2–13.
29. Martin AG, Depoix F, Stohr M, Meissner U, Hagner-Holler S, Hammouti K, Burmester T, Heyd J, Wriggers W, Markl J: ***Limulus polyphemus* Hemocyanin: 10 Å Cryo-EM structure, sequence analysis, molecular modelling and rigid-body fitting reveal the interfaces between the eight hexamers.** *J Mol Biol* 2007, **366:**1332–1350.
30. Ballweber P, Markl J, Burmester T: **Complete hemocyanin subunit sequences of the hunting spider *Cupiennius salei*: recent hemoglobin remodeling in enelegyne spiders.** *J Biol Chem* 2002, **277:**14451–14457.
31. Rehm P, Pick C, Borner J, Markl J, Burmester T: **The diversity and evolution of chelicerate hemocyanins.** *BMC Evol Biol* 2012, **12:**19.
32. Jeram AJ: **Phylogeny, classifications and evolution of Silurian and Devonian scorpions.** In *Proceedings of the 17th European Colloquium of Arachnology.* Edited by Selden PA. Edinburgh (UK): British Arachnological Society, Burnham Beeches; 1998:17–31.
33. Dunlop JA, Kamenz C, Scholtz G: **Reinterpreting the morphology of the Jurassic scorpion *Liassoscorpionides*.** *Arthropod Struct Dev* 2007, **36:**245–252.
34. Dunlop JA: **Geological history and phylogeny of Chelicerata.** *Arthropod Struct Dev* 2010, **39:**124–142.
35. Coddington JA, Giribet G, Harvey MS, Prendini L, Walter DE: **Arachnida.** In *Assembling the Tree of Life.* Edited by Cracraft J, Donoghue MJ. New York (NY): Oxford: University Press; 2004:296–318.
36. Obst M, Faurby S, Bussarawit S, Funch P: **Molecular phylogeny of extant horseshoe crabs (Xiphosura, Limulidae) indicates Paleogene diversification of Asian species.** *Mol Phylogenet Evol* 2012, **62:**21–26.
37. Starrett J, Hedin M, Ayoub N, Hayshi CY: **Hemocyanin gene family evolution in spiders (Araneae), with implications for phylogenetic relationships and divergence times in the infraorder Mygalomorphae.** *Gene* 2013, **524:**175–186.
38. Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW: **Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences.** *Nature* 2010, **463:**1079–1083.
39. Cook CE, Smith ML, Telford MJ, Bastianello A, Akam M: ***Hox* genes and the phylogeny of the arthropods.** *Curr Biol* 2001, **11:**759–763.
40. Khadjeh S, Turetzek N, Pechmann M, Schwager EE, Wimmer EA, Damen WGM, Prpic N-M: **Divergent role of the Hox gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression.** *Proc Natl Acad Sci U S A* 2012, **109:**4921–4926.
41. Sharma PP, Schwager EE, Extavour CG, Giribet G: **Hox gene expression in the harvestman *Phalangium opilio* reveals divergent patterning of the chelicerate opisthosoma.** *Evol Dev* 2012, **14:**450–463.
42. Ayoub NA, Garb JE, Hedin M, Hayashi CY: **Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1γ), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae).** *Mol Phylogenet Evol* 2007, **42:**394–409.
43. Selden PA, Shear WA, Bonamo PM: **A spider and other arachnids from the Devonian of New York, and reinterpretations of Devonian Araneae.** *Palaeontology* 1991, **34:**241–281.
44. Selden PA, Shear WA, Sutton MD: **Fossil evidence of the origin of spider spinnerets and a proposed arachnid order.** *Proc Natl Acad Sci U S A* 2008, **105:**20781–20785.
45. Selden PA, Gall J-C: **A Triassic mygalomorph spider from the northern Vosges, France.** *Palaeontology* 1992, **35:**211–235.
46. Selden PA, Anderson JM, Anderson HM, Fraser NC: **Fossil araneomorph spiders from the Triassic of South Africa and Virginia.** *J Arachnol* 1999, **27:**401–414.
47. Selden PA: **First fossil mesothele spider, from the Carboniferous of France.** *Rev Suisse Zool* 1996, **2:**585–596.
48. Béthoux O: **The earliest beetle identified.** *J Paleontol* 2009, **83:**931–937.
49. Béthoux O, Klass KD, Schneider JW: **Tackling the Protoblattoidea problem: Revision of *Protoblattinopsis stubblefieldi* (Dictyoptera; Late Carboniferous).** *Eur J Entomol* 2009, **106:**145–152.
50. Béthoux O, Cui Y-Y, Kondratieff B, Stark B, Ren D: **At last, a Pennsylvanian stem-stonefly (Plecoptera) discovered.** *BMC Evol Biol* 2011, **11:**248.
51. Kenrick P, Wellman CH, Schneider H, Edgecombe GD: **A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic.** *Philos T Roy Soc B* 2012, **367:**519–536.
52. Legg DA, Garwood RJ, Dunlop JA, Sutton MD: **A taxonomic revision of Orthosternous scorpions from the English Coal-Measures aided by X-ray micro-tomography.** *Palaeontol Electron* 2012, **15:**1–16.
53. Nel P, Azar D, Prokop J, Roques P, Hodebert G, Nel A: **From Carboniferous to Recent: wing venation enlightens evolution of thysanopteran lineage.** *J Syst Palaeontol* 2012, **10:**385–399.
54. Garwood RJ, Sharma PP, Dunlop JA, Giribet G: **A Paleozoic Stem Group to Mite Harvestmen Revealed through Integration of Phylogenetics and Development.** *Curr Biol* 2014, **24:**1017–1023.
55. Bond JE, Garrison NL, Hamilton CA, Godwin RL, Hedin M, Agnarsson I: **Phylogenomics resolves a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution.** *Curr Biol* 2014. doi:10.1016/j.cub.2014.06.034.
56. Sharma PP, Giribet G: **A revised dated phylogeny of the arachnid order Opiliones.** *Front Genet* 2014, **5:**255.
57. Menon F: **Higher systematics of scorpions from the Crato Formation, Lower Cretaceous of Brazil.** *Palaeontology* 2007, **50:**185–195.
58. Lourenço WR, Gall J-C: **Fossil scorpions from the Buntsandstein (Early Triassic) of France.** *Syst Palaeontol* 2004, **3:**369–378.
59. Prendini L: **Phylogeny and classification of the superfamily Scorpionoidea Latreille 1802 (Chelicerata, Scorpiones): an exemplar approach.** *Cladistics* 2000, **16:**1–78.
60. Markl J, Markl A, Schartau W, Linzen B: **Subunit heterogeneity in arthropod hemocyanins: I. Chelicerata.** *J Comp Physiol B* 1979, **130:**283–292.
61. Shultz JW: **A phylogenetic analysis of the arachnid orders based on morphological characters.** *Zool J Linn Soc* 2007, **150:**221–265.
62. Burmester T: **Evolution and adaptation of hemocyanin within spiders.** In *Spider Ecophysiology.* Edited by Nentwig W. Heidelberg: Springer Press; 2013:3–14.
63. Riesgo A, Andrade SCS, Sharma PP, Novo M, Pérez-Porro AR, Vahtera V, González VL, Kawauchi GY, Giribet G: **Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa.** *Front Zool* 2012, **9:**33.
64. Hedin M, Starett J, Akhter S, Schönhofer AL, Shultz JW: **Phylogenomic resolution of Paleozoic divergences in harvestmen (Arachnida, Opiliones)**

via analysis of next-generation transcriptome data. *PLoS One* 2012, **7:**e42888.
65. Giribet G, Vogt L, Pérez González A, Sharma P, Kury AB: **A multilocus approach to harvestmen (Arachnida: Opiliones) phylogeny with emphasis on biogeography and the systematics of Laniatores.** *Cladistics* 2010, **26:**408–437.
66. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32:**1792–1797.
67. Stamatakis A: **RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics* 2006, **22:**2688–2690.
68. Whelan S, Goldman N: **A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach.** *Mol Biol Evol* 2001, **18:**691–699.
69. Yang Z: **Among-site rate variation and its impact on phylogenetic analyses.** *Trends Ecol Evol* 1996, **11:**367–372.
70. Darriba D, Taboada GL, Doallo R, Posada D: **ProtTest 3: fast selection of best-fit models of protein evolution.** *Bioinformatics* 2011, **27:**1164–1165.
71. Stamatakis A, Hoover P, Rougemont J: **A rapid bootstrap algorithm for the RAxML Web servers.** *Syst Biol* 2008, **57:**758–771.
72. Rudkin DM, Young GA, Nowlan GS: **The oldest horseshoe crab: a new xiphosurid from the Late Ordovician Konservat-Lagerstätten deposits, Manitoba, Canada.** *Palaeontology* 2008, **51:**1–9.
73. Waloszek D, Dunlop JA: **A larval sea spider (Arthropoda: Pycnogonida) from the Upper Cambrian 'Orsten' of Sweden and the phylogenetic position of pycnogonids.** *Palaeontology* 2002, **45:**421–446.
74. Rambaut A, Drummond AJ: **Tracer v. 1.5.** 2009. program and documentation available from: <http://tree.bio.ed.ac.uk/software/tracer/>.

doi:10.1186/s12983-014-0057-x
Cite this article as: Sharma and Wheeler: **Cross-bracing uncalibrated nodes in molecular dating improves congruence of fossil and molecular age estimates.** *Frontiers in Zoology* 2014 **11**:57.