Molecular Phylogenetics and Evolution 95 (2016) 171–182

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenomic analyses resolve an ancient trichotomy at the base of Ischyropsalidoidea (Arachnida, Opiliones) despite high levels of gene tree conflict and unequal minority resolution frequencies ☆

Casey H. Richart [a,b,*], Cheryl Y. Hayashi [b], Marshal Hedin [a]

[a] Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA
[b] Department of Biology, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Article history:
Received 13 March 2015
Revised 16 September 2015
Accepted 13 November 2015
Available online 9 December 2015

Keywords:
Multispecies coalescence
Concatenation
Concordance factors
Unequal minority resolution frequencies
Incomplete lineage sorting
Gene trees

ABSTRACT

Phylogenetic resolution of ancient rapid radiations has remained problematic despite major advances in statistical approaches and DNA sequencing technologies. Here we report on a combined phylogenetic approach utilizing transcriptome data in conjunction with Sanger sequence data to investigate a tandem of ancient divergences in the harvestmen superfamily Ischyropsalidoidea (Arachnida, Opiliones, Dyspnoi). We rely on Sanger sequences to resolve nodes within and between closely related genera, and use RNA-seq data from a subset of taxa to resolve a short and ancient internal branch. We use several analytical approaches to explore this succession of ancient diversification events, including concatenated and coalescent-based analyses and maximum likelihood gene trees for each locus. We evaluate the robustness of phylogenetic inferences using a randomized locus sub-sampling approach, and find congruence across these methods despite considerable incongruence across gene trees. Incongruent gene trees are not recovered in frequencies expected from a simple multispecies coalescent model, and we reject incomplete lineage sorting as the sole contributor to gene tree conflict. Using these approaches we attain robust support for higher-level phylogenetic relationships within Ischyropsalidoidea.

1. Introduction

Harvestmen (Opiliones) are among the most species-rich arachnid orders (Harvey, 2002), and have an ancient diversification history (Hedin et al., 2012; Sharma and Giribet, 2014). Within Opiliones there are 46 recognized families, approximately 1500 genera, and more than 6500 described species (Machado et al., 2007; Kury et al., 2014). Phylogenomics of higher-level relationships within Opiliones (Hedin et al., 2012) has found strong support for four primary clades (suborders Cyphophthalmi, Laniatores, Dyspnoi, and Eupnoi), with recent amendments based on novel fossil discoveries (Giribet and Sharma, 2015). Here we build upon the research of Hedin et al. (2012) to analyze relationships within the Dyspnoi superfamily Ischyropsalidoidea. The Ischyropsalidoidea are confined to the northern hemisphere, with the age of the root estimated to be as recent as 137 MYA (Schönhofer et al., 2013) or as old as 240–360 MYA (Sharma and Giribet, 2014). Currently, 85 species are classified into seven genera: *Ischyropsalis*, *Sabacon*, *Taracus*, *Ceratolasma*, *Acuclavella*, *Hesperonemastoma*, and *Crosbycus* (Kury, 2013). The superfamily is defined on the basis of genitalic characters (Martens, 1976), palpal morphology (Martens et al., 1981), and by having metapeltidial sensory cones (Shear, 1986; though see Shultz, 1998). Each genus is morphologically distinct (Fig. 1), and there has been little controversy regarding their respective monophyly. In fact, monogeneric families have been proposed for a number of genera (e.g., Dresco, 1970; Martens, 1976; Shear, 1986; Schönhofer, 2013), but this taxonomic solution was criticized by Gruber (1978). Conversely, family level hypotheses within Ischyropsalidoidea have been problematic, with two of the last three non-monogeneric familial hypotheses (Sabaconidae of Giribet et al., 2010, Taracidae of Schönhofer, 2013) failing to identify diagnostic morphological synapomorphies.

Abbreviations: AA, amino acid; AT3, AT content at the 3rd base-pair position; BCA, Bayesian concordance analysis; BIC, Bayesian information criterion; BLAST, Basic Local Alignment Search Tool; BSV, bootstrap value; CF, concordance factor; ESS, effective sample size; ILS, incomplete lineage sorting; LBA, long-branch attraction; ML, maximum likelihood; MYA, millions of years ago; NGS, next generation sequencing; NNI, nearest neighbor interchange; OTU, operational taxonomic unit; PP, posterior probability; SPR, subtree-pruning-regrafting; UMRFs, unequal minority resolution frequencies.

☆ This paper was edited by the Associate Editor Francesco Frati.

* Corresponding author at: Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.
E-mail address: pileated@gmail.com (C.H. Richart).

http://dx.doi.org/10.1016/j.ympev.2015.11.010
1055-7903/

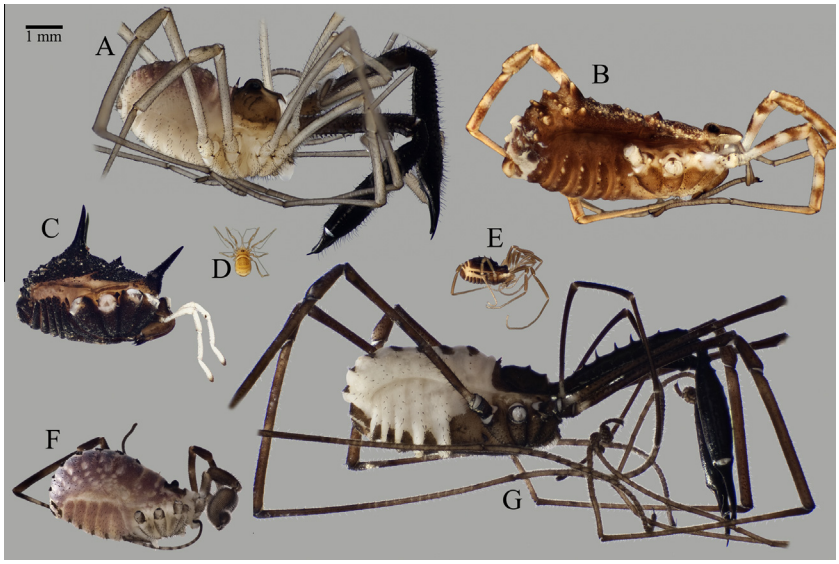

Fig. 1. Ischyropsalidoidea. Generic representatives from the superfamily Ischyropsalidoidea. A. *Taracus gertschi* (851092), B. *Ceratolasma tricantha* (850889), C. *Acuclavella makah* (829726), D. *Crosbycus dasycnemus* (851086), E. *Hesperonemastoma* sp. (851085), F. *Sabacon* sp. (851091), and G. *Ischyropsalis h. hellwigi* (851090). Full sized high-resolution images can be seen at MorphBank using the specimen identification numbers listed behind each name. Specimens were imaged using a Visionary Digital BK Plus system (http://www.visionarydigital.com) with composite images combined using Zerene Stacker 1.04 (http://www.zerenesystems.com), and edited with Adobe Photoshop CS6.

Systematics has moved into an era where phylogenetic hypotheses are being resolved at an unprecedented rate. For example, well-studied systems, such as mammalian interordinal relationships, now contain few controversial nodes (e.g., Murphy et al., 2001; Meredith et al., 2011). This development is due in part to the arrival of next-generation sequencing (NGS) technologies and continued advance in statistical phylogenetics. With the ability to generate matrices containing hundreds of loci (e.g. Hedin et al., 2012; Faircloth et al., 2012), NGS data have both supported pre-existing hypotheses, and recovered novel taxonomic hypotheses that are robustly supported. Furthermore, NGS technologies have allowed for the identification of rapid, ancient radiations (e.g., McCormack et al., 2013; Teeling and Hedges, 2013). These radiations are notoriously hard to resolve (e.g. Faircloth et al., 2012; Springer and Gatesy, 2014), with short internal branches that are in part characterized by high levels of gene tree conflict. Gene trees can conflict with a species tree for numerous reasons (Maddison, 1997; Maddison and Knowles, 2006; Degnan and Rosenberg, 2009) including undetected paralogy, recombination, hybridization, saturation, and long-branch attraction (LBA). A major source of gene tree conflict inherent to rapid diversifications is coalescent stochasticity – the random sorting of ancestral polymorphisms across successive speciation events (Kingman, 1982; Degnan and Rosenberg, 2009). This phenomenon is known as incomplete lineage sorting (ILS), and its occurrence is expected to increase as a function of shorter internal branch lengths and larger ancestral population sizes (Maddison, 1997). ILS can occur to such an extent that the most likely gene tree is incongruent with the species tree, a situation that has been defined as an "anomaly zone" (Degnan and Rosenberg, 2006).

A problem inherent to molecular phylogenetic reconstruction of ancient and rapid successive diversification events is that short internal branches do not provide enough time for slowly evolving loci to accumulate informative substitutions, whereas more rapidly evolving loci accumulate homoplastic substitutions along descending long branches (Regier et al., 2008). Such internal branches can have so little phylogenetic signal that even small amounts of non-phylogenetic signal can yield support for an incorrect phylogeny (Huelsenbeck and Hillis, 1993; Swofford et al., 2001; Philippe et al., 2011), and this can occur to such an extent that it is positively misleading (Huelsenbeck and Hillis, 1993; Bull et al., 1993). Therefore, although the selective use of slowly evolving coding regions (nucleotides or amino acids) has been a successful approach for reconstructing the backbone of numerous higher-level phylogenies (e.g., Iwabe et al., 1989; Hedin et al., 2012; Zhang et al., 2012; Lang et al., 2013; Raymann et al., 2014), these loci are not expected to contain enough informative characters to resolve the branching order of ancient internal nodes across short branches, and simply adding more data does not guarantee that analyses will resolve the correct topology (Swofford et al., 2001; Philippe et al., 2011).

Longer loci have been shown to improve phylogenetic inference in situations where long terminal branches relative to short internal branches cause "zones" of inconsistent estimation (Swofford et al., 2001). Accordingly, independent loci have traditionally been concatenated into a supermatrix, with the assumption that this will allow for the emergence of hidden support, or the increased support for a clade relative to the sum of support for the clade when data partitions are analyzed separately (Gatesy et al., 1999). The theoretical argument against the supermatrix approach is that recombination and coalescent stochasticity result in genes having different evolutionary histories, and that concatenating these loci into a supermatrix (which in effect treats all data as a single locus) violates the assumption of recombination (Kubatko and Degnan, 2007). Simulation studies under these conditions (short internal branches with high levels of gene tree conflict) have supported this contention, showing that concatenation can result in support for incorrect topologies (Seo, 2008), with support increasing as more loci are added (Kubatko and Degnan, 2007). To address these concerns, methods of phylogenetic inference have been developed under multispecies coalescent models that co-estimate gene trees, divergence times, population sizes, and a species tree from multiple unlinked loci (e.g., BEST, Liu and Pearl, 2007; *BEAST, Heled and Drummond, 2010). Simulation studies have shown that fully-parametric multispecies coalescent methods can be highly accurate even with high levels of gene tree incongruence (Liu and Edwards, 2009), and outperform supermatrix methods (Heled and Drummond, 2010). Currently, the major

shortcoming of fully-parametric coalescent analyses is that the very large parameter space is too computationally demanding to be applied to analyses with a large number of loci and/or taxa (Edwards et al., 2007; Liu et al., 2009; Bayzid and Warnow, 2013; O'Neill et al., 2013). As such, many phylogenomic analyses use partially-parametric coalescent analyses such as STAR (Liu et al., 2009) or MP-EST (Liu et al., 2010). These "short-cut" coalescent analyses use gene trees constructed in isolation as input for phylogenetic inference, and assume that gene trees are correctly inferred and all gene tree discordance is due to ILS. However, gene trees can conflict at ancient and rapid divergences for numerous reasons (see above). Further, careful reanalyses of short-cut coalescent results (e.g., the data from Song et al., 2012) have shown that such methods can provide high support for likely erroneous topologies when the assumption of correctly inferred gene trees is violated (Springer and Gatesy, 2016). Simulations have long shown that phylogenetic inference of deep divergences associated with short internal branches produces a large number of incorrect gene trees simply do to sampling error (e.g., Huelsenbeck and Hillis, 1993; Swofford et al., 2001); empirically this problem is more extensive with shorter loci (Gatesy and Springer, 2014). As such, there is continued debate as to which phylogenomic methods are preferable for resolution of ancient rapid diversifications (Liu et al., 2010; Leaché and Rannala, 2011; Song et al., 2012; Lemmon and Lemmon, 2013; Patel et al., 2013; Gatesy and Springer, 2014; Lanier and Knowles, 2015; Springer and Gatesy, 2016; Edwards et al., 2016).

In this paper we report on a combined phylogenetic approach utilizing transcriptome data in conjunction with Sanger sequences (e.g., Leaché et al., 2014b) to analyze two nearly independent phylogenetic matrices, including an "expanded panel" that contains 14 loci for 12 ingroup terminals (with some missing data), and a "transcriptome panel" that contains 672 loci for 3 ingroup terminals (no missing data). Preliminary phylogenetic analysis of the expanded panel identified a weakly supported topology deep in the ischyropsalidoid species tree. Therefore, we used transcriptome data from ischyropsalidoid exemplars descending from these ancient and rapid diversifications to specifically target these problematic nodes. With this combined strategy we reconstruct a robustly supported phylogeny for every node sampled within Ischyropsalidoidea to the rank of genus. We are able to identify a near trichotomy at the base of the superfamily that has resulted in high levels of gene tree incongruence, and show that the minority resolution frequencies of alternative topologies are unequal. For this ancient and short internal branch we compare multiple phylogenetic methods that are congruent in their support for a topology not previously recovered for ischyropsalidoids. Further, analyses of this dataset suggest that the supermatrix approach recovers the agreed upon phylogeny with fewer loci and higher support than do partially-parametric coalescent analyses.

2. Materials and methods

2.1. Primer design, PCR, and Sanger sequencing

Protein-coding genes annotated as single-copy single-exon in *Ixodes scapularis*, a well-annotated arachnid genome, were downloaded and filtered from VectorBase (http://iscapularis.vectorbase.org/). BLAST was used to query these loci against three published transcriptome assemblies (*Hesperonemastoma*, *Ortholasma*, *Trogulus*; Hedin et al., 2012) to generate alignments for PCR primer design. PCR primers were manually designed based on these alignments in Geneious Pro 5.5 (Kearse et al., 2012) and characterized using Primer3 (Rozen and Skaletsky, 2000). Primers were tested against *Hesperonemastoma*, *Ortholasma* and *Trogulus* (*HOT*) genomic DNA extractions, and primer combinations successful on any member of the *HOT* panel were then tested on an expanded panel of ischyropsalidoid genera (*Sabacon*, *Taracus*, *Acuclavella*, *Ceratolasma*, *Ischyropsalis*, and an additional *Hesperonemastoma*). In addition to newly designed loci, the expanded panel and outgroups (*Ortholasma* and *Trogulus*) were amplified for gene regions previously used at deeper levels in Opiliones. These included EF-1α (Hedin et al., 2010), 18S and 28S (Giribet et al., 1999; Shultz and Regier, 2001), COI (many authors, e.g., Richart and Hedin, 2013; Derkarabetian and Hedin, 2014), polII (Shultz and Regier, 2001), and wingless (Wnt2; Richart and Hedin, 2013). Detailed methods regarding locus selection, primer design, PCR conditions, and Sanger sequencing are available in the Supplement (s1.1).

2.2. Expanded panel phylogenetics

To evaluate ischyropsalidoid intergeneric relationships we targeted an expanded panel of six ischyropsalidoid genera (see above), including two species from each genus, plus outgroups. The intrageneric sampling scheme targeted species spanning the root node of each genus with the intention of subdividing long branches. This sampling was informed by previous research in *Acuclavella* (Richart and Hedin, 2013), *Sabacon* (Schönhofer et al., 2013), *Ischyropsalis* (Schönhofer et al., 2015), and *Hesperonemastoma* (unpublished: Richart, Hayashi, and Hedin). Exemplars of *Taracus* and *Ceratolasma* were chosen from distant localities within their respective geographic distributions. Original sequence data were augmented with GenBank sequences. Also, the 14 OTUs in the expanded matrix were occasionally represented by multiple intraspecific individuals or relatively closely related species (Appendix). Expanded panel specimens were field-collected and stored at −80 °C in 100% EtOH (Vink et al., 2005) with the exception of *Ischyropsalis* which was preserved in a urea buffer (Asahida et al., 1996). All extractions were conducted using the Qiagen DNeasy Blood & Tissue Kit, per manufacturer's protocol; most extractions were performed using half of a bilaterally divided individual, with the other half saved as a voucher.

Expanded panel alignments were generated from newly developed markers (eight loci) and six previously-used loci (see above). Some alignments were further populated using transcriptome-derived sequence data, and trimmed to the start at the nearest first base pair of an open reading frame. GenBank accession numbers, data matrix coverage, and alignment lengths are provided in Table 1. All alignments were conducted in Geneious using MAFFT 6 (Katoh et al., 2002), and regions of alignment uncertainty were removed with GBlocks 0.91b (Castresana, 2000). Partitions and models of evolution were jointly estimated using PartitionFinder 1.1.1 (Lanfear et al., 2012) for protein coding loci using linked branch lengths, BIC criterion, and a greedy search algorithm, with analyses run separately to inform *BEAST, RAxML, and MrBayes analyses. Substitution models for translated AA sequences for the eleven nuclear protein-coding loci were estimated using MEGA 6.06, using ML model selection (Tamura et al., 2013). Evolutionary models for the ribosomal regions 28S and 18S utilized jModelTest 2.1.6 (Guindon and Gascuel, 2003; Darriba et al., 2012), considering 24 models evaluated using AIC criterion to choose optimal models under a ML search. Further methods for model selection and resulting models are available as Supplementary material (Table S2).

ML gene trees and concatenated phylogenetic analyses were run using RAxML-HPC2 8.0.24 (Stamatakis et al., 2008) on the CIPRES Science Gateway 3.3 (Miller et al., 2010). A rapid bootstrap analysis and search for the best-scoring ML tree (-f a) was conducted using the GTRGAMMA model. The RAxML concatenated phylogeny was repeated three times. Bayesian phylogenetic reconstruction used both concatenation (via MrBayes 3.2.1; Ronquist

Table 1
GenBank numbers, expanded panel matrix, and alignment lengths. Italicized GenBank accession numbers represent sequences downloaded from GenBank. See Appendix for additional voucher information. See Table S1 for gene annotations.

OTU	8	17	49	69	156	281	300	334	COI	EF1a	polII	r18S	r28S	Wnt2	*n* =
Ortholasma	KU168429	KU168438	KU168456	KU168457	KU168473	KU168483	KU168491	KU168498	*GQ912870*	KU168506	KU168516	KU168520	KU168533		13
Trogulus	KU168430	KU168439	KU168446	KU168458	KU168474	KU168476	KU168492	KU168499	*GQ912872*	*AF240880*	KU168517	KU168521	KU168534	KU168542	14
Hesperonemastoma A	KU168431	KU168440	KU168447	KU168459	KU168466	KU168477	KU168493	KU168500	*EF108588*	*AF240869*	KU168518	KU168522	KU168535	KU168543	14
Hesperonemastoma B	KU168432	KU168441	KU168448		KU168467	KU168478		KU168495	KU168503	KU168507		KU168523	KU168536	KU168544	11
Taracus A			KU168449	KU168460	KU168468	KU168479	KU168485	KU168496		KU168508		KU168524	*JX573592*	KU168545	10
Taracus B			KU168450		KU168469	KU168480	KU168486	KU168497	*GQ912867*	KU168509	*AH010475*	KU168525		KU168546	10
Sabacon A	KU168437	KU168442	KU168451	KU168464	KU168475	KU168481	KU168494	KU168501	KU168505	KU168510	*AH010471*	KU168526	KU168537	KU168547	14
Sabacon B		KU168443	KU168452	KU168461		KU168482			*JX573670*	KU168511		KU168527	KU168538	KU168548	9
Acuclavella A	KU168433		KU168453	KU168462	KU168470		KU168487		*GQ870647*	KU168512		KU168528	KU168539	KU168549	10
Acuclavella B	KU168434	KU168445	KU168454	KU168465	KU168471	KU168484	KU168488	KU168502	*GQ870645*	KU168513	KU168519	KU168529	KU168540	KU168550	14
Ceratolasma A	KU168435		KU168455				KU168489		*GQ912865*	KU168514	*AH010458*	KU168530	KU168541		8
Ceratolasma B	KU168436			KU168463	KU168472		KU168490			KU168515		KU168531		KU168551	7
Ischyropsalis A		KU168444							KU168504	*JX573604*	AH010464	KU168532	*JX573546*	KU168552	7
Ischyropsalis B									*JX573639*	*JX573603*			*JX573545*		3
n=	9	8	11	9	10	9	10	8	12	14	8	13	12	11	
Align. Length (BPs)	729	456	588	324	438	612	789	517	1098	672	1137	547	1110	390	Total Length (BP): **9407**

et al., 2012) and the coalescent-based *BEAST (Heled and Drummond, 2010). MrBayes 3.2.1 was run for 10 million generations at which point the average standard deviation of split frequencies was <0.0001. Parameters were logged every thousand generations, and the first 25% of samples were discarded as burn-in. Lognormal relaxed clock models were specified for all 14 loci in preliminary *BEAST analyses in order to determine whether a strict clock could be statistically rejected for any of the loci. Estimates of the coefficient of variation for each locus indicated that a strict clock model could not be rejected for 9 of the 14 loci; a strict clock model was specified for each of these 9 loci in subsequent analyses. Six independent *BEAST runs were conducted for 100,000,000 generations, sampling every 10,000 generations. ESS values and stationarity of likelihood values were evaluated in Tracer 1.6, then combined using LogCombiner 1.8.0 with a 60% burn-in. This output was used to reconstruct a maximum clade credibility tree in TreeAnnotator 1.8.0. The concatenated and unpartitioned AA alignment was analyzed in PhyML 3.0 under the JTT+G model and using the best of NNI and SPR tree search algorithm (Guindon et al., 2010).

2.3. Transcriptome generation and phylogenomics

Preliminary expanded panel phylogenetic analyses showed a near trichotomy with long-branches separating three ingroup lineages including (1) *Sabacon*, (2) *Hesperonemastoma* and *Taracus*, and (3) *Ischyropsalis*, *Acuclavella*, and *Ceratolasma*. Thus we selected exemplars representing each of these lineages (*Sabacon*, *Hesperonemastoma*, and *Acuclavella*) for transcriptome-based phylogenomics. The transcriptome panel also included two outgroup taxa (*Trogulus* and *Ortholasma*). Three of the panel transcriptomes were previously published (Hedin et al., 2012; SRX450964, SRX451776, and SRX450937 for *Trogulus*, *Ortholasma*, and *Hesperonemastoma* respectively). RNA extractions for *Sabacon* and *Acuclavella* were conducted using TRIzol, and purified with the Qiagen RNeasy MinElute Cleanup Kit. Extractions used whole body tissues from single individuals with the midgut removed to reduce contamination. RNAs were sent to the Genomic Services Lab at HudsonAlpha Institute for Biotechnology (www.hudsonalpha.org) where non-normalized libraries were prepared using the Illumina TruSeq RNASeq kit and then sequenced using Illumina HiSeq technology with paired-end, 50-bp reads. For all transcriptomes adaptors were trimmed using TrimGalore! 0.2.7 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), and sequences with >1% ambiguity or less than thirty base pairs were removed using PRINSEQ Lite 0.20.2 (Schmieder and Edwards, 2011). Sequence reads passing these filters were assembled *de novo* using the Trinity platform (Grabherr et al., 2011; Haas et al., 2013).

Sets of orthologous sequences were filtered from the five transcriptomes by first identifying putative homologs to 5470 *Ixodes* loci annotated as single-exon, and also the 367 harvestmen loci used previously by Hedin et al. (2012). Identification of putative homologs was conducted with a liberal BLAST e-value (1e−1; Altschul et al., 1990). Queries resulting in missing data ($n = 3931$) or with multiple highly-overlapping hits from a single transcriptome ($n = 985$) were not considered further. Paralogy was further assessed in two ways. First, a representative sequence from each alignment was again subjected to BLAST against the Dyspnoi transcripts, with a conservative e-value (1e−50). If this BLAST returned multiple sequences per exemplar, then alignments were discarded ($n = 2$). Second, gene trees not recovering ($n = 20$) or not supporting (with a BSV < 70; $n = 30$) a monophyletic Ischyropsalidoidea were discarded under the assumption that incongruent nodes for this otherwise well-supported clade are the result of paralogy or LBA. These criteria were not mutually exclusive, and a total of 672 loci passed query and paralogy filters.

Gene trees for 672 loci were estimated using ML in PhyML 3.0 (Guindon et al., 2010) using default parameters including the HKY85+G substitution model and the NNI tree search algorithm. Nodal support was assessed via 100 bootstrap replicates, which were rooted using the reroot tool on the STRAW web server (Shaw et al., 2013). STRAW was also used to conduct partially-parametric coalescent-based analyses, using PhyML gene trees as input. For these analyses we used both MP-EST (Liu et al., 2010), which uses the frequency of triplets of taxa to estimate the topology and branch lengths, and STAR (Liu et al., 2009), which computes the pairwise topological distance among pairs of taxa to determine the average placement of nodes across a collection of gene trees. The coalescent arises as a large-population approximation of the Wright-Fisher model (Nordborg, 2001), thus coalescent analyses make the same simplifying assumptions including constant populations sizes and no selection, and attribute all gene tree incongruence to ILS (Kubatko and Degnan, 2007; Springer and Gatesy, 2016). These short-cut coalescent methods were chosen because the size of the transcriptome panel was too computationally demanding to implement fully-parametric coalescent analyses (e.g., *BEAST). Additionally, we analyzed the transcriptome panel via concatenation using RAxML-HPC2 on CIPRES. This supermatrix was partitioned by gene using the default "new rapid hill-climbing" tree search algorithm, with a GTRGAMMA model applied to each partition.

2.4. Comparison of concatenation versus coalescent phylogenomic analyses

We evaluated the performance and consistent estimation of concatenation versus partially-parametric methods by randomly sub-sampling transcriptome-derived loci. Ten replicates each of 25, 50, 100, 200, 300, 400, 500, and 600 loci were selected, resulting in a total of 80 replicates. Phylogenetic analyses for each of these replicates was performed using MP-EST, STAR, and RAxML, using parameters as outlined above.

2.5. Evaluation of unequal minority resolution frequencies

Under the basic multispecies coalescent model the frequency of minority resolution gene trees should be equal (Pamilo and Nei, 1988). We used the 672 PhyML gene trees to test the equality of minority resolution frequencies using a two-sided binomial test. In order to evaluate if UMRFs are caused by methodological bias, we summarized the symmetry of gene tree frequency across different attributes of our data. Additionally, we analyzed a subset of our loci that were retrieved from transcriptomes using a different methodological pipeline (loci from Hedin et al., 2012). If minority asymmetry persists across different attributes or treatments of the data, we assume that UMRFs are not a methodological artifact, but are caused by biological aspects such as structured ancestral populations or paraphyletic gene flow (see Discussion). We choose two locus attributes that have been suggested to improve phylogenomic pipelines, including high AT3 (Romiguier et al., 2013), and high phylogenetic support values (Salichos and Rokas, 2013, though see Betancur-R et al., 2014). We further evaluated the frequency of alternative topologies in our data by analyzing CFs, or the proportion of the sampled genome that agree with a given bipartition, within the 672-locus dataset (Baum, 2007). This was done using a Bayesian Concordance Analysis (BCA; Ané et al., 2007) in the program BUCKy (Larget et al., 2010). BUCKy uses independent Bayesian analysis of each gene as input. These analyses were conducted in MrBayes 3.2.1 (Ronquist et al., 2012), based on 100,000 generation runs, sampling every 100 trees, and discarding the first 250 trees as burn-in. Each locus included two partitions, one combining the 1st and 2nd bp position, and another for the 3rd, with nst = 6 and rates = gamma. BUCKy was used to map the posterior sample of trees to alternative topologies using an *a priori* expectation of gene tree discordance. For this analysis, the prior level of discordance (α) was chosen to give equal likelihood to each of the three possible rooted triplets. The probability that two loci share the same tree is about $1/(1+\alpha)$, thus we set $\alpha = 2$. This analysis can be used to reject the hypothesis that all gene tree discordance is due to incomplete lineage sorting (Ané, 2010).

3. Results

3.1. Expanded panel phylogenetics

Primer design resulted in the development of eight molecular markers with phylogenetic utility in Ischyropsalidoidea (Table s1). Data augmented from *Sabacon* and *Acuclavella* transcriptomes improved the average percentage of loci sampled per OTU in the expanded panel (Table 1). The final expanded panel included 14 loci for 12 ingroup taxa, with a concatenated alignment length of 9407 bp (26.5% missing). All phylogenetic analyses were rooted with the troguloid genera *Ortholasma* and *Trogulus* except for the Wnt2 matrix that contained data for *Trogulus* only. Expanded panel RAxML gene trees are deposited in the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.3mr26) and available in the Supplementary material (s4). Ischyropsalidoidea and all genera are recovered with high support in the majority of gene trees. Twelve of the 14 loci recovered Ischyropsalidoidea with BSVs $\geqslant$ 97. Occasionally genera were not recovered as monophyletic, though paraphyly was always with respect to closely related genera. Furthermore, gene trees tended to recover (*Ischyropsalis*, (*Acuclavella*, *Ceratolasma*)) with strong support. It should be noted that we found no evidence supporting EF-1α paralogy in this study, and the rampant gene duplication of EF-1α known from a single species of Cyphophthalmi (Clouse et al., 2013) appears not to be problematic in Dyspnoi (see also Schönhofer et al., 2015, Supplement s2.1).

Despite recovery of superfamily and "tip" relationships, nearly all backbone nodes within Ischyropsalidoidea lack support in gene tree analyses – i.e., it is unclear how most genera are related by examining individual gene trees. In contrast, combined phylogenetic analyses of the expanded panel recovered the same topology across most methods of inference (Fig. 2). The one exception is the PhyML analysis of translated AAs which recovered *Sabacon* sister to remaining ischyropslidoids, though this relationship is not well supported (BSV = 52; Supplement s4). This analysis also fails to recover the genus *Ischyropsalis* as monophyletic, with these two taxa by far having the most missing data (Table 1). All other combined analysis nodes were strongly supported by MrBayes, RAxML and *BEAST, with the exception of a node associated with a short branch deep in the ischyropsalidoid phylogeny. This node was most strongly recovered in the MrBayes concatenated analysis, which had only one tree in the 99% credible set – the only other sampled tree recovered (*Sabacon*, (*Hesperonemastoma*, *Taracus*)). This node was less well-supported in RAxML (BSV = 68) and *BEAST (PP = 0.82) analyses.

3.2. Transcriptome panel phylogenomics

Transcriptome assembly statistics, and comparison to previously published (Hedin et al., 2012) transcriptome assemblies, are reported in Table 2. The final transcriptome panel included 672 loci, 3 ingroup and 2 outgroup taxa, with a concatenated alignment length of 536,124 bp. These data are nearly distinct from the expanded panel, with 5 loci with various levels of overlap totaling 1335 bp. Phylogenetic analyses of the transcriptome panel further resolved phylogenetic relationships at the base of Ischyropsalidoidea despite high levels of gene tree conflict. Evaluation of

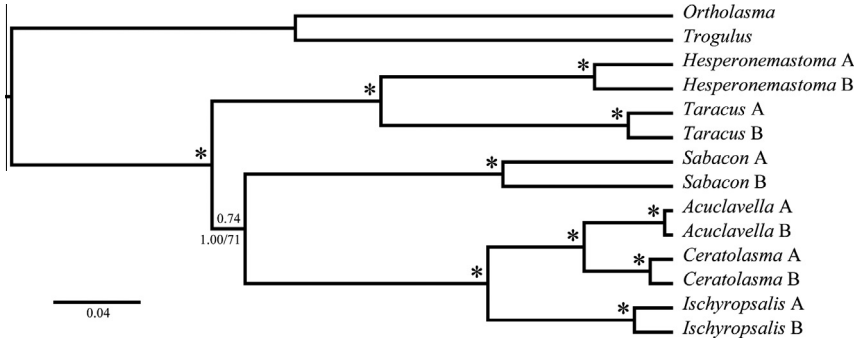

Fig. 2. Expanded panel phylogeny. Phylogeny of Ischyropsalidoidea based on 14 genes analyzed via coalescent (*BEAST; topology pictured) and concatenated (MrBayes, RAxML) methods. The node with support values shows *BEAST above, and MrBayes/RAxML support values below the parent branch. All other nodes were recovered with *BEAST posteriors ⩾0.96, MrBayes posteriors of 1.0, and RAxML bootstrap values ⩾99. Identification of samples used to populate OTUs follows Appendix and Table 1.

Table 2
Transcriptome data and assembly information.

Taxon	# of paired-end reads	# Gb	# Transcripts >200 bp	Mean length (>200 bp)	Max length
Ortholasma	80.7 M (50-bp)	4.04	34,357	839.6	11,074
Trogulus	54.7 M (50-bp)	2.74	46,840	937.4	9614
Hesperonemastoma	120.0 M (50-bp)	6.00	42,007	999.1	8952
Acuclavella	60.9 M (100-bp)	6.09	20,926	1494.9	36,044
Sabacon	43.3 M (100-bp)	4.33	24,135	1121.7	12,424

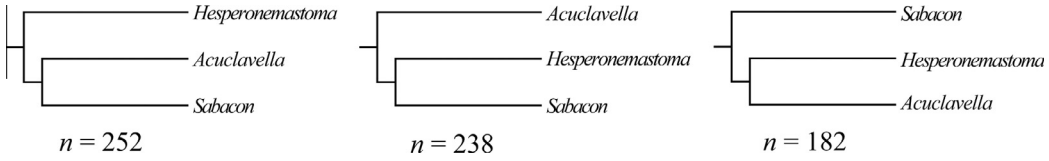

Fig. 3. Gene tree synopsis. Results of PhyML gene tree analyses of 672 loci.

PhyML gene trees indicate high levels of gene tree incongruence (Fig. 3), with (*Hesperonemastoma*, (*Acuclavella*, *Sabacon*)) recovered in 37.5%; (*A*, (*H*, *S*)) in 35.4%; and (*S*, (*H*, *A*)) in 27.1% of gene trees. Concatenated analysis of all data (RAxML), and coalescent-based analyses of PhyML gene trees (MP-EST and STAR) were congruent in their recovery of *Hesperonemastoma* as sister to the other remaining ischyropsalidoid lineages, but differed in their support and inferred branch lengths (Fig. 4). Partially-parametric coalescent analyses tended to recover a very short branch just inside of Ischyropsalidoidea with only moderate support values. The concatenated analysis recovered a longer internal ischyropsalidoid branch with higher support for this topology.

3.3. Comparison of phylogenomic analyses from sub-sampled loci

Concatenated and partially-parametric coalescent analyses tended to recover the same topology in any particular replicate, though concatenated analyses more consistently recovered (*H*, (*A*, *S*)) with higher support than in coalescent analyses, which did not settle on this topology until after 300 or more loci were analyzed. Perhaps most conspicuous is a 600-loci replicate that was recovered as (*A*, (*H*, *S*)) by both STAR and MP-EST. Examination of the results from sub-sampled loci shows that recovering *Hesperonemastoma* as sister to *Sabacon* + *Acuclavella* could not have reasonably been recovered without using over 400 loci (Fig. 5).

3.4. Unequal minority resolution frequencies

The minority resolution frequencies in the 672-locus dataset were unequal (two-sided binomial test, $p = 0.0072$). This trend persisted across treatments (Table 3), though this was not significant for the subset of loci that were generated by Hedin et al., 2012 ($p = 0.1203$). BCA analyses reject the hypothesis that all gene tree discordance is due to ILS with 99% confidence (Fig. 6). The BUCKy concordance tree (Fig. 4) recovers the same topology as concatenation and coalescent-based analyses, with a CF of 0.475. The 99% highest posterior density interval of trees in the posterior sample (0.394–0.475) does not overlap with either of the minority resolution topologies. Both of the alternative topologies (*A*, (*H*, *S*)) and (*S*, (*H*, *A*)) were frequently recovered with non-overlapping CFs in the 99% posterior tree sample, with CFs of 0.351 (0.313–0.390) and 0.214 (0.179–0.250) respectively (Fig. 6).

3.5. Data availability

A spreadsheet characterizing the 672 loci alignments, as well as all alignments, matrices, trees, and partition files are deposited in the Dryad Digital Repository <http://dx.doi.org/10.5061/dryad.3mr26>. Illumina raw reads for *Sabacon* (SRR2924723) and *Acuclavella* (SRR2924718) have been submitted to NCBI Short Read Archive. All Sanger sequence data generated in this study have been deposited to GenBank (Table 1).

4. Discussion

4.1. Resolution of an ancient trichotomy is aided by increased taxon sampling

Randomly sampling loci from the transcriptome panel shows that hundreds of loci were necessary to reliably infer the topology at the base of Ischyropsalidoidea. The 25-loci sub-sample analyses recovered (*H*, (*A*, *S*)) in only 50% of replicates (Fig. 5). As such, with

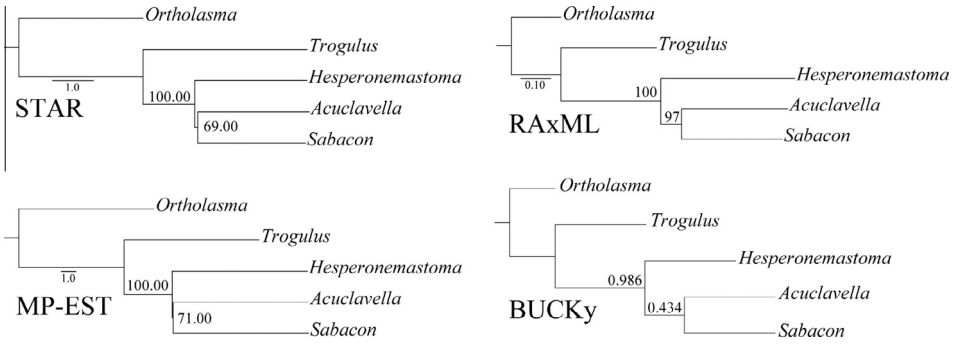

Fig. 4. Transcriptome panel phylogenies and concordance tree. Phylogenies from analyses of 672 loci derived from transcriptomics. Bootstrap support values are shown for partially-parametric coalescent (STAR and MP-EST) and concatenation (RAxML) analyses. Also shown is the BUCKy primary concordance tree and associated concordance factors. Scale bars for STAR and MP-EST are in coalescent units; the RAxML scale depicts the number of substitutions per site.

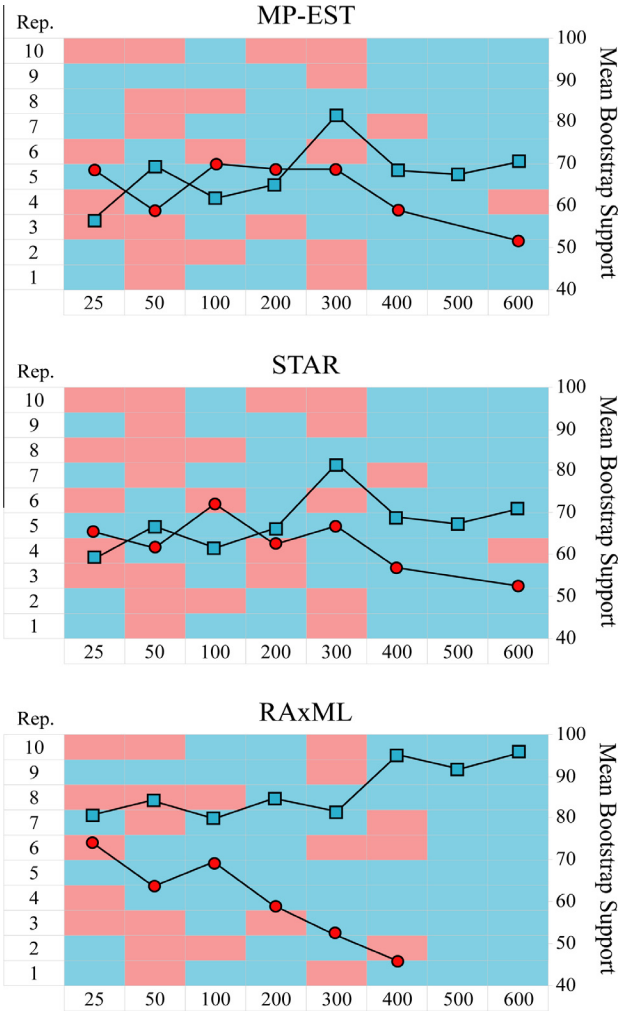

Fig. 5. Comparison across inference methods of randomly sampled loci. Comparison of partially-parametric coalescent (MP-EST and STAR) and concatenated supermatrix (RAxML) methods of phylogenetic inference across randomly sampled loci. Replicates are color-coded to represent recovered topologies. Blue: (*Hesperonemastoma*, (*Sabacon*, *Acuclavella*)). Red: (*A*, (*S*, *H*)). The *x*-axis is the number of loci per replicate. The *y*-axis is the mean bootstrap value, the average values of alternate topologies recovered from each replicate. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

only 14 loci in the expanded panel, the topological congruence between the expanded panel and transcriptome panel may simply be due to chance. An alternative explanation is that the increased taxon sampling in the expanded panel aids phylogenetic inference by shortening the branches leading from the base of Ischyropsalidoidea (Pollock et al., 2002). To explore this possibility, we trimmed the expanded panel to only include OTUs represented in the transcriptome panel (this matrix includes only 1.4% missing data). Phylogenetic analysis of this reduced matrix with RAxML, MrBayes, and *BEAST using the same parameters as on the full expanded panel returns mixed results (Supplement, s4.2). RAxML does not recover *Hesperonemastoma* as sister to *Acuclavella* and *Sabacon* (BSV = 52 for *Sabacon* as sister), MrBayes does (PP = 49), though both of these concatenated analyses are weakly supported. On the other hand, *BEAST recovers *Hesperonemastoma* as sister to the other ischyropsalidoids with PP = 97.8. For ancient radiations, *BEAST may be more robust to reduced taxon sampling compared to other methods of phylogenetic inference used here, though we provide just a single example and this should be further explored. Conversely, *BEAST is typically used to infer shallow evolutionary events, and sampling more than one individual per species is explicitly recommended. In *BEAST, sampling multiple individuals per species allows for more accurate population size estimation, and this in turn may allow for better estimates of divergence times and topology (Heled and Drummond, 2010).

Our findings suggest that increased taxon sampling along descending branches from an ancient near-trichotomy helps with the phylogenetic inference of these diversifications. It has been assumed this would not be the case, because the number of lineages to evaluate sorting along the short critical branch is not increased (Degnan and Rosenberg, 2006; Kubatko and Degnan, 2007). Likely this is due to the additional taxa diminishing phylogenetic artifacts by breaking up long external branches, thus resulting in less LBA (Hillis, 1998). For deep phylogenetic questions, variant sites can become saturated, resulting in abundant homoplasy due to convergence, which is thought to be positively correlated with branch length (Felsenstein, 2004). It is likely that increased taxon sampling diminishes the amount of saturation, which in effect unmasks synapomorphic information along the short internal branch. To this end, our data support the findings of Heled and Drummond (2010) that increased taxon sampling contributes to accurate species tree estimations of rapid radiations. However, Heled and Drummond (2010) couched this argument for shallow phylogenetic inferences, and suggested that increased locus sampling is more important for accurate estimation of deep phylogenetic questions. Additionally, our results suggest using caution when attempting to resolve ancient diversifications using few terminals.

4.2. Emergence of support with supermatrix analyses

High levels of gene tree conflict (Fig. 3) characterize the root of Ischyropsalidoidea. Despite the reported success of partially-parametric coalescent analyses compared to concatenation (e.g.,

Table 3
Occurrence of minority gene trees and probability of equal RFs using a two-sided binomial test.

Treatment	*A*, (*H*, *S*)	*S*, (*H*, *A*)	*p*=
All Loci	238	181	0.0062
BS Values	67	28	<0.0001
% AT3 Content	63	42	0.0504
Hedin et al. (2012)	69	51	0.1203

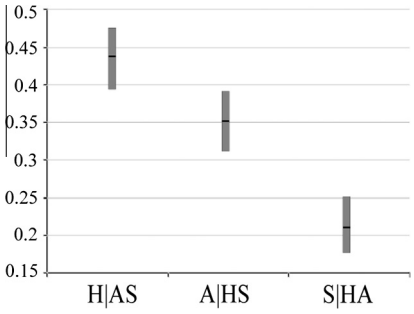

Fig. 6. Concordance factors. Concordance factors of three alternate hypotheses. (H): *Hesperonemastoma*, (A): *Aculavella*, and (S): *Sabacon*. Bars show the estimated concordance factor (black) and 99% confidence interval (gray).

Liu et al., 2009, 2010; Kubatko et al., 2009), and claims that widespread ILS can result in poor performance of concatenated methods (Kubatko and Degnan, 2007), our results seem to indicate that hidden support emerges when analyzing data via concatenated analyses (Gatesy and Springer, 2014), and concatenation appears more robust than shortcut coalescent methods (e.g., Leaché and Rannala, 2011; Patel et al., 2013). Our randomly sub-sampled loci shows that the supermatrix approach to phylogenetic inference, even in the face of high levels of gene tree conflict, results in robust support for nodes otherwise weakly supported by partially-parametric coalescent analyses. This is notable, since others (Edwards et al., 2016) have claimed that sub-sampling phylogenomic data as a test for phylogenetic consistency produces a condition where "the deficiencies of concatenation become even more glaring". They make this claim despite only a single supporting example (from Song et al., 2012); our example does not support this claim. In analyses of mammal relationships, Gatesy and Springer (2014) were able to uncover extreme hidden support for 10 clades by concatenating loci (from McCormack et al., 2012) that strictly conflict with the species tree. To evaluate if hidden support is present in our data, we combined the 420 loci from the transcriptome panel with gene trees strictly conflicting with the species tree, and analyzed this partitioned supermatrix in RAxML using the parameters outlined above. Our analysis did not recover extreme hidden support (Supplement s4.1). However, this may be caused by the unequal minority frequency of the gene trees generated from analysis of these loci. The analyses of Gatesy and Springer included 25 ingroup taxa, and the 5 OTUs representing members of Glires were scattered through the gene trees. Thus, it is difficult to calculate the symmetric frequencies of the many alternate gene tree topologies in their data.

4.3. Interpreting gene tree incongruence and unequal minority resolution frequencies

We infer that divergence at the base of Ischyropsalidoidea happened rapidly enough to allow for ILS to occur. However, gene tree incongruence can arise from factors other than ILS. Springer and Gatesy (2016) were able to show that the high levels of incongruent gene trees, attributed by Song et al. (2012) to ILS, were instead caused by other factors. Among their criticisms of the phylogenetic analyses of gene trees by Song et al. (2012) are factors that potentially apply to this research, including inadequate search algorithms and LBA. First, our PhyML gene tree analyses used default parameter settings including the HKY85 model of nucleotide substitution and the NNI tree search algorithm. More rigorous model selection would likely return gene trees with better log-likelihood values (e.g., Springer and Gatesy, 2016), and NNI has been shown to be prone to being stuck in local optima (Money and Whelan, 2011). However, we also ran each gene tree using more rigorous methods using MrBayes for the BUCKy analysis. The data for both sets of gene tree analyses agree that there is a large amount of gene-tree discordance in our data set, and they agree in the frequency that these alternative topologies are recovered. Second, higher frequencies of a gene tree than its true frequency can be caused by LBA. In our data however, the longest ingroup branch in both the PhyML and RAxML analyses is associated with *Hesperonemastoma*. The average branch lengths in PhyML analyses is 0.400891 for *Hesperonemastoma*, 0.299665 for *Aculavella*, and 0.299444 for *Sabacon*; in the RAxML phylogeny it is 0.443753, 0.333928, and 0.350993, respectively. Under LBA, longer *Hesperonemastoma* branches would be pulled toward outgroup sequences, causing the gene tree stoichiometry to be skewed toward an (*H*, (*A*, *S*)) topology. If this is the case, there is more gene tree incongruence in this system than we report, for (*H*, (*A*, *S*)) is our most frequently recovered topology. As such, we assert that ILS is a likely source of gene tree incongruence in our data.

That being said, the lack of symmetry in the frequency of minority gene trees in the 672 loci does not match theoretical expectations. Since ILS is a random process, expectations are that one topology would be favored, with alternative topologies occurring at equal frequencies (Pamilo and Nei, 1988; Knowles and Kubatko, 2010). For example, in great ape relationships, a study using 11,945 loci found that 76.6% of gene trees support a (*Gorilla*, (*Homo*, *Pan*)) topology, with 11.5% and 11.4% supporting (*H*, (*G*, *P*)) and (*P*, (*H*, *G*)) respectively (Ebersberger et al., 2007). In Ischyropsalidoidea we do not observe such symmetry (Fig. 3): 37.5% of gene trees recover (*Hesperonemastoma*, (*Acuclavella*, *Sabacon*)), 35.4% recover (*A*, (*H*, *S*)), and 27.1% recover (*S*, (*H*, *A*)). This is a classic rooted triplet case, where the probabilities of the three possible gene trees should all equal 33.3% when $\tau = 0$ – a hard polytomy. UMRFs in rooted triplets have been shown to arise from violation of the assumptions of the multispecies coalescent model (Zwickl et al., 2014). Potentially, any of the artifactual or biological causes of gene tree/species tree discordance (Maddison, 1997; Degnan and Rosenberg, 2006), other than ILS, could lead to UMRFs. A recent study by Zwickl et al. (2014) uncovered UMRFs due to both analytical artifacts and biological processes. In some cases, they were able to restore equal minority frequencies of gene trees by using alternative alignment strategies, suggesting that unequal frequencies were caused by artifacts in their methodological pipeline. In other cases, UMRFs persisted across alternative alignment strategies, suggesting that they are the result of biological processes (e.g., introgression).

Zwickl et al. (2014) were able to reduce or eliminate UMRFs that were the product of methodological biases by including intron sequences in their alignments. We are unable to employ their strategy here, for our transcriptome-generated data do not include intronic sequences. Therefore, we attempted to remove UMRFs from our dataset by considering loci that were generated using a different methodological pipeline (Hedin et al., 2012), or by considering locus/gene tree traits thought to improve phylogenomic analyses, including phylogenetic support values (Salichos and Rokas, 2013) and high AT3 content (Romiguier et al., 2013). We found that UMRFs were consistent across these treatments (Table 3), and thus conclude that UMRFs are likely the result of biological processes. However, it is possible that methodological biases persist across these subsets of our data. For example,

convergence in the base composition between two taxa could skew the stoichiometry of a topology combining these taxa to be more common that its true frequency (Springer and Gatesy, 2016).

The cause of UMRFs in our system could result from any violation of the multispecies coalescent model. Although we don't consider undetected paralogy as the likely cause of this discordance, due to low frequency of paralogs detected by our filtering criteria, this is one possibility. Other biological processes that are more likely to apply to Ischyropsalidoidea include ancestral population structure and paraphyletic gene flow. Population structure has been shown to cause UMRFs when subdivision is present in the ancestor of three lineages and persists through both speciation events (Slatkin and Pollack, 2008). This may initially seem unlikely, but many harvestmen lineages are known to show extreme population structure, as are many nonvagile terrestrial arthropods (e.g., Derkarabetian et al., 2011; Keith and Hedin, 2012). Likewise, it can be inferred from Leaché et al. (2014a) that paraphyletic gene flow, or gene flow between species that are not sister taxa, can increase the frequency of gene trees grouping these taxa together. Also, the total branch lengths of the transcriptome panel RAxML analysis from the base of Ischyropsalidoidea to the tip of *Hesperonemastoma*, *Acuclavella*, and *Sabacon* are 35.0%, 33.2%, and 31.8% of the total of these sums respectively, which more closely matches the asymmetry of gene trees above than do theoretical expectations, suggesting that selection or evolutionary rates may be playing a role.

To us, inferring the population genetics of lineages that underwent successive diversifications around 200 MYA (Schönhofer et al., 2013; Sharma and Giribet, 2014) seems a near-futile effort. Thus far, species tree analyses for the most part have been robust in their inference of nodes with UMRFs (Zwickl et al., 2014), though estimating the correct species tree can become difficult when the biological processes underlying this discordance are severe (e.g. Leaché et al., 2014a). Since species tree resolution within Ischyropsalidoidea is our primary goal, we do not further seek out the source the UMRFs recovered here, under the assumption that it is not severely affecting our phylogenetic inference. Future research on difficult phylogenetic nodes could employ a strategy where the likelihood of the data are analyzed with respect to *a priori* models that vary in ancestral population structure, timing of divergences, etc., with the most likely model selected using a criterion score (e.g., Carstens et al., 2013). Clearly the causes of UMRFs and the consequences to phylogenetic inference should be the focus of future research.

4.4. Ischyropsalidoidea systematics

Giribet and Kury (2007) suggested waiting on taxonomic amendments within Ischyropsalidoidea until inclusion of *Acuclavella* and *Crosbycus* allowed for rigorous testing of family-level relationships. Despite this suggestion, ischyropsalidoid familial amendments with poor morphological and molecular diagnoses have continued (e.g., Giribet et al., 2010; Schönhofer, 2013). For example, "Sabaconidae" *sensu* Giribet et al. (2010) was erected in spite of the authors not being able to find a single morphological synapomorphy and low support (jackknife value < 50) for a clade comprising *Sabacon* sister to *Hesperonemastoma* + *Taracus*. Note that we do not recover this clade here. Likewise, "Taracidae" as defined by Schönhofer (2013) is not supported by morphological synapomorphies. Our Fig. 2 shows that *Taracus* and *Hesperonemastoma*, two morphologically very different genera, are recovered as a clade with high support, but that the root for this "family" is about twice as ancient as his definitions of Sabaconidae and Ischyropsalididae. In light of these recent failures, we do not propose new familial diagnoses, for we agree with Giribet and Kury (2007) that such amendments and definitions would be premature without the inclusion of all ischyropsalidoid genera. Additionally, the CF for (*H*, (*A*, *S*)) is less than 0.5 with 99% confidence, and thus perhaps does not warrant formal taxonomic recognition (Baum, 2007).

In comparison, our results corroborate other taxonomic hypotheses and suggestions. For example, recent work on *Sabacon* (Schönhofer et al., 2013; Martens, 2015) recommended splitting this genus into multiple genera. The expanded panel phylogeny (Fig. 2) recovers the divergence of *Sabacon* approximately twice as deep in time as *Hesperonemastoma*, the next most divergent genus. In fact, *Sabacon* is almost as internally divergent as the genus *Ischyropsalis* is from *Acuclavella* plus *Ceratolasma*. Additionally, the redefinition of Ischyropsalididae (Schönhofer, 2013) to include *Acuclavella* and *Ceratolasma* (formally in Ceratolasmatidae) results in this family having a very similar crown age with his definition of Sabaconidae. Thus, our results compliment the taxonomic conclusions of Schönhofer et al. (2013) and Schönhofer (2013). Explicitly testing family hypotheses within Ischyropsalidoidea with the inclusion of all genera and defining genera within *Sabacon* should be higher-level taxonomic research priorities within Ischyropsalidoidea.

5. Conclusions

We have recovered a single short branch deep in the phylogeny of Ischyropsalidoidea. Despite high levels of gene tree conflict, we consistently recover *Hesperonemastoma* + *Taracus* sister to remaining ischyropsalidoids across different analytical methods with different strategies of taxon and locus sampling. Though the short internal branch deep in the ischyropsalidoid phylogeny is consistently recovered, it is associated with high levels of gene tree conflict and relatively poor support values. These characteristics are precisely those associated with topological conflict between coalescent- and supermatrix-based methods of phylogenetic inference (Lambert et al., 2015). That being said, the causes of gene tree conflict associated with ancient short internal branches should continue to be explored. Particularly, simulation analyses should explore the effects of ILS and UMRFs at the time of divergence on phylogenetic analyses, with these simulations extended to deep time to assess if this signal degrades though time. If gene tree conflict initially associated with ILS degrades to gene tree conflict associated with sampling error in ancient diversifications, than coalescent methods that attribute all such conflict to ILS may be inappropriate. As phylogenomics comes of age, transposon insertions have been used to independently assess incongruence at ancient nodes, suggesting that incongruence in gene trees is largely due to sampling error (Gatesy and Springer, 2014), and assessing this signal could inform the appropriateness of phylogenomic analyses. Further, these analyses should manipulate the underlying assumptions of the simple coalescent model, such as selection and variation in population size, to assess the impacts of these parameters on the frequency of gene tree incongruence (Nordborg, 2001; Scally et al., 2012; Springer and Gatesy, 2016).

Arguably the strongest evidence for phylogenetic hypotheses occurs when clades are recovered from independent lines of evidence (Rota-Stabelli et al., 2011). The expanded panel results suggest that increased taxon sampling may be as important for estimating ancient radiations as increased locus sampling, for these additional taxa may decrease saturation and increase phylogenetic signal along problematic branches. The impact of taxon sampling on the phylogenetic reconstruction of ancient radiations should also be an area of future research.

Acknowledgments

Alexa Feist and Yixuan Xia helped characterize alignments generated via transcriptomics. Kristen Emata, Angela DiDomenico, and

Timothy Shaw helped troubleshoot analyses. William A. Shear helped identify *Taracus* species. Axel Schönhofer and Jochen Martens provided *Ischyropsalis* genomic material. Shahan Derkarabetian, Jim Starrett, and Amy Hubert aided in RNA extractions, and S. Derkarabetian and Dave Carlson aided in transcriptome assembly. This research was improved via discussions with Kevin Burns, Dean Leavitt, John Gatesy, Rob Meredith, Tod Reeder, Bill Shear, Mark Springer, and J. Starrett. We thank Jean Valjean for asking tough questions. Specimens were collected with aid from D. Carlson, S. Derkarabetian, Damian Elias, Marc P. Hayes, S. Huber, Robert Keith, Steve Merkley, T. Novak, Adrienne Richart, Dan Richart, Jordan Satler, A. Schönhofer, L. Slana, J. Starrett, B. Shear, and Jeff Underwood. This manuscript was improved through comments from Mercedes Burns, Allan Cabrero, D. Carlson, S. Derkarabetian, K. Emata, J. Starrett, and two anonymous reviewers. This project was funded in part by the University of California Riverside Newell Travel Award, and the American Arachnological Society Vincent Roth Fund for Research in Systematics. Some data collection was supported by a California State University Program for Education and Research in Biotechnology (CSUPERB) grant to M. Hedin.

Appendix. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.11.010.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

Ané, C., 2010. Reconstructing concordance trees and testing the coalescent model from genome-wide data sets. In: Knowles, L.L., Kubatko, L.S. (Eds.), Estimating Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell, Hoboken, New Jersey (Chapter 3).

Ané, C., Larget, B., Baum, D.A., Smith, S.D., Rokas, A., 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24 (2), 412–426. http://dx.doi.org/10.1093/molbev/msl170.

Asahida, T., Kobayashi, T., Saitoh, K., Nakayama, I., 1996. Tissue preservation and total DNA extraction from fish stored at ambient temperature using buffers containing high concentration of urea. Fish. Sci. 62 (5), 727–730.

Baum, D.A., 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. Taxon 56 (2), 417–426.

Bayzid, M.S., Warnow, T., 2013. Naive binning improves phylogenomic analyses. Bioinformatics 29 (18), 2277–2284. http://dx.doi.org/10.1093/bioinformatics/btt394.

Betancur-R, R., Naylor, G.J.P., Ortí, G., 2014. Conserved genes, sampling error, and phylogenomic inference. Syst. Biol. 63 (2), 257–262. http://dx.doi.org/10.1093/sysbio/syt073.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42 (3), 384–397. http://dx.doi.org/10.2307/2992473.

Carstens, B.C., Brennan, R.S., Chua, V., Duffie, C.V., Harvey, M.G., Koch, R.A., McMahan, C.D., Nelson, B.J., Newman, C.E., Satler, J.D., Seeholzer, G., Posbic, K., Tank, D.C., Sullivan, J., 2013. Model selection as a tool for phylogeographic inference: an example from the willow *Salix melanopsis*. Mol. Ecol. 22 (15), 4014–4028. http://dx.doi.org/10.1111/mec.12347.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17 (4), 540–552.

Clouse, R.M., Sharma, P.P., Giribet, G., Wheeler, W.C., 2013. Elongation factor-1α, a putative single-copy nuclear gene, has divergent sets of paralogs in an arachnid. Mol. Phylogenet. Evol. 68 (3), 471–481. http://dx.doi.org/10.1016/j.ympev.2013.04.025.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9 (8), 772.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2 (5), e68. http://dx.doi.org/10.1371/journal.pgen.0020068.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24 (6), 332–340. http://dx.doi.org/10.1016/j.tree.2009.01.009.

Derkarabetian, S., Hedin, M., 2014. Integrative taxonomy and species delimitation in harvestmen: a revision of the western North American genus *Sclerobunus* (Opiliones: Laniatores: Travunioidea). PLoS One 9 (8), e104982. http://dx.doi.org/10.1371/journal.pone.0104982.

Derkarabetian, S., Ledford, J., Hedin, M., 2011. Genetic diversification without obvious genitalic morphological divergence in harvestmen (Opiliones, Laniatores, *Sclerobunus robustus*) from montane sky islands of western North America. Mol. Phylogenet. Evol. 61, 844–853.

Dresco, E., 1970. Recherches sur la variabilité et la phylogénie chez les Opiliones du genre *Ischyropsalis* CL Koch (Fam. Ischyropsalidae), avec la creation de la famille nouvelle des Sabaconidae. Bull. Mus. Natl. d'Histoire Naturelle, 2e Serie 41, 1200–1213.

Ebersberger, I., Galgoczy, P., Taudien, S., Taenzer, S., Platzer, M., von Haeseler, A., 2007. Mapping human genetic ancestry. Mol. Biol. Evol. 24 (10), 2266–2276. http://dx.doi.org/10.1093/molbev/msm156.

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. 104 (14), 5936–5941.

Edwards, S.V., Xi, Z., Janke, A., Faircloth, B.C., McCormack, J.E., Glenn, T.C., Zhong, B., Wu, S., Lemmon, E.M., Lemmon, A.R., Leaché, A.D., Liu, L., Davis, C.C., 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. Part A 94, 447–462. http://dx.doi.org/10.1016/j.ympev.2015.10.027.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61 (5), 717–726. http://dx.doi.org/10.1093/sysbio/sys004.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates.

Gatesy, J., Springer, M.S., 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80, 231–266. http://dx.doi.org/10.1016/j.ympev.2014.08.013.

Gatesy, J., O'Grady, P., Baker, R.H., 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level Artiodactyl Taxa. Cladistics 15 (3), 271–313. http://dx.doi.org/10.1111/j.1096-0031.1999.tb00268.x.

Giribet, G., Kury, A.B., 2007. Phylogeny and biogeography. In: Pinto-da-Rocha, R., Machado, G., Giribet, G. (Eds.), Harvestmen: The Biology of Opiliones. Harvard University Press, Cambridge, MA, pp. 62–87 (Chapter 3).

Giribet, G., Sharma, P.P., 2015. Evolutionary biology of harvestmen (Arachnida, Opiliones). Annu. Rev. Entomol. 60 (1), 157–175. http://dx.doi.org/10.1146/annurev-ento-010814-021028.

Giribet, G., Rambla, M., Carranza, S., Baguñà, J., Riutort, M., Ribera, C., 1999. Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S and partial 28S ribosomal DNA sequences and morphology. Mol. Phylogenet. Evol. 11 (2), 296–307. http://dx.doi.org/10.1006/mpev.1998.0583.

Giribet, G., Vogt, L., Pérez González, A., Sharma, P., Kury, A.B., 2010. A multilocus approach to harvestman (Arachnida: Opiliones) phylogeny with emphasis on biogeography and the systematics of Laniatores. Cladistics 26 (4), 408–437. http://dx.doi.org/10.1111/j.1096-0031.2009.00296.x.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29 (7), 644–652. http://dx.doi.org/10.1038/nbt.1883.

Gruber, J., 1978. Redescription of *Ceratolasma tricantha* Goodnight and Goodnight, with notes on the family Ischyropsalidae (Opiliones, Palpatores). J. Arachnol. 6, 105–124.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52 (5), 696–704. http://dx.doi.org/10.1080/10635150390235520.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321. http://dx.doi.org/10.1093/sysbio/syq010.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M., Eccles, D., Li, B., Lieber, M., MacManes, M., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8 (8), 1494–1512. http://dx.doi.org/10.1038/nprot.2013.084.

Harvey, M.S., 2002. The neglected cousins: What do we know about the smaller arachnid orders? J. Arachnol. 30 (2), 357–372. http://dx.doi.org/10.1636/0161-8202(2002)030[0357:TNCWDW]2.0.CO;2.

Hedin, M., Derkarabetian, S., McCormack, M., Richart, C., Shultz, J.W., 2010. The phylogenetic utility of the nuclear protein-coding gene EF-1α for resolving recent divergences in Opiliones, emphasizing intron evolution. J. Arachnol. 38 (1), 9–20. http://dx.doi.org/10.1636/HA09-49.1.

Hedin, M., Starrett, J., Akhter, S., Schönhofer, A.L., Shultz, J.W., 2012. Phylogenomic resolution of Paleozoic divergences in Harvestmen (Arachnida, Opiliones) via analysis of next-generation transcriptome data. PLoS One 7 (8), e42888. http://dx.doi.org/10.1371/journal.pone.0042888.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27 (3), 570–580. http://dx.doi.org/10.1093/molbev/msp274.

Hillis, D.M., 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47 (1), 3–8. http://dx.doi.org/10.1080/106351598260987.

Huelsenbeck, J.P., Hillis, D.M., 1993. Success of phylogenetic methods in the four-taxon case. Syst. Biol. 42 (3), 247–264. http://dx.doi.org/10.1093/sysbio/42.3.247.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., Miyata, T., 1989. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. 86 (23), 9355–9359.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30 (14), 3059–3066. http://dx.doi.org/10.1093/nar/gkf436.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649. http://dx.doi.org/10.1093/bioinformatics/bts199.

Keith, R., Hedin, M., 2012. Extreme mitochondrial population subdivision in southern Appalachian paleoendemic spiders (Araneae: Hypochilidae: Hypochilus), with implications for species delimitation. J. Arachnol. 40 (2), 167–181. http://dx.doi.org/10.1636/A11-49.1.

Kingman, J.F., 1982. The coalescent. Stoch. Process. Appl. 13 (3), 235–248. http://dx.doi.org/10.1016/0304-4149(82)90011-4.

Knowles, L.L., Kubatko, L.S., 2010. Estimating species trees: an introduction to concepts and models. In: Knowles, L.L., Kubatko, L.S. (Eds.), Estimating Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell, Hoboken, New Jersey (Chapter 1).

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56 (1), 17–24. http://dx.doi.org/10.1080/10635150601146041.

Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25 (7), 971–973. http://dx.doi.org/10.1093/bioinformatics/btp079.

Kury, A.B., 2013. Order Opiliones Sundevall, 1833. In: Zhang, Z.-Q. (Ed.), Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Zootaxa 3703 (1), 27–33. http://dx.doi.org/10.11646/zootaxa.3703.1.7.

Kury, A., Mendes, A., Souza, D., 2014. World checklist of Opiliones species (Arachnida). Part 1: Laniatores–Travunioidea and Triaenonychoidea. Biodivers. Data J. 2, e4094. http://dx.doi.org/10.3897/BDJ.2.e4094.

Lambert, S.M., Reeder, T.W., Wiens, J.J., 2015. When do species-tree and concatenated estimates disagree? An empirical analysis with higher-level scincid lizard phylogeny. Mol. Phylogenet. Evol. Part A 82, 146–155. http://dx.doi.org/10.1016/j.ympev.2014.10.004.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29 (6), 1695–1701. http://dx.doi.org/10.1093/molbev/mss020.

Lang, J.M., Darling, A.E., Eisen, J.A., 2013. Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. PLoS One 8 (4), e62510. http://dx.doi.org/10.1371/journal.pone.0062510.

Lanier, H.C., Knowles, L.L., 2015. Applying species-tree analyses to deep phylogenetic histories: challenges and potential suggested from a survey of empirical phylogenetic studies. Mol. Phylogenet. Evol. 83, 191–199. http://dx.doi.org/10.1016/j.ympev.2014.10.022.

Larget, B.R., Kotha, S.K., Dewey, C.N., Ané, C., 2010. BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. Bioinformatics 26 (22), 2910–2911. http://dx.doi.org/10.1093/bioinformatics/btq539.

Leaché, A.D., Rannala, B., 2011. The accuracy of species tree estimation under simulation: a comparison of methods. Syst. Biol. 60 (2), 126–137. http://dx.doi.org/10.1093/sysbio/syq073.

Leaché, A.D., Harris, R.B., Rannala, B., Yang, Z., 2014a. The influence of gene flow on species tree estimation: a simulation study. Syst. Biol. 63 (1), 17–30. http://dx.doi.org/10.1093/sysbio/syt049.

Leaché, A.D., Wagner, P., Linkem, C.W., Böhme, W., Papenfuss, T.J., Chong, R.A., Lavin, B.R., Bauer, A.M., Nielsen, S.V., Greenbaum, E., Rödel, M.-O., Schmitz, A., LeBreton, M., Ineich, I., Chirio, L., Ofori-Boateng, C., Eniang, E.A., Baha El Din, S., Lemmon, A.R., Burbrink, F.T., 2014b. A hybrid phylogenetic–phylogenomic approach for species tree estimation in African Agama lizards with applications to biogeography, character evolution, and diversification. Mol. Phylogenet. Evol. 79, 215–230. http://dx.doi.org/10.1016/j.ympev.2014.06.013.

Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44 (1), 99–121. http://dx.doi.org/10.1146/annurev-ecolsys-110512-135822.

Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. 58 (4), 452–460. http://dx.doi.org/10.1093/sysbio/syp034.

Liu, L., Pearl, D.K., 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56 (3), 504–514. http://dx.doi.org/10.1080/10635150701429982.

Liu, L., Yu, L., Pearl, D.K., Edwards, S.V., 2009. Estimating species phylogenies using coalescence times among sequences. Syst. Biol. 58 (5), 468–477. http://dx.doi.org/10.1093/sysbio/syp031.

Liu, L., Yu, L., Edwards, S., 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10 (1), 302. http://dx.doi.org/10.1186/1471-2148-10-302.

Machado, G., Pinto-da-Rocha, R., Giribet, G., 2007. What are harvestmen? In: Pinto-da-Rocha, R., Machado, G., Giribet, G. (Eds.), Harvestmen: The Biology of Opiliones. Harvard University Press, Cambridge, MA, pp. 1–13 (Chapter 1).

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46 (3), 523–536. http://dx.doi.org/10.1093/sysbio/46.3.523.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55 (1), 21–30. http://dx.doi.org/10.1080/10635150500354928.

Martens, J., 1976. Genitalmorphologie, system und phylogenie der Weberknechte (Arachnida: Opiliones). Entomol. Germanica 3 (1–2), 51–68.

Martens, J., 2015. *Sabacon* Simon, 1879 in the Palaearctic: a survey of new and known species from France, Nepal, India, China, Russia and Japan (Arachnida: Opiliones: Sabaconidae). In: Biodiversität und Naturausstattung im Himalaya V. Hartmann and Weipert, Erfurt, pp. 167–210.

Martens, J., Hoheisel, U., Götze, M., 1981. Vergleichende anatomie der legeröhren der Opiliones als beitrag zur phylogenie der ordnung (Arachnida). Zoologische Jahrbücher. Abteilung für Anatomie und Ontogenie der Tiere Abteilung für Anatomie und Ontogenie der Tiere 105, 13–76.

McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 22 (4), 746–754. http://dx.doi.org/10.1101/gr.125864.111.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLoS One 8 (1), e54848.

Meredith, R.W., Janečka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simão, T.L.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M., Murphy, W.J., 2011. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 334 (6055), 521–524. http://dx.doi.org/10.1126/science.1211028.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 November 2010, New Orleans, LA, pp. 1–8. http://dx.doi.org/10.1109/GCE.2010.5676129.

Money, D., Whelan, S., 2011. Characterizing the phylogenetic tree-search problem. Syst. Biol., syr097 http://dx.doi.org/10.1093/sysbio/syr097.

Murphy, W.J., Eizirik, E., O'Brien, S.J., Madsen, O., Scally, M., Douady, C.J., Teeling, E., Ryder, O.A., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294 (5550), 2348–2351. http://dx.doi.org/10.1126/science.1067179.

Nordborg, M., 2001. Coalescent theory. In: Balding, D., Bishop, M., Cannings, C. (Eds.), Handbook of Statistical Genetics. Wiley, Chichester, pp. 179–212.

O'Neill, E.M., Schwartz, R., Bullock, C.T., Williams, J.S., Shaffer, H.B., Aguilar-Miguel, X., Parra-Olea, G., Weisrock, D.W., 2013. Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. Mol. Ecol. 22, 111–129. http://dx.doi.org/10.1111/mec.12049.

Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5 (5), 568–583.

Patel, S., Kimball, R.T., Braun, E.L., 2013. Error in phylogenetic estimation for bushes in the tree of life. J. Phylogenet. Evol. Biol. 1 (2), 110. http://dx.doi.org/10.4172/2329-9002.1000110.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9 (3), e1000602. http://dx.doi.org/10.1371/journal.pbio.1000602.

Pollock, D.D., Zwickl, D.J., McGuire, J.A., Hillis, D.M., 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51 (4), 664–671. http://dx.doi.org/10.1080/10635150290102357.

Raymann, K., Forterre, P., Brochier-Armanet, C., Gribaldo, S., 2014. Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in Archaea. Genome Biol. Evol. 6 (1), 192–212. http://dx.doi.org/10.1093/gbe/evu004.

Regier, J.C., Shultz, J.W., Ganley, A.R.D., Hussey, A., Shi, D., Ball, B., Zwick, A., Stajich, J.E., Cummings, M.P., Martin, J.W., Cunningham, C.W., 2008. Resolving arthropod phylogeny: exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequence. Syst. Biol. 57 (6), 920–938. http://dx.doi.org/10.1080/10635150802570791.

Richart, C., Hedin, M., 2013. Three new species in the harvestmen genus *Acuclavella* (Opiliones, Dyspnoi, Ischyropsalidoidea), including description of male *Acuclavella quattuor* Shear, 1986. ZooKeys 311, 19–68. http://dx.doi.org/10.3897/zookeys.311.2920.

Romiguier, J., Ranwez, V., Delsuc, F., Galtier, N., Douzery, E.J., 2013. Less is more in mammalian phylogenomics: AT-rich genes minimize tree conflicts and unravel the root of placental mammals. Mol. Biol. Evol. 30 (9), 2134–2144. http://dx.doi.org/10.1093/molbev/mst116.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61 (3), 539–542. http://dx.doi.org/10.1093/sysbio/sys029.

Rota-Stabelli, O., Campbell, L., Brinkmann, H., Edgecombe, G.D., Longhorn, S.J., Peterson, K.J., Pisani, D., Philippe, H., Telford, M.J., 2011. A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support

monophyletic Mandibulata. Proc. R. Soc. B: Biol. Sci. 278 (1703), 298–306. http://dx.doi.org/10.1098/rspb.2010.0590.
Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. In: Misener, S., Krawetz, S.A. (Eds.), Bioinformatics Methods and Protocols. Humana Press, pp. 365–386.
Salichos, L., Rokas, A., 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497 (7449), 327–331. http://dx.doi.org/10.1038/nature12130.
Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., et al., 2012. Insights into hominid evolution from the gorilla genome sequence. Nature 483 (7388), 169–175. http://dx.doi.org/10.1038/nature10842.
Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27 (6), 863–864. http://dx.doi.org/10.1093/bioinformatics/btr026.
Schönhofer, A.L., 2013. A taxonomic catalogue of the Dyspnoi Hansen and Sørensen, 1904 (Arachnida: Opiliones). Zootaxa 3679 (1), 1–68.
Schönhofer, A.L., McCormack, M., Tsurusaki, N., Martens, J., Hedin, M., 2013. Molecular phylogeny of the harvestmen genus *Sabacon* (Arachnida: Opiliones: Dyspnoi) reveals multiple Eocene–Oligocene intercontinental dispersal events in the Holarctic. Mol. Phylogenet. Evol. 66 (1), 303–315. http://dx.doi.org/10.1016/j.ympev.2012.10.001.
Schönhofer, A.L., Vernesi, C., Martens, J., Hedin, M., 2015. Molecular phylogeny, biogeographic history, and evolution of cave-dwelling taxa in the European harvestmen genus *Ischyropsalis* (Opiliones: Dyspnoi). J. Arachnol. 43 (1), 40–53. http://dx.doi.org/10.1636/H14-39.1.
Seo, T.K., 2008. Calculating bootstrap probabilities of phylogeny using multilocus sequence data. Mol. Biol. Evol. 25 (5), 960–971. http://dx.doi.org/10.1093/molbev/msn043.
Sharma, P., Giribet, G., 2014. A revised dated phylogeny of the arachnid order Opiliones. Evol. Popul. Genet. 5, 255. http://dx.doi.org/10.3389/fgene.2014.00255.
Shaw, T.I., Ruan, Z., Glenn, T.C., Liu, L., 2013. STRAW: species TRee Analysis Web server. Nucleic Acids Res. 41 (W1), W238–W241. http://dx.doi.org/10.1093/nar/gkt377.
Shear, W.A., 1986. A cladistic analysis of the opilionid superfamily Ischyropsalidoidea, with descriptions of the new family Ceratolasmatidae, the new genus *Acuclavella*, and four new species. Am. Mus. Novitates, 2844, 1–29. URI: http://hdl.handle.net/2246/3579.
Shultz, J.W., 1998. Phylogeny of Opiliones (Arachnida): an assessment of the "Cyphopalpatores" Concept. J. Arachnol. 26 (3), 257–272.
Shultz, J.W., Regier, J.C., 2001. Phylogenetic analysis of Phalangida (Arachnida, Opiliones) using two nuclear protein-encoding genes supports monophyly of Palpatores. J. Arachnol. 29 (2), 189–200.
Slatkin, M., Pollack, J.L., 2008. Subdivision in an ancestral species creates asymmetry in gene trees. Mol. Biol. Evol. 25 (10), 2241–2246. http://dx.doi.org/10.1093/molbev/msn172.
Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl. Acad. Sci. 109 (37), 14942–14947. http://dx.doi.org/10.1073/pnas.1211733109.
Springer, M.S., Gatesy, J., 2014. Land plant origins and coalescence confusion. Trends Plant Sci. 19 (5), 267–269. http://dx.doi.org/10.1016/j.tplants.2014.02.012.
Springer, M.S., Gatesy, J., 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94, 1–33. http://dx.doi.org/10.1016/j.ympev.2015.07.018.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57 (5), 758–771. http://dx.doi.org/10.1080/10635150802429642.
Swofford, D.L., Waddell, P.J., Huelsenbeck, J.P., Foster, P.G., Lewis, P.O., Rogers, J.S., 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Syst. Biol. 50 (4), 525–539, http://www.jstor.org/stable/3070852.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30 (12), 2725–2729. http://dx.doi.org/10.1093/molbev/mst197.
Teeling, E.C., Hedges, S.B., 2013. Making the impossible possible: rooting the tree of placental mammals. Mol. Biol. Evol. 30 (9), 1999–2000. http://dx.doi.org/10.1093/molbev/mst118.
Vink, C.J., Thomas, S.M., Paquin, P., Hayashi, C.Y., Hedin, M., 2005. The effects of preservatives and temperatures on arachnid DNA. Inver. Syst. 19 (2), 99–104. http://dx.doi.org/10.1071/IS04039.
Zhang, N., Zeng, L., Shan, H., Ma, H., 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. New Phytol. 195 (4), 923–937. http://dx.doi.org/10.1111/j.1469-8137.2012.04212.x.
Zwickl, D.J., Stein, J.C., Wing, R.A., Ware, D., Sanderson, M.J., 2014. Disentangling methodological and biological sources of gene tree discordance on *Oryza* (Poaceae) Chromosome 3. Syst. Biol. 63 (5), 645–659. http://dx.doi.org/10.1093/sysbio/syu027.