Molecular Phylogenetics and Evolution 127 (2018) 907–918

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea)

Hannah M. Wood[a,*], Vanessa L. González[a], Michael Lloyd[a], Jonathan Coddington[a], Nikolaj Scharff[b]

[a] Smithsonian Institution, National Museum of Natural History, 10th and Constitution Ave. NW, Washington, D.C. 20560-0105, U.S.A.
[b] Biodiversity Section, Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark

ARTICLE INFO

Keywords:
Ultra conserved elements
Exon
Ethanol
Araneomorphae

ABSTRACT

Historical museum specimens are invaluable for morphological and taxonomic research, but typically the DNA is degraded making traditional sequencing techniques difficult to impossible for many specimens. Recent advances in Next-Generation Sequencing, specifically target capture, makes use of short fragment sizes typical of degraded DNA, opening up the possibilities for gathering genomic data from museum specimens. This study uses museum specimens and recent target capture sequencing techniques to sequence both Ultra-Conserved Elements (UCE) and exonic regions for lineages that span the modern spiders, Araneomorphae, with a focus on Palpimanoidea. While many previous studies have used target capture techniques on dried museum specimens (for example, skins, pinned insects), this study includes specimens that were collected over the last two decades and stored in 70% ethanol at room temperature. Our findings support the utility of target capture methods for examining deep relationships within Araneomorphae: sequences from both UCE and exonic loci were important for resolving relationships; a monophyletic Palpimanoidea was recovered in many analyses and there was strong support for family and generic-level palpimanoid relationships. Ancestral character state reconstructions reveal that the highly modified carapace observed in mecysmaucheniids and archaeids has evolved independently.

1. Introduction

Spiders are important predators in terrestrial ecosystems, and with over 47,000 described species (World Spider Catalog, 2018) they are notable in terms of global ubiquity, diversity in behavior and ecology, and medical importance. Spiders rank as the seventh most diverse arthropod order, and total species diversity will likely triple as more species are discovered and named (Coddington and Levi, 1991). However, until very recently research in spider phylogenetics was hampered by a scarcity of genetic markers with few (e.g., six or less) available for phylogenetic inference using traditional Sanger sequencing approaches (e.g., Wheeler et al., 2016; Dimitrov et al., 2017). This small set of genetic markers has been mostly unsuccessful at resolving deeper relationships at the interfamilial level (Wheeler et al., 2016). Phylogenetic resolution has been further confounded by insufficient taxon sampling, which is complicated by the difficulty field biologists face in gathering genetic samples for taxonomically diverse organisms that are rare and elusive, or that live in remote, inaccessible areas. Natural history museums, with the goal of documenting the diversity of life, provide an additional source of genetic resources as they have amassed large collections over a long period of time, which include rare or extinct species and specimens from remote areas of the world. Historical museum specimens are invaluable for morphological and taxonomic research, but typically their DNA is degraded making traditional sequencing techniques difficult to impossible for many specimens (Wandeler et al., 2007).

Recent advances in Next-Generation Sequencing (NGS), specifically target capture, makes use of short fragment sizes typical of degraded DNA, opening up the possibility of gathering genomic data for phylogenetic analysis from existing museum specimens (Jones and Good, 2016). Target capture techniques have been successfully performed on dried (i.e., pinned insects or dried skins) museum specimens over 100 years old for birds (McCormack et al., 2016), mammals (Bi et al., 2013; Guschanski et al., 2013; Bailey et al., 2016), and insects (Blaimer

* Corresponding author.
E-mail address: woodh@si.edu (H.M. Wood).

https://doi.org/10.1016/j.ympev.2018.06.038
Received 22 December 2017; Received in revised form 14 June 2018; Accepted 22 June 2018
Available online 30 June 2018
1055-7903/ Published by Elsevier Inc.

et al., 2016). Genetic fragments have even been recovered from fluid preserved museum specimens (Ruane and Austin, 2017), including a lizard that was collected 145 years ago (McGuire et al., 2018). Target capture techniques have also been used to infer arachnid phylogeny (Starrett et al., 2017), including using some museum material (Hamilton et al., 2016; Hedin et al., 2018).

Arachnid specimens in museum collections are typically stored in 70–80% ethanol and at these concentrations at room temperature DNA degrades by oxidation and hydrolysis (Quicke et al., 1999; Vink et al., 2005). For such specimens with fragmented DNA, NGS target capture techniques are preferred to traditional techniques, such as Sanger sequencing of PCR amplified specific genes or gene fragments, because large amounts of data can be generated rapidly and at relatively low cost. Target capture techniques can gather data from numerous loci throughout the genome. This approach can dramatically advance research in spider systematics in two ways: first, the ability to sequence hundreds of regions throughout the genome may resolve deep interfamilial relationships more accurately; second, legacy and archival museum collections offer cheap and efficient ways to include rare lineages that currently are difficult to collect.

Palpimanoidea is an ancient spider lineage that has evolved some remarkable morphologies in the carapace and chelicerae (functionally equivalent to jaws or mandibles) compared to other spiders (Wood et al., 2012, 2016) (Fig. 1). Their strange morphology complements their unusual and highly specialized predatory behaviors: for example, mecysmaucheniids use ballistic trap-jaw strikes (Wood et al., 2016), and archaeids are specialists that attack other spiders at a distance with their long chelicerae (Wood et al., 2012). As a possible sister clade (Wood et al., 2012; Wheeler et al., 2016; Fernández et al., 2018) to the more modern, derived spiders, the Entelegynae, palpimanoids are an important group for understanding spider evolution. Palpimanoids are paleoendemics, and are currently mostly confined to the Southern Hemisphere, although there is an extensive fossil record from the Northern Hemisphere, with distribution patterns of some extant lineages likely relating to Gondwanan vicariance (Wood et al., 2013), . Phylogenetic analysis using both morphological data and four genetic markers strongly supports a monophyletic Palpimanoidea that consists of five families (Wood et al., 2012, 2013): Archaeidae, Huttoniidae, Mecysmaucheniidae, Palpimanidae, and Stenochilidae. Another analysis using only molecular data from six genetic markers (not all markers were successfully sequenced for all terminals) suggested a paraphyletic Palpimanoidea, although with weak branch support (Wheeler et al., 2016), but a recent phylogeny based on transcriptomic data strongly supported the monophyly of Palpimanoidea (Fernández et al., 2018). Phylogenetic relationships among palpimanoid families were weakly supported in these previous studies, corroborating that the standard genetic markers used in spider studies are not adequate for resolving deep relationships (Agnarsson et al., 2013), and although transcriptomic data helped confirm the monophyly of Palpimanoidea it did not help resolve internal nodes. Palpimanoids occur in many remote areas of the world that are hard to access, so that thorough taxon sampling for the group is difficult. However, there are many representatives of palpimanoids in museum collections, mostly collected over the last two decades, stored in 75–95% ethanol. Because palpimanoids are ancient spiders that were once more widespread, because of their phylogenetic placement as possible sister to a major clade of spider (Entelegynae), and because of their unusual predatory behaviors and morphology, they are an excellent group to study in order to examine phylogeny, trait evolution, biogeography, and diversification patterns in relictual lineages.

This study uses such museum specimens, combined with recent target capture sequencing techniques, to examine phylogenetic relationships among palpimanoid spiders. We use the recently designed Ultra-Conservative-Elements (UCE) Arachnid bait-set (Faircloth, 2017; Starrett et al., 2017) and an exon-based bait-set that was designed from spider transcriptomes. The UCE method makes use of short, highly conserved DNA sequences that span major lineages (Siepel et al., 2005; Faircloth et al., 2012). Hundreds of independent UCE loci can be extracted from diverse taxa, with the regions flanking the UCE cores providing signal for phylogenetic inference at multiple taxonomic scales. Sequence variation in flanking regions increases with distance from the UCE core (Faircloth et al., 2012). This increasing variation makes UCE useful for reconstructing phylogenetic relationships across a variety of timescales, from recently diverged populations to distantly related groups (Faircloth et al., 2012; McCormack et al., 2012; Smith et al., 2013). An exon-based bait-set was designed for this study to complement the UCE data by providing additional markers for protein-coding regions within the genome that do not overlap with the UCE data. We used these two bait-sets to gather molecular sequences of lineages that spanned the modern spiders (Araneomorphae), but that focused on the palpimanoids. Fourteen museum specimens that were collected over the last two decades and stored in 75% ethanol were included in this study. We also perform an ancestral character state reconstruction to examine evolution of the tubular carapace in palpimanoids.

2. Methods

2.1. Taxa selection and DNA extraction

To examine relationships among palpimanoids we included 34 terminals representing the five palpimanoid families (number of genera included/total number of genera): Archaeidae (5/5), Huttoniidae (1/1), Mecysmaucheniidae (5/7), Palpimanidae (11/18), Stenochilidae (1/2). We included 14 additional non-palpimanoid Araneomorphae taxa representing 13 families that represent major clades within the Araneomorphae: Hypochilidae, Filistatidae, Pholcidae, Scytodidae, Segestriidae, Austrochilidae, Eresidae, Araneidae, Theridiosomatidae, Uloboridae, Agelenidae, Lycosidae, and Salticidae. The tree was rooted with the clade containing Hypochilidae (*Hypochilus*), Filistatidae (*Kukulcania*), and Synspermiata taxa (*Physocyclus*, *Scytodes*, *Segestria*), the earliest-diverging araneomorphs (Wheeler et al., 2016). Some African Palpimanidae genera were difficult to diagnose based on the literature, which contains few illustrations, and the following identifications should be treated with caution: *Boagrius* was diagnosed based on eye pattern (Simon, 1893), however the specimen is a male and there are no illustrations of males in the literature; specimens listed as Chediminae may represent a new genus.

Thirty (out of 48) specimens had DNA extractions available from previous studies (Wood et al., 2012, 2015, 2016), including six extractions that came from the National Museum of Natural History Biorepository at the Smithsonian Institution. For the remaining specimens, total genomic DNA was isolated using the DNeasy™ Tissue Kit (Qiagen Inc., Valencia, CA). For each specimen, 1–4 legs were ground up in a lysis buffer with a grinding implement. The homogenate was incubated at 55 °C for 6–8 h and then purified following the manufacturer's protocols. See Table 1 for a complete list of all specimens used in the study. Extracted DNA was quantified using high sensitivity Qubit fluorometry (Life Technologies, Inc.) and quality was checked using gel electrophoresis on a 1.5% agarose gel.

2.2. Exon-based bait design

Unprocessed raw transcriptome reads were obtained from the SRA database (http://www.ncbi.nlm.nih.gov/sra) for five taxa: *Calileptoneta californica* (SRR3144085), *Hypochilus pococki* (SRR1514889), *Uloborus glomosus* (SRR1328334, SRR3144086), *Deinopis longipes* (SRR1514879), and *Loxosceles deserta* (SRR3144077). FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) visualized quality of reads. Initial removal of low-quality reads and TruSeq multiplex index adaptor sequences (Illumina) was performed with Trim Galore! v.0.4.0 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore), setting the

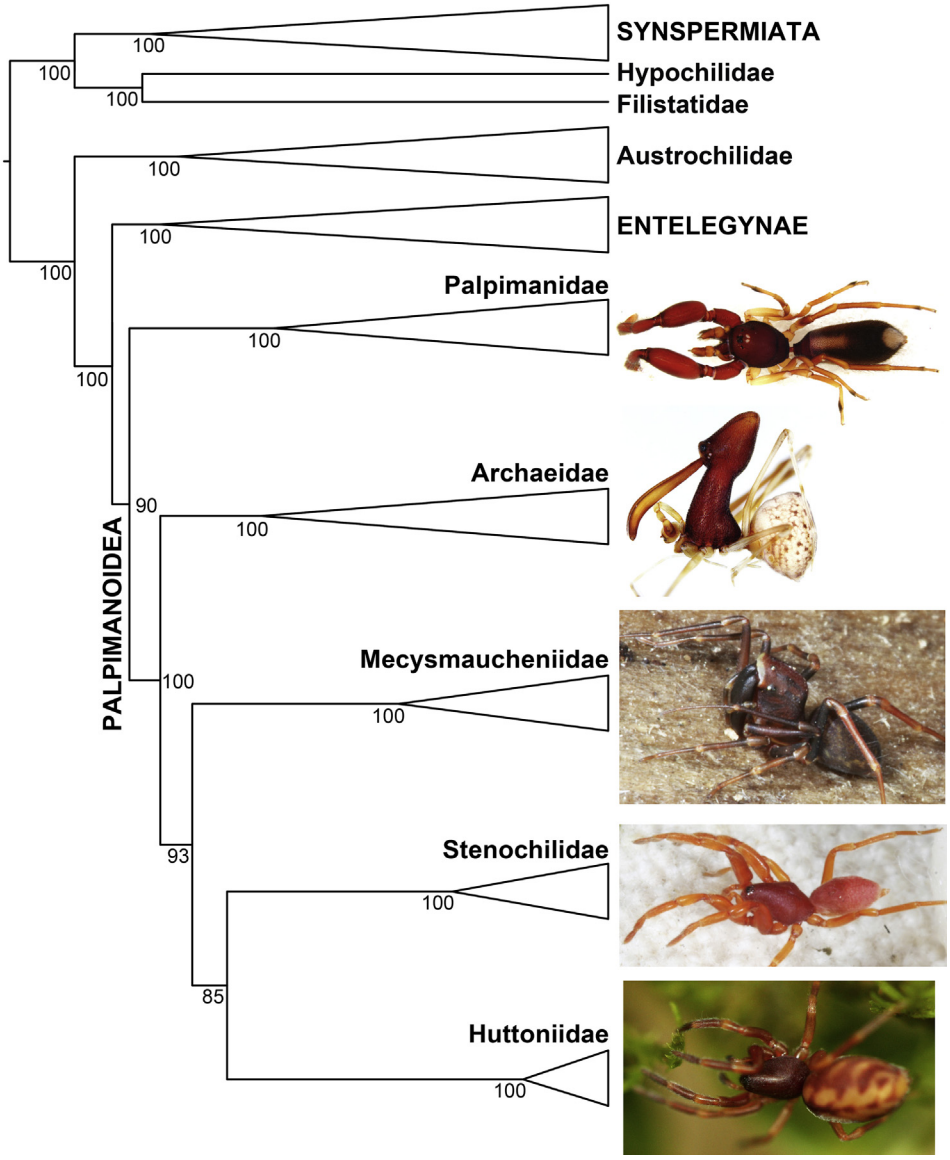

Fig. 1. Summary phylogeny from Maximum Likelihood analysis of combined (UCE + exon) partitioned data at 70% taxon coverage.

quality threshold to minimum Phred score of 30. Illumina TruSeq multiplex adaptor sequences were trimmed, specific to the adaptor used in sequencing with the paired-end data flagged. A second round of quality threshold filtering (minimum Phred score of 35) as well as removal of rRNA sequence contamination was conducted in Agalma v. 0.5.0 using the 'transcriptome' pipeline (Dunn et al., 2013). Processed reads were assembled with the Trinity de novo Assembler (Grabherr et al., 2011b) (release 13-07-2011) with 100 GB of memory and a path reinforcement distance of 50 (Haas et al., 2013). Contigs were mapped against the Swissprot database (Bairoch and Apweiler, 2000) using the blastx program of the BLAST suite (Altschul et al., 1990) and the number of contigs returning blast hits was quantified. All nucleotide sequences were translated with Transdecoder using default parameters (Grabherr et al., 2011a). Subsequent peptide translations were filtered for redundancy and uniqueness using CD-Hit v.4.6.3 under default parameters, and a 95% similarity threshold (Fu et al., 2012). Genome data from *Parasteatoda tepidariorum* (PRJNA316108) and *Stegodyphus mimosarum* (PRJNA222714) were incorporated using predicted peptide sequences obtained from NCBI protein database (http://www.ncbi.nlm.nih.gov/genome/).

Orthology assessment was conducted on 100 CPUs of the High-Performance Computing Cluster (Smithsonian Institution, confluence.si.edu/display/HPC/), using OMA standalone v.1.05 (Roth et al., 2008; Altenhoff et al., 2013), with the parameters set to default, except with a minimum alignment score of 200, a length tolerance ratio of 0.75 and a minimum sequence length of 100. A total of 7866 informative putative orthogroups (more than four taxa) were obtained; orthogroups and genes are referred to interchangeably. Of these, 4809 putative orthogroups were shared by all taxa. In order to include all putative loci that spanned the greatest phylogenetic distance within spiders, we proceeded with orthogroups that contained sequences from both *Parasteatoda tepidariorum* and *Hypochilus pococki* (5753 orthogroups). In order to ensure only single copy orthologs were used for downstream probe design, orthogroups were mapped against a curated "custom" spider core ortholog reference protein set for *Acanthoscurria geniculate* from Garrison et al. (2015) using the blastx program of the BLAST suite. Only orthogroups with unique hits to this database were kept, resulting in 4926 putative loci. Nucleotide alignments were generated for all orthogroups in MAFFT (Katoh et al., 2002; Katoh and Toh, 2008). Loci were then filtered based on global pairwise distance, selecting loci with 70% or higher percent pairwise identity value in Geneious v10.2.3 (Biomatters Ltd.), resulting in 1050 putative loci. Each putative locus

Table 1
Voucher information and sequencing statistics. USNMENT = Smithsonian Institution, National Museum of Natural History, Entomology Department; CASENT = California Academy of Sciences, Entomology Department; BR = National Museum of Natural History Biorepository. Specimens in bold were stored in 70–80% ethanol at room temperature; * the year the specimen was collected.

Voucher number	Extraction number	Family	Species	Country	DNA input (ng)	Data Set	Raw read pairs	Contigs	Sequence Coverage	n50_size	Loci	Accession Number
USNMENT879960	HW_0164	Hypochilidae	*Hypochilus pococki*	United States	100.0	UCE	36,95,002	50,655	30.6	239	440	SRR7363166
						Exon	34,48,556	36,201	81.8	319	117	SRR7387525
CASENT9034219	HW_0079	Filistatidae	*Kukulcania hibernalis*	United States	100.0	UCE	12,85,121	13,822	4.9	273	127	SRR7363165
						Exon	9,90,720	9,270	7.4	264	212	SRR7387526
USNMENT01147069	BR: AH4WZ39	Pholcidae	*Physocyclus globosus*	Costa Rica	100.0	UCE	25,93,684	66,049	20.3	252	493	SRR7363168
						Exon	33,06,266	48,718	97.0	294	287	SRR7387523
USNMENT00786521	BR: AH2UN92	Scytodidae	*Scytodes thoracica*	Slovenia	8.1	UCE	7,96,724	11,772	6.8	261	254	SRR7363167
						Exon	10,21,248	14,959	17.5	286	275	SRR7387524
CASENT9024001	HW_Sege43	Segestriidae	*Segestria* sp.	Madagascar	100.0	UCE	31,21,408	39,843	5.9	264	283	SRR7363162
						Exon	12,71,277	22,939	3.4	291	192	SRR7387521
CASENT9027630	HW_Achil44	Austrochilidae	Austrochilidae sp.	Chile	100.0	UCE	25,95,223	26,525	24.8	235	518	SRR7363161
						Exon	25,55,924	30,226	31.1	269	293	SRR7387522
CASENT9020543	HW_0197	Austrochilidae	***Hickmania troglodytes* *2009**	Tasmania	100.0	UCE	16,68,681	20,484	16.6	320	313	SRR7363164
						Exon	20,97,255	24,569	29.7	322	253	SRR7387499
USNMENT00879068	HW_0162	Uloboridae	*Philoponella tingens*	Costa Rica	24.1	UCE	8,64,707	19,045	4.8	224	338	SRR7363163
						Exon	6,40,552	12,445	6.7	294	247	SRR7387520
USNMENT00786563	BR: AH2UN75	Agelenidae	*Coelotes terrestris*	Slovenia	100.0	UCE	13,64,879	10,717	7.4	296	149	SRR7363170
						Exon	8,26,487	7,862	12.0	292	248	SRR7387517
USNMENT00786358	BR: AH2UP76	Lycosidae	*Pardosa blanda*	Switzerland	78.4	UCE	20,20,330	22,724	10.5	284	355	SRR7363169
						Exon	13,41,735	19,078	40.9	314	254	SRR7387518
USNMENT00786348	BR: AH2UP77	Salticidae	*Hasarius adansoni*	Slovenia	67.2	UCE	16,89,285	17,645	11.7	225	338	SRR7363158
						Exon	11,19,599	15,185	33.5	271	252	SRR7387507
CASENT9028232	HW_Steg42	Eresidae	*Stegodyphus* sp.	South Africa	100.0	UCE	14,61,517	14,357	19.3	295	682	SRR7363157
						Exon	15,27,340	19,108	27.3	365	220	SRR7387508
USNMENT00786593	BR: AH2UN31	Araneidae	*Araneus diadematus*	Slovenia	100.0	UCE	19,89,506	18,078	19.1	280	281	SRR7363156
						Exon	14,11,084	15,154	46.6	272	247	SRR7387505
USNMENT00879029	HW_0163	Theridiosomatidae	*Wendilgarda clara*	Costa Rica	25.8	UCE	7,87,579	12,863	3.2	236	242	SRR7363155
						Exon	5,90,647	7,777	3.6	289	207	SRR7387506
CASENT9028364	HW_0080	Archaeidae	*Zephyrarchaea mainae*	Australia	48.4	UCE	7,91,706	12,711	4.1	273	160	SRR7363154
						Exon	12,70,705	11,275	12.1	268	272	SRR7387503
CASENT9028388	HW_0074	Archaeidae	*Austrarchaea nodosa*	Australia	37.1	UCE	10,23,868	21,145	7.8	267	300	SRR7363153
						Exon	14,29,053	16,175	21.4	290	294	SRR7387504
CASENT9018992	HW_0014	Archaeidae	*Madagascarchaea legendrei*	Madagascar	25.9	UCE	9,54,739	20,937	16.6	238	420	SRR7363152
						Exon	29,85,556	18,731	69.3	298	322	SRR7387501
CASENT9019010	HW_0026	Archaeidae	*Madagascarchaea jeanneli*	Madagascar	17.0	UCE	6,37,608	15,119	6.7	227	346	SRR7363151
						Exon	14,61,485	12,914	22.0	293	288	SRR7387502
CASENT9028446	HW_0110	Archaeidae	*Afrarchaea woodae*	South Africa	11.7	UCE	12,82,939	21,826	4.9	247	305	SRR7363160
						Exon	10,43,205	7,873	13.0	338	319	SRR7387515
CASENT9028297	HW_0092	Archaeidae	*Eriauchenius maharaensis*	Madagascar	23.0	UCE	50,28,554	1,13,528	11.0	258	386	SRR7363159
						Exon	33,95,269	26,688	44.2	259	280	SRR7387516
CASENT9028340	HW_0129	Archaeidae	*Eriauchenius andrianampoinimerina*	Madagascar	53.7	UCE	11,49,043	17,612	9.4	243	384	SRR7363185
						Exon	15,10,353	27,271	13.6	275	306	SRR7387498
CASENT9028231	HW_Hutt41	Huttoniidae	*Huttonia* sp.	New Zealand	28.6	UCE	25,58,763	35,578	29.7	278	432	SRR7363186
						Exon	22,59,241	36,417	33.5	317	301	SRR7387497
CASENT9034307	HW_0126	Huttoniidae	*Huttonia* sp.	New Zealand	39.6	UCE	7,62,537	11,683	13.8	260	375	SRR7363183
						Exon	6,25,777	11,917	15.3	305	308	SRR7387496
CASENT9028424	HW_0081	Stenochilidae	*Colopea* sp.	Brunei	50.0	UCE	11,70,260	16,243	6.8	291	241	SRR7363184
						Exon	12,51,520	8,809	15.0	326	279	SRR7387495
CASENT9034305	HW_0136	Stenochilidae	*Colopea* sp.	Philippines	50.0	UCE	32,66,266	29,112	24.2	262	457	SRR7363181
						Exon	28,92,058	33,911	22.7	279	311	SRR7387494
CASENT9028246	HW_0050	Mecysmaucheniidae	*Aotearoa magna*	New Zealand	100.0	UCE	35,99,379	45,995	15.3	252	270	SRR7363182

(*continued on next page*)

Table 1 (*continued*)

Voucher number	Extraction number	Family	Species	Country	DNA input (ng)	Data Set	Raw read pairs	Contigs	Sequence Coverage	n50_size	Loci	Accession Number
						Exon	24,76,535	26,533	26.1	306	238	SRR7387493
CASENT9028244	HW_0052	Mecysmaucheniidae	*Zearchaea sp.*	New Zealand	10.8	UCE	7,26,059	13,818	5.7	251	269	SRR7363179
						Exon	10,50,990	12,228	10.9	295	279	SRR7387492
CASENT9027870	HW_0040	Mecysmaucheniidae	*Mesarchaea bellavista*	Chile	17.4	UCE	15,55,603	21,742	10.4	237	317	SRR7363180
						Exon	15,93,182	15,628	16.7	276	279	SRR7387491
no voucher, whole body used	HW_0028	Mecysmaucheniidae	*Chilarchaea quellon*	Chile	100.0	UCE	28,73,548	43,220	11.5	291	352	SRR7363177
						Exon	15,38,639	17,859	17.2	314	263	SRR7387490
CASENT9034609	HW_0114	Mecysmaucheniidae	*Mecysmauchenius segmentatus*	Chile	50.0	UCE	5,92,199	8,475	3.4	338	109	SRR7363178
						Exon	7,93,387	6,028	8.6	308	280	SRR7387489
CASENT9053300	HW_0148	Palpimanidae	*Anisaedus* sp.	Chile	50.0	UCE	8,85,662	16,280	6.0	251	344	SRR7363148
						Exon	6,61,069	17,003	6.4	290	276	SRR7387509
CASENT9070025	HW_0189	Palpimanidae	***Otiothops* sp. *2005**	French Guiana	37.1	UCE	12,95,122	22,073	4.8	279	194	SRR7363147
						Exon	13,83,918	21,175	5.3	315	227	SRR7387510
CASENT9024279	HW_0082	Palpimanidae	*Palpimanus* sp.	South Africa	100.0	UCE	74,87,203	42,637	9.3	322	300	SRR7363150
						Exon	21,69,825	13,923	19.7	296	246	SRR7387511
CASENT9065070	HW_0180	Palpimanidae	***Hybosida* sp. *2010**	Seychelles	7.0	UCE	43,92,763	1,06,351	12.1	235	529	SRR7363149
						Exon	29,71,509	56,377	14.0	243	282	SRR7387512
CASENT9065049	HW_0181	Palpimanidae	***Chedimanops* sp. *2012**	Uganda	34.4	UCE	16,42,312	29,057	6.1	269	309	SRR7363144
						Exon	7,38,942	12,941	5.7	301	227	SRR7387513
CASENT9064935	HW_0191	Palpimanidae	***Chedimanops* sp. *2012**	Uganda	50.0	UCE	11,08,168	20,536	3.7	273	176	SRR7363143
						Exon	10,48,072	17,121	4.6	312	228	SRR7387514
CASENT9065134	HW_0193	Palpimanidae	***Diaphorocellus* sp. *2012**	Uganda	36.8	UCE	51,45,899	86,074	13.6	257	386	SRR7363146
						Exon	39,39,657	55,806	14.0	275	287	SRR7387519
CASENT9030406	HW_0084	Palpimanidae	*Diaphorocellus* sp.	Madagascar	50.0	UCE	12,24,051	14,768	4.7	279	229	SRR7363145
						Exon	13,60,895	9,897	12.8	306	254	SRR7387500
CASENT9029145	HW_0086	Palpimanidae	*Diaphorocellus* sp.	Madagascar	37.4	UCE	32,91,189	47,508	9.1	279	335	SRR7363142
						Exon	38,47,608	21,216	33.1	287	262	SRR7387488
CASENT9070036	HW_0194	Palpimanidae	***Sarascelis* sp. *2000**	Gabon	16.4	UCE	34,68,776	59,631	10.2	249	377	SRR7363141
						Exon	24,43,617	35,629	10.1	275	287	SRR7387527
CASENT9070035	HW_0176	Palpimanidae	***Sarascelis* sp. *2000**	Gabon	6.7	UCE	15,20,119	16,962	7.6	253	190	SRR7363175
						Exon	11,58,272	12,079	8.3	297	176	SRR7387529
CASENT9070040	HW_0196	Palpimanidae	**Chediminae sp. *2000**	Cameroon	21.3	UCE	12,65,189	27,736	4.5	255	248	SRR7363176
						Exon	9,87,383	18,067	5.0	286	223	SRR7387528
CASENT9070034	HW_0190	Palpimanidae	**Chediminae sp. *2000**	Cameroon	24.9	UCE	8,06,009	17,028	3.4	243	201	SRR7363139
						Exon	7,10,898	13,603	4.1	270	200	SRR7387531
CASENT9064933	HW_0188	Palpimanidae	***Boagrius* sp. *2012**	Uganda	50.0	UCE	14,77,490	26,285	4.0	241	196	SRR7363140
						Exon	10,07,113	15,249	4.0	262	195	SRR7387530
CASENT9011042	HW_0178	Palpimanidae	***Steriphopus* sp. *2003**	Myanmar	16.1	UCE	13,91,938	25,592	5.6	258	259	SRR7363171
						Exon	9,40,633	16,676	5.1	295	219	SRR7387533
CASENT9070038	HW_0179	Palpimanidae	***Scelidocteus* sp. *2000**	Cameroon	29.6	UCE	57,65,322	85,189	11.1	287	409	SRR7363172
						Exon	38,96,044	57,335	9.6	352	270	SRR7387532
CASENT9065043	HW_0192	Palpimanidae	***Scelidocteus* sp. *2012**	Uganda	18.6	UCE	20,33,022	42,258	5.2	256	276	SRR7363173
						Exon	19,14,944	34,595	5.2	283	269	SRR7387535
CASENT9039578	HW_0198	Palpimanidae	***Fernandezina* sp. *2007**	Peru	< 0.05	UCE	1,55,551	4,788	0.6	270	7	SRR7363174
						Exon	1,68,110	5,317	5.2	276	28	SRR7387534

and associated probes were then mapped to the *Parasteatoda tepidariorum* genome (Schwager et al., 2017) to test for fidelity. A custom mybaits final probe set was generated by Arbor Biosciences that contained 58,585 baits that were 120 base pairs long with around 1.3× tiling density, resulting in an on average 22 base pair overlap, and targeted 1050 putative loci.

2.3. Library preparation, enrichment, and sequencing

Libraries were prepared and enriched following protocols in Faircloth et al. (2015), but following modifications in Blaimer et al. (2015). Depending on prior degradation and quality of the DNA, between 5 and 100 ng of DNA were sheared between 0 and 60 s (amp = 25%, pulse = 10–10, to a target size of approximately 250–600 bp) by sonication (Q800R, Qsonica Inc.). Sheared DNA was dried completely and rehydrated to the required input volume (13 μL) and used as input for DNA library preparation (Kapa Hyper Prep Library kit, Kapa Biosystems). Library preparation was done following the protocols of Faircloth et al. (2015) in the absence of SPRI beads. After ligation of universal stubs (Faircloth and Glenn, 2012), a 0.7× SPRI bead clean was done (Kapa Pure Beads, Kapa Biosystems) on a Wafergen Apollo liquid handler (Wafergen Biosystems), resulting in 30 μL of post-ligation library. For adapter ligation, we used TruSeq-style adapters (Faircloth and Glenn, 2012). PCR conditions were as follows: 15 μL post ligation library, 25 μL HiFi HotStart polymerase (Kapa Biosystems), 2.5 μL each of Illumina TruSeq-style i5 and i7 primers, and 5 μL double-distilled water (ddH_2O). We used the following thermal protocol (Kapa Biosystems): 98 °C for 45 s; 13 cycles of 98 °C for 15 s, 65 °C for 30 s, 72 °C for 60 s, and final extension at 72 °C for 5 m. We generated two PCR amplified libraries from each post-ligation library, with each PCR containing a different i5-i7 adapter combination. This allowed one DNA sample to be equally divided into two separate enrichment reactions (i.e., UCE & exon).

PCR cleanup was done with a 1.0X SPRI bead clean (Kapa Pure Beads, Kapa Biosystems) on a Wafergen Apollo (Wafergen Biosystems) with a final library volume of 20μL. Following clean-up, libraries were divided into enrichment pools containing 8 libraries combined at equimolar ratios with final concentrations of 137–184 ng/μL. The duplicated PCR products formed two different groups of pools, one for UCE enrichment and one for exon enrichment.

We enriched one set of pools for UCE using the Arachnid-specific probes-set (Arachnida 1.1Kv1) targeting 1120 UCE loci (Faircloth, 2017; Starrett et al., 2017). The duplicate set of pools with different indices was enriched for the custom designed exon probes targeting 1050 putative loci (as described above). For both UCE and exon enrichments, we followed version 3.02 of the mybaits protocol (Arbor Biosciences). Hybridization reactions were incubated for 24 h at 65 °C, subsequently all pools were bound to streptavidin beads (MyOne C1; Life Technologies), and washed according to the MYBait protocol (Arbor Biosciences). We combined 15 μL of streptavidin bead-bound, washed, enriched library with 25 μL HiFi HotStart Taq (Kapa Biosystems), 5 μL of Illumina TruSeq primer mix (5 μM forward and reverse primers) and 5 μL of ddH^2O. Post-enrichment PCR used the following thermal profile: 98 °C for 45 s; 18 cycles of 98 °C for 15 s, 60 °C for 30 s, 72 °C for 60 s; and a final extension of 72 °C for 5 m. We purified resulting reactions using 1.0X bead clean using Kapa Pure Beads (Kapa Biosystems), and resuspended the enriched pools in 22 μL.

We then quantified pools using qPCR library quantification (Kapa Biosystems) with two serial dilutions of each pool (1:200,000, 1:1,000,000), assuming an average library fragment length of 600 bp. Based on the size-adjusted concentrations estimated by qPCR, we combined all pools at an equimolar concentration of 30 nM, and size-selected for 250–600 bp with a BluePippin (SageScience). We sequenced the pooled libraries in a single lane of a paired-end run on an Illumina HiSeq (2x150bp rapid run) at the University of Utah Huntsman Cancer Institute.

2.4. Data processing

Data processing steps were conducted with the PHYLUCE pipeline (Faircloth, 2016). The UCE and exon datasets were analyzed separately. We removed low-quality bases and adapter sequence from the demultiplex raw read FASTQ data using Illumiprocessor (Faircloth, 2013; Bolger et al., 2014). Sample specific sequences were assembled into contigs using Trinity (Grabherr et al., 2011b; version r20140717). Contigs from all samples were mapped to FASTA files containing enrichment probes for either UCE or exon using the LASTZ (Harris, 2007) wrapper script "phyluce_assembly_match_contigs_to_probes" with - minimum coverage and minimum identity values of 65 for the UCE dataset (Starrett et al., 2017), and 80 for the exon dataset. Loci obtained from the preceding script step were aligned using MAFFT (min-length = 20, no-trim) (Katoh and Standley, 2013) and trimmed with GBLOCKS (b1 = 0.5, b2 = 0.5, b3 = 12, b4 = 7; Castresana, 2000; Talavera and Castresana, 2007). We selected 25%, 50% and 70% taxa per locus completeness subsets for the UCE and exon alignments. Data were concatenated into UCE-only, exon-only, and combined UCE + exon matrices for subsequent phylogenetic analysis. The best-fit substitution model and data partitions were estimated using Partition Finder v.2.1.1 (Lanfear et al., 2014, 2016) after implementing the Sliding-Window Site Characteristics method (Tagliacollo and Lanfear, 2018).

2.5. Phylogenetic analysis

Phylogenetic analyses were conducted on three matrices for each concatenated data set (UCE-only, exon-only, and combined – UCE + exon), which consisted of taxon coverages of 70%, 50%, and 25% for each locus. Phylogenetic analysis was performed on the partitioned and non-partitioned data matrices using raxmlHPC-MPI-SSE3 v8.2 (Stamatakis, 2014) implementing the rapid bootstrap algorithm (Stamatakis et al., 2008) plus Maximum Likelihood tree search (100 bootstrap replicates and the GTRGAMMA model). Coalescent analyses were also performed on all data sets at different taxon coverages using ASTRAL-III (Zhang et al., 2017) on individual unrooted gene trees estimated by RAxML and 500 bootstrap replicates. A Bayesian analysis was performed on the partitioned data sets at the different taxon coverages using MrBayes v.3.2 (Ronquist et al., 2012): Analyses were run for 10–50 million generations, depending on how quickly the analysis converged, with sampling every 1000th generation; simultaneous analyses were evaluated for convergence and burn-in was assessed using Tracer v.1.7.0 (Rambaut et al., 2018), resulting in a final consensus tree with node support expressed as posterior probabilities. All analyses were conducted on the Smithsonian Institution High Performance Cluster (SI/HPC).

To examine evolution of the highly modified, tubular carapace present in Archaeidae and Mecysmaucheniidae we performed a likelihood ancestral trait reconstruction in Mesquite (Maddison and Maddison, 2010) on the phylogeny that resulted from the RAxML analysis of the partitioned, combined dataset at 70% gene occupancy. Branch lengths were smoothed using penalized likelihood (Sanderson, 2002) in the R package 'ape' (Paradis et al., 2004) using the 'cronopl' function (lambda = 1). Carapace shape was treated as a discrete trait and modeled using the Markov k-state 1 parameter model (Lewis, 2001). Archaeids and mecysmaucheniids were scored as "present" for having a tubular carapace and all other taxa were scored as "absent." The phylogeny was pruned so that only one representative per family remained.

3. Results

The majority of DNA extractions were from specimens that had been collected directly into 95% ethanol, and upon completion of the field expedition, were stored at below freezing. Fourteen extractions were

from museum specimens in 70–75% ethanol that were stored at room temperature (5 from the year 2000, 1 from 2003, 1 from 2005, 1 from 2010, and 5 from 2012). An additional seven museum specimens in 70–75% ethanol were extracted but were not sequenced for the phylogenetic analysis. Of these extractions of museum specimens, four failed, meaning there was too little DNA (less than 0.05 ng/μL) to be detected by the Qubit Fluorometer. See Supplementary Fig. S1 to compare total DNA extracted by year, coded for specimen size, for the museum specimens. For sequencing, 48 DNA extractions were enriched using both the exon baits and the UCE baits, for a total of 96 samples sequenced in the Illumina Hi-Seq lane. Sequences were produced for all samples except one: *Fernandezina* sp. (Palpimanidae), a museum specimen collected in the year 2007; the DNA concentration of this extraction was not detectable by the Qubit Fluorometer (< 0.05 ng/μL). The specimen was included in the library preparation, enrichment and sequencing, but was excluded from the phylogenetic analysis due to a very low number of sequences.

Table 1 provides the summary sequencing statistics for each of the 48 taxa sequenced. We obtained an average of 2,047,135 reads and an average coverage of 10.2 per sample for the UCE-based data and 1,689,045 and an average coverage of 20.3 for the exon-based data. The reads were assembled into an average of 30,918 contigs per sample for the UCE data and 20,995 contigs per sample for the exon-based data. For one specimen, *Fernandezina*, which did not have detectable levels of DNA in the extraction, we only recovered 7 and 28 UCE and exon target loci, respectively. Excluding *Fernandezina*, the number of loci recovered ranged from 109 to 682 out of 1120 UCE targets, and 117 to 322 of 1050 exon targets. Average capture efficiencies were 28% (317/1120) for UCE and 24% (256/1050) for exon-based loci. Many of the exon-based loci that were recovered during data processing were discarded after the validation step in PHYLUCE where multiple contigs that map to a single locus are removed. We suspect multiple contigs per representative loci were recovered but not assembled into one continuous contig due to large intron lengths between coding regions, which are present in some arachnid genomes (Sanggaard et al., 2014; Babb et al., 2017; Schwager et al., 2017). As a result, the average capture efficiency of the exon-based targets was lower than that of the UCE targets (see Supplementary Info for confirmation of putative loci within the target taxa based on Palpimanidae transcriptomes). See Table 2 for summary statistics of the final concatenated data sets at 50% taxon coverage. Quality-trimmed sequence reads are available from the NCBI Sequence Read Archive (BioProject PRJNA475748, SRP150838).

There were only minor differences between the RAxML, Bayesian, and ASTRAL phylogenetic analyses (Figs. 1–3, S2-3). Furthermore, the analysis of the UCE-only datasets and the exon-only datasets resulted in phylogenies that are mostly congruent with the combined datasets. However the exon-only datasets sometimes failed to recover clades that were well supported in the UCE-only and combined analyses, including Synspermiata and Araneoidea (Figs. 2, 3 and S2-3). In the following analyses we recovered a well-supported monophyletic Palpimanoidea (that fell sister to the Entelegynae): the Bayesian, RAxML partitioned, and RAxML unpartitioned analyses of the combined and the UCE-only datasets at 70% taxon coverage. In all ASTRAL analyses Palpimanoidea was monophyletic, but with low support (bootstrap < 80%). Palpimanoidea was also monophyletic for the RAxML and Bayesian analyses of the exon-only data set at 50% taxon coverage. In all other analyses Palpimanoidea was paraphyletic, with Palpimanidae falling as sister to the Entelegynae. All five palpimanoid families were monophyletic in all analyses with strong support. Palpimanidae was consistently supported as the earliest-diverging lineage. In the majority of analyses the remaining palpimanoids (Archaeidae, Huttoniidae, Mecysmaucheniidae, Stenochilidae) formed a well-supported monophyletic group. Archaeidae was the next lineage to diverge, leaving Huttoniidae, Stenochilidae and Mecysmaucheniidae forming a well-supported clade in most analyses. The majority of analyses recovered a strongly supported South American clade (*Otiothops* + *Anisaedus*) that is sister to an Old-World clade that includes specimens collected from parts of Africa, Madagascar, the Seychelles, and Myanmar. In the majority of analyses: *Palpimanus* is the first Old-World genus to diverge; *Diaphorocellus*, *Hybosida*, and *Chedimanops* form a monophyletic group; *Steriphopus* + *Scelidocteus* is sister to a clade that is difficult to diagnose, but may be *Boagrius* and a new genus. The aligned data matrices and resulting phylogeny (.tre) files are deposited in **the Dryad Digital Repository**: https://doi.org/10.5061/dryad.387c58v.

The ancestral character state reconstruction revealed that the tubular carapace in archaeids and mecysmaucheniids evolved two times independently (Supp. Fig. S4).

Table 2
Sequence statistics for 47 taxa at 50% gene occupancy.

Data set	Total loci	Average length of loci	Total alignment	Informative sites
UCE	211	187.44	39,549	14,086
Exon	35	296.54	10,379	3,933
Combined	246	202.96	49,928	18,019

4. Discussion

Palpimanoidea classification. Based on a total evidence analysis of morphological data and four molecular markers Palpimanoidea was restricted to the following five families (Wood et al., 2012): Archaeidae, Huttoniidae, Mecysmaucheniidae, Palpimanidae, and Stenochilidae. This study also placed Palpimanoidea as sister to the Entelegynae and recovered the well-supported relationships of (Archaeidae + Stenochilidae, both long branches) and (Palpimanidae + Huttoniidae) (other relationships among palpimanoid families were not well supported). Later, Wheeler et al (2016), using data from six molecular markers to build a phylogeny that was constrained by transcriptomic data, and with thorough taxon sampling across spiders, recovered Palpimanoidea as paraphyletic, although with low branch support. Recently, Fernández et al (2018) used transcriptomic data to examine relationships across spiders, which included four out of five palpimanoid families, and recovered a well-supported, monophyletic Palpimanoidea that was sister to the Entelegynae.

The present study is the first to examine generic-level relationships among palpimanoids using Next-Generation Sequencing. Our study recovered a well-supported monophyletic Palpimanoidea in some analyses, while in other analyses Palpimanoidea was paraphyletic, with Palpimanidae falling as sister to the Entelegynes. However, there is strong morphological support for a monophyletic Palpimanoidea (Wood et al., 2012). Further, the Fernández et al. (2018) transcriptomic study had a much greater degree of taxon sampling throughout Araneae and recovered Palpimanoidea monophyly with strong support, although it included only 5 palpimanoids. Palpimanoidea is likely an ancient clade, with fossils going back to the Middle Jurassic (Selden et al., 2008), that likely diversified in the Triassic-Jurassic (Wood et al., 2013; Fernández et al., 2018). Palpimanoids tend to be spider specialists, with archaeids and palpimanids documented to prey on the Entelegynes (Cerveira and Jackson, 2005; Wood et al., 2012), their putative sister group that contains the bulk of spider diversity. Rapid, ancient diversification of palpimanoids as the Entelegynes concurrently diversified may explain why it is difficult to tease apart these relationships and why results at different gene occupancies conflict. We continue to accept the monophyly of Palpimanoidea until there is strong evidence otherwise.

The current study suggests that Palpimanidae, which has the most widespread distribution, occurring throughout South America, Africa, and the Mediterranean, and in parts of Asia, is likely the earliest-diverging palpimanoid clade, sister to the remaining families. Archaeidae, which is currently restricted to Australia, South Africa and Madagascar,

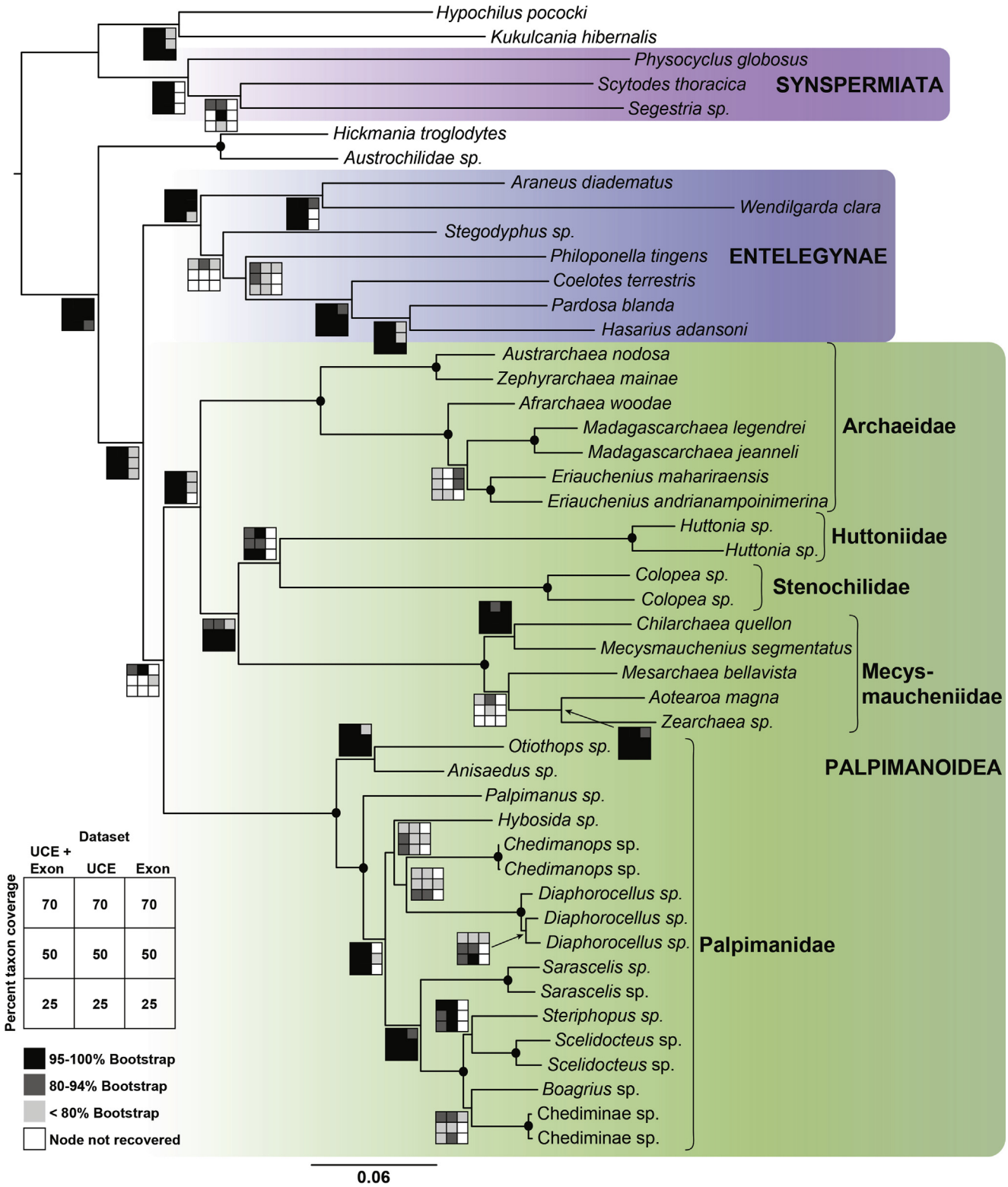

Fig. 2. Phylogeny from the Maximum Likelihood analysis of the combined (UCE + exon) and partitioned data at 70% taxon coverage. Square plots at nodes summarize branch support for the nine different analyses of the three different data sets (combined, UCE-only, exon-only) at three different amounts of taxon coverage (70%, 50%, 25%). Nodes with black circles were supported with bootstrap values > 95% in all analyses. The phylogeny was rooted with the clade containing Hypochilidae, Filistatidae, and Synspermiata.

although fossils occur in several areas in the Northern Hemisphere, is the next lineage that branches off. Mecysmaucheniidae, restricted to southern South America and New Zealand, forms a clade with Huttoniidae and Stenochilidae. Huttoniids are restricted to New Zealand and stenochilids occur from India, including Sri Lanka, through southeast Asia, an into northeastern Australia (R. Raven, *pers. comm.*). Relationships among Mecysmaucheniidae and Archaeidae genera were consistent with previous studies that used only four molecular markers

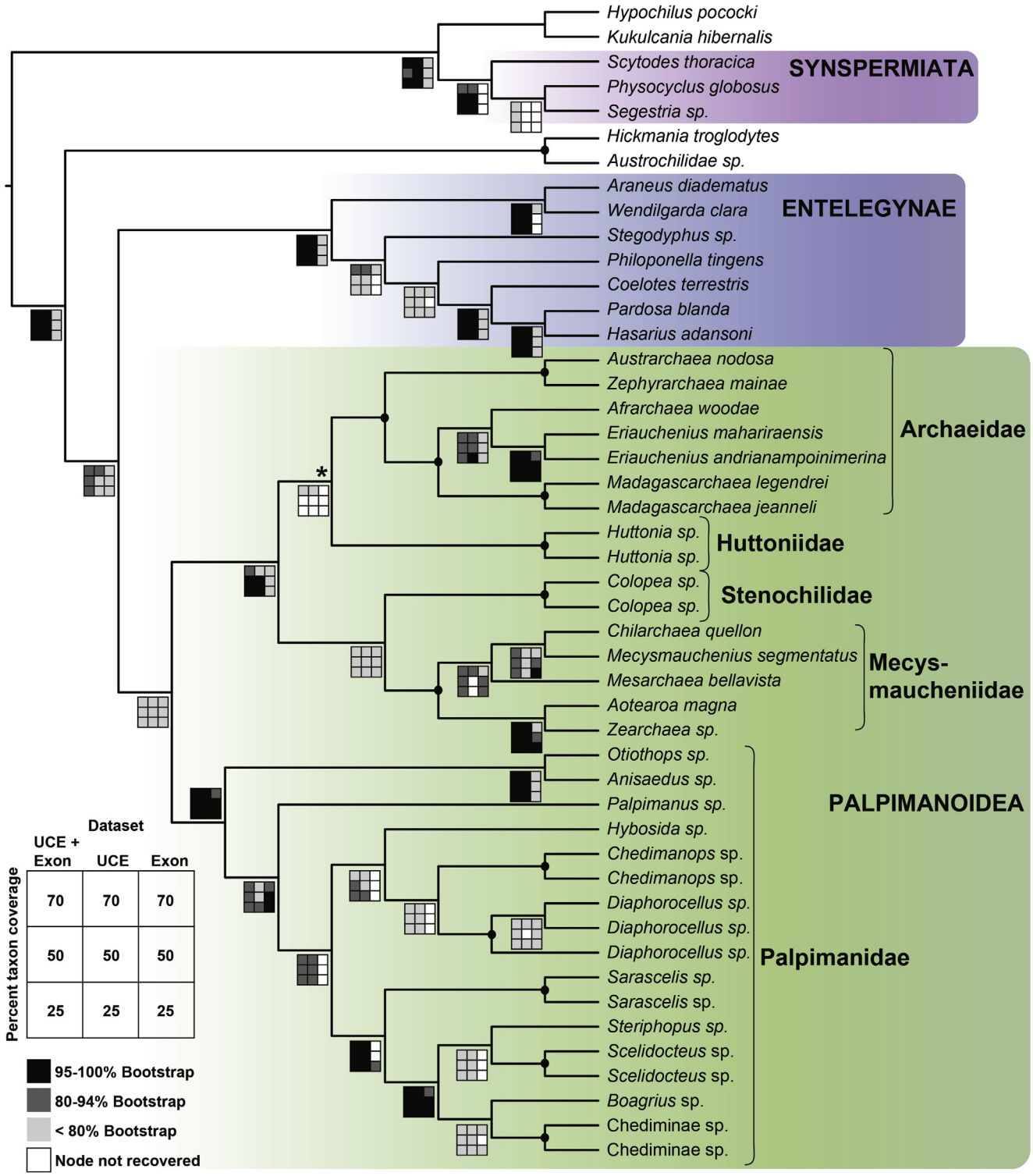

Fig. 3. ASTRAL species tree based on the Maximum Likelihood analysis of individual unrooted gene trees and 500 bootstrap replicates. The phylogeny was rooted with the clade containing Hypochilidae, Filistatidae, and Synspermiata. *In the analysis of the combined dataset and the UCE-only dataset, both at 70% taxon coverage, Archaeidae is sister to Huttoniidae; in all other ASTRAL analyses Archaeidae diverges first and Huttoniidae, Mecysmaucheniidae, and Stenochilidae form a monophyletic group with the following bootstrap support: combined dataset at 50% taxon coverage, bootstrap 77%; combined dataset at 25% taxon coverage, bootstrap 88%; UCE-only dataset at 50% taxon coverage, UCE-only dataset at 25% taxon coverage, exon-only dataset at 70% taxon coverage, exon-only dataset at 50% taxon coverage, exon-only dataset at 25% taxon coverage – all with bootstrap < 80%.

(Wood et al., 2012, 2015, 2016). However, target capture techniques were particularly useful for examining relationships among Palpimanidae genera, the current study being the first phylogenetic analysis produced for the family. Using these techniques, we were able to include museum specimens from many remote parts of the world that would be otherwise nearly impossible to collect without a great amount of time and resources. The African palpimanids need taxonomic revision: the genera are difficult to identify as the literature contains few

illustrations and a dichotomous key does not exist. Recent taxonomic work (Zonstein and Marusik, 2013, 2017; Zonstein et al., 2016) is helping to clarify the diagnosis of these genera, including documenting new genera, but is still only a first step. Field expeditions starting in the 1990s have documented many palpimanid specimens from Madagascar, a part of the world where palpimanids have not previously been documented. Some of these specimens were included in the current study (two *Diaphorocellus* specimens) and are currently being revised. Current and future work will build upon this study in order to elucidate world-wide diversification patterns of Palpimanidae.

Archaeids and mecysmaucheniids have evolved a similarly shaped tubular carapace, as well as elongated chelicerae ("jaws"). On the other hand, palpimanids, stenochilids, and huttoniids have more typical looking appearance, although, minor modifications to the carapace have evolved compared to non-palpimanoid spiders (Wood et al., 2012). The tubular carapace has shifted the orientation of the cheliceral muscles (Wood et al., 2012, 2016), supposedly allowing for the highly-maneuverable chelicerae observed in both archaeids and mecysmaucheniids. Both families can move their chelicerae 90° away from their body, atypical for spiders. Archaeids use this maneuverability to attack other spiders at a distance by swinging their long chelicerae out and holding the impaled prey away from their body until the prey is immobilized (Legendre, 1961; Forster and Platnick, 1984; Wood, 2008). Mecysmaucheniids execute a high-speed, ballistic, trap-jaw predatory strike that is the fastest arachnid movement documented to date (Wood et al., 2016). Our finding of independent evolution of the tubular carapace in each family is suggestive that palpimanoids may have evolved an innovation that predisposed them to carapace/chelicerae diversification. Current research is examining morphological modifications of internal carapace structures in palpimanoids.

Perspective on target capture methods. Next-generation sequencing has revolutionized arachnid systematics, allowing for increased resolution at very deep nodes, which was not previously possible using Sanger sequencing approaches (Bond et al., 2014; Fernández et al., 2014; Sharma et al., 2014; Garrison et al., 2015; Hamilton et al., 2016; Starrett et al., 2017; Fernández et al., 2018; Godwin et al., 2018; Hedin et al., 2018). This study further supports the utility of target capture methods in Araneomorphae spiders. Recent studies have used UCE baits to resolve deep relationships within Arachnids (Starrett et al., 2017), and to resolve family and generic relationships among mygalomorphs (tarantulas and their relatives) (Hedin et al., 2018). Here we corroborate the findings that UCE baits have phylogenetic utility, and we also include additional exon-based data for examining deep relationships within Araneomorphae with a specific focus on Palpimanoidea. Exon and UCE baits were both important for resolving relationships, although since there were more UCE loci than exon-based loci, UCE loci likely made a greater contribution to the results. Synspermiata (early-diverging araneomorphs with primitive genitalia) is a well-supported clade based on recent morphological and molecular studies (Michalik and Ramírez, 2014; Garrison et al., 2015; Fernández et al., 2018), and the monophyly of Araneoidea (orb-weavers and their relatives) is well established (Bond et al., 2014; Fernández et al., 2014; Hormiga and Griswold, 2014; Dimitrov et al., 2017). However, in the current study, only the UCE datasets successfully recovered these clades with strong support in all phylogenetic analyses, while the exon-only datasets failed to recover these clades in most analyses. Because there were fewer exon-based loci in our study than there were UCE loci it is difficult to draw more general conclusions about the usefulness of UCE markers compared to exonic markers in phylogenetic analysis.

We developed an exon-based probe set based on transcriptomes and our phylogenetic results indicate that exon-based targets span variation sufficient to resolve ancient and recent divergences in Palpimanoidea. Though an average of 310 loci from the 1050 targeted exon-based targets were recovered, the number of shared loci populating the matrices were 43, 35, and 11 for the 25%, 50% and 70% occupancy analyses, respectively. As previously discussed in Hamilton et al. (2016), we can expect a number of loci recovered to differ among samples due to variation in DNA preservation quality, library preparation, and sequencing depth. However, the number of loci recovered is also determined by the fidelity of the probes to the target clade of interest, which in turn, is affected by the long divergence times of the taxa used in probe design. Here, we constructed target exon loci and probe sequences from seven non-palpimanoid taxa, that likely diverged from each other over 200 million years ago (Ayoub et al., 2007; Dimitrov et al., 2012; Wood et al., 2013), where no representative from the in-group Palpimanoidea was present, as transcriptome sequence information was not available at the time. Although the taxa used in probe design bracketed the Palpimanoidea, with some lineages more basal and some lineages more derived, designing probes from non-in-group taxa may have resulted in a lower/no capture of palpimanoid specific sequence. Furthermore, for the different araneomorph taxa, the designed probes for the same loci ranged from 69 to 87% pairwise identity at the nucleotide level, which may have exceeded or been at the upper limit of physical capture for the probes (best performance ≤ 10%, though probes may be effective at up to 30% sequence divergence, Arbor Biosciences, *pers. comm*). Lastly, because we designed loci targets spanning the phylogenetic distance of Araneomorphae spiders, further relaxing stringency in analysis parameters may allow for an increased capture of non-orthologous sequences to be populated into the matrix (but see Supplementary Info for confirmation of putative loci within the target taxa based on Palpimanidae transcriptomes). As genomic or transcriptomic resources become available these exon-based probes or sequences should be redesigned or filtered to increase capture efficiency within the Palpimanoidea.

Our results show that target capture techniques successfully allow DNA from museum specimens to be used in molecular phylogenetics. Arachnid specimens in museum collections are typically stored in 70–80% ethanol and these concentrations at room temperature degrades DNA by oxidation and hydrolysis (Quicke et al., 1999; Vink et al., 2005). Spiders from museum collections that were collected back to the 1950s have been DNA barcoded, however, barcoding was typically more successful for larger-bodied specimens (> 3 mm) that were collected less than 20 years ago, and for smaller spiders that were only a few years old (Miller et al., 2013). In contrast, a different study showed the opposite, that PCR amplification of museum material works significantly better in spiders with smaller body sizes, and showed in addition, that the best predictor of success was the long term storage conditions of the museum collection, with age of the museum material being less important (Krehenwinkel and Pekar, 2015). While these studies offer encouragement that fluid-preserved specimens can be used for PCR amplification, high-throughput sequencing techniques actually capitalize on the fragmented DNA that characterizes museum specimens. Furthermore, high-throughput sequencing produces vastly more genomic data compared to traditional Sanger sequencing: using a novel DNA extraction method partial mitochondrial genomes have been assembled for spiders that were collected in the 1940s and preserved in ethanol (Cotoras et al., 2017); using target capture for UCE sequencing including several museum specimens collected over 30 years ago, systematic relationships among mygalomorph spiders have been revealed (Hedin et al., 2018). The museum specimens used in the current study were only collected as far back as the year 2000. However, results from these studies are encouraging and suggest that target capture is an exciting direction for sequencing rare taxa known only from museum collections. There is also great potential for generating taxon rich data sets with world-wide representation at a fraction of the effort and cost of re-collecting fresh material.

Acknowledgements

Funding for this study came from an Exploratory Award (year 2016) from the Global Genome Initiative at the Smithsonian Institution National Museum of Natural History. NS received support by the Danish

National Research Foundation grant DNRF96 to the Center for Macroecology, Evolution and Climate.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.06.038.

References

Agnarsson, I., Coddington, J.A., Kuntner, M., 2013. Systematics: progress in the study of spider diversity and evolution. In: Penney, D. (Ed.), Spider Research in the 21st Century: Trends and Perspectives. Siri Scientific Press, Manchester, pp. 58–111.

Altenhoff, A.M., Gil, M., Gonnet, G.H., Dessimoz, C., 2013. Inferring hierarchical orthologous groups from orthologous gene pairs. PloS One 8, e53786.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Ayoub, N.A., Garb, J.E., Hedin, M., Hayashi, C.Y., 2007. 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. 42, 394–409.

Babb, P.L., Lahens, N.F., Correa-Garhwal, S.M., Nicholson, D.N., Kim, E.J., Hogenesch, J.B., Kuntner, M., Higgins, L., Hayashi, C.Y., Agnarsson, I., 2017. The Nephila clavipes genome highlights the diversity of spider silk genes and their complex expression. Nat. Genet. 49, 895–903.

Bailey, S.E., Mao, X., Struebig, M., Tsagkogeorga, G., Csorba, G., Heaney, L.R., Sedlock, J., Stanley, W., Rouillard, J.M., Rossiter, S.J., 2016. The use of museum samples for large-scale sequence capture: a study of congeneric horseshoe bats (family Rhinolophidae). Biol. J. Linnean Soc. 117, 58–70.

Bairoch, A., Apweiler, R., 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucl. Acids Res. 28, 45–48.

Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R., Moritz, C., 2013. Unlocking the vault: next-generation museum population genomics. Mol. Ecol. 22, 6018–6032.

Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., Ward, P.S., 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evol. Biol.* 15.

Blaimer, B.B., Lloyd, M.W., Guillory, W.X., Brady, S.G., 2016. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. PloS One 11, e0161531.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Bond, J.E., Garrison, N.L., Hamilton, C.A., Godwin, R.L., Hedin, M., Agnarsson, I., 2014. Phylogenomics resolves a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution. Curr. Biol. 24, 1765–1771.

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

Cerveira, A.M., Jackson, R.R., 2005. Specialised predation by *Palpimanus* sp. (Araneae: Palpimanidae) on jumping spiders (Araneae: Salticidae). J. East African Nat. History 94, 303–317.

Coddington, J.A., Levi, H.W., 1991. Systematics and evolution of spiders (Araneae). Annu. Rev. Ecol. System. 22, 565–592.

Cotoras, D.D., Murray, G.G., Kapp, J., Gillespie, R.G., Griswold, C., Simison, W.B., Green, R.E., Shapiro, B., 2017. Ancient DNA Resolves the History of *Tetragnatha* (Araneae, Tetragnathidae) Spiders on Rapa Nui. Genes 8, 403.

Dimitrov, D., Benavides, L.R., Arnedo, M.A., Giribet, G., Griswold, C.E., Scharff, N., Hormiga, G., 2017. Rounding up the usual suspects: a standard target-gene approach for resolving the interfamilial phylogenetic relationships of ecribellate orb-weaving spiders with a new family-rank classification (Araneae, Araneoidea). Cladistics-Int. J. Willi Hennig Soc. 33, 221–250.

Dimitrov, D., Lopardo, L., Giribet, G., Arnedo, M.A., Álvarez-Padilla, F., Hormiga, G., 2012. Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling. Proc. Royal Soc. B: Biol. Sci. 279, 1341–1350.

Dunn, C.W., Howison, M., Zapata, F., 2013. Agalma: an automated phylogenomics workflow. BMC Bioinform. 14, 330.

Faircloth, B.C., 2017. Identifying conserved genomic elements and designing universal probe sets to enrich them. Methods Ecol. Evol. 8, 1103–1112.

Faircloth, B.C., 2013. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. https://doi.org/10.6079/J9ILL.

Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32, 786–788.

Faircloth, B.C., Branstetter, M.G., White, N.D., Brady, S.G., 2015. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol. Ecol. Resour. 15, 489–501.

Faircloth, B.C., Glenn, T.C., 2012. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS One* 7.

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, 717–726.

Fernández, R., Hormiga, G., Giribet, G., 2014. Phylogenomic analysis of spiders reveals nonmonophyly of orb weavers. Curr. Biol. 24, 1772–1777.

Fernández, R., Kallal, R.J., Dimitrov, D., Ballesteros, J.A., Arnedo, M.A., Giribet, G., Hormiga, G., 2018. Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. Curr. Biol. 28, 1489–1497.

Forster, R.R., Platnick, N.I., 1984. A review of the archaeid spiders and their relatives, with notes on the limits of the superfamily Palpimanoidea (Arachnida, Araneae). Bull. Am. Museum Nat. History 178, 1–106.

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152.

Garrison, N.L., Rodriguez, J., Agnarsson, I., Coddington, J.A., Griswold, C.E., Hamilton, C. A., Hedin, M., Kocot, K.M., Ledford, J.M., Bond, J.E., 2015. Spider phylogenomics: untangling the Spider Tree of Life. In: PeerJ PrePrints.

Godwin, R.L., Opatova, V., Garrison, N.L., Hamilton, C.A., Bond, J.E., 2018. Phylogeny of a cosmopolitan family of morphologically conserved trapdoor spiders (Mygalomorphae, Ctenizidae) using Anchored Hybrid Enrichment, with a description of the family, Halonoproctidae Pocock 1901. Mol. Phylogenet. Evol. 126, 303–313.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., 2011a. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., 2011b. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29, 644–652.

Guschanski, K., Krause, J., Sawyer, S., Valente, L.M., Bailey, S., Finstermeier, K., Sabin, R., Gilissen, E., Sonet, G., Nagy, Z.T., 2013. Next-generation museomics disentangles one of the largest primate radiations. Syst. Biol. 62, 539–554.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protocols 8, 1494–1512.

Hamilton, C.A., Lemmon, A.R., Lemmon, E.M., Bond, J.E., 2016. Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. BMC Evol. Biol. 16, 212.

Harris, R.S., 2007. Improved pairwise alignment of genomic DNA. [Pennsylvania State University].

Hedin, M., Derkarabetian, S., Ramírez, M.J., Vink, C., Bond, J.E., 2018. Phylogenomic reclassification of the world's most venomous spiders (Mygalomorphae, Atracinae), with implications for venom evolution. Sci. Reports 8, 1636.

Hormiga, G., Griswold, C.E., 2014. Systematics, phylogeny, and evolution of orb-weaving spiders. Annu. Rev. Entomol. 59, 487–512.

Jones, M.R., Good, J.M., 2016. Targeted capture in evolutionary and ecological genomics. Mol. Ecol. 25, 185–202.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucl. Acids Res. 30, 3059–3066.

Katoh, K., Standley, D.M., 2013. Mafft multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings Bioinform. 9, 286–298.

Krehenwinkel, H., Pekar, S., 2015. An analysis of factors affecting genotyping success from museum specimens reveals an increase of genetic and morphological variation during a historical range expansion of a European spider. PloS One 10, e0136337.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14, 82.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.

Legendre, R., 1961. Études sur les *Archaea* (Aranéides). – II. La capture des proies et la prise de nourriture. Bull. Zool. Soc. France 86, 316–319.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.

Maddison, W.P., Maddison, D.R., 2010. Mesquite: a modular system for evolutionary analysis, version 2.7. http://mesquiteproject.org.

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, 746–754.

McCormack, J.E., Tsai, W.L., Faircloth, B.C., 2016. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16, 1189–1203.

McGuire, J.A., Cotoras, D.D., O'Connell, B., Lawalata, S.Z., Wang-Claypool, C.Y., Stubbs, A., Huang, X., Wogan, G.O., Hykin, S.M., Reilly, S.B., 2018. Squeezing water from a stone: high-throughput sequencing from a 145-year old holotype resolves (barely) a cryptic species problem in flying lizards. PeerJ 6, e4470.

Michalik, P., Ramírez, M., 2014. Evolutionary morphology of the male reproductive system, spermatozoa and seminal fluid of spiders (Araneae, Arachnida) - current knowledge and future directions. Arthropod Struct. Develop. 43, 291–322.

Miller, J.A., Beentjes, K.K., van Helsdingen, P., IJland, S., 2013. Which specimens from a museum collection will yield DNA barcodes? A time series study of spiders in alcohol. *Zookeys* 245–261.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.

Quicke, D.L., Lopez-Vaamonde, C., Belshaw, R., 1999. Preservation of hymenopteran specimens for subsequent molecular and morphological study. Zool. Scripta 28, 261–267.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*:syy032.

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, 539–542.

Roth, A.C., Gonnet, G.H., Dessimoz, C., 2008. Algorithm of OMA for large-scale orthology inference. BMC Bioinform. 9, 518.
Ruane, S., Austin, C.C., 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. Mol. Ecol. Resour. 17, 1003–1008.
Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.
Sanggaard, K.W., Bechsgaard, J.S., Fang, X., Duan, J., Dyrlund, T.F., Gupta, V., Jiang, X., Cheng, L., Fan, D., Feng, Y., 2014. Spider genomes provide insight into composition and evolution of venom and silk. Nat. Commun. 5 (3765).
Schwager, E.E., Sharma, P.P., Clarke, T., Leite, D.J., Wierschin, T., Pechmann, M., Akiyama-Oda, Y., Esposito, L., Bechsgaard, J., Bilde, T., 2017. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. bioRxiv:106385.
Selden, P.A., Huang, D.Y., Ren, D., 2008. Palpimanoid spiders from the Jurassic of China. J. Arachnol. 36, 306–321.
Sharma, P.P., Kaluziak, S.T., Pérez-Porro, A.R., González, V.L., Hormiga, G., Wheeler, W.C., Giribet, G., 2014. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. Mol. Biol. Evol. 31, 2963–2984.
Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034–1050.
Simon, E., 1893. Histoire naturelle des araignées. Paris 1, 256–488.
Smith, B.T., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Brumfield, R.T., 2013. Target capture and massively parallel sequencing of ultraconserved elements (UCEs) for comparative studies at shallow evolutionary time scales. Syst. Biol. 63, 83–95.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771.
Starrett, J., Derkarabetian, S., Hedin, M., Bryson, R.W., McCormack, J.E., Faircloth, B.C., 2017. High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. Mol. Ecol. 17, 812–823.
Tagliacollo, V.A., Lanfear, R., 2018. Estimating improved partitioning schemes for ultraconserved elements. *Mol. Biol. Evol.*:msy069.
Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564–577.
Vink, C.J., Thomas, S.M., Paquin, P., Hayashi, C.Y., Hedin, M., 2005. The effects of preservatives and temperatures on arachnid DNA. Invertebrate System. 19, 99–104.
Wandeler, P., Hoeck, P.E., Keller, L.F., 2007. Back to the future: museum specimens in population genetics. Trends Ecol. Evol. 22, 634–642.
Wheeler, W.C., Coddington, J.A., Crowley, L.M., Dimitrov, D., Goloboff, P.A., Griswold, C.E., Hormiga, G., Prendini, L., Ramírez, M.J., Sierwald, P., et al., 2016. The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. Cladistics-Int. J. Willi Hennig Soc. 1–43.
Wood, H., 2008. A revision of the assassin spiders of the *Eriauchenius gracilicollis* group, a clade of spiders endemic to Madagascar (Araneae : Archaeidae). Zool. J. Linnean Soc. 152, 255–296.
Wood, H.M., Gillespie, R.G., Griswold, C.E., Wainwright, P.C., 2015. Why is Madagascar special? The extraordinarily slow evolution of pelican spiders (Araneae, Archaeidae). Evolution 69, 462–481.
Wood, H.M., Griswold, C.E., Gillespie, R.G., 2012. Phylogenetic placement of pelican spiders (Archaeidae, Araneae), with insight into evolution of the “neck” and predatory behaviours of the superfamily Palpimanoidea. Cladistics-Int. J. Willi Hennig Soc. 28, 598–626.
Wood, H.M., Matzke, N.J., Gillespie, R.G., Griswold, C.E., 2013. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. Syst. Biol. 62, 264–284.
Wood, H.M., Parkinson, D.Y., Griswold, C.E., Gillespie, R.G., Elias, D.O., 2016. Extremely rapid predatory strikes evolved repeatedly in trap-jaw spiders. Curr. Biol. 26, 1–5.
World Spider Catalog, 2018. Natural History Museum Bern, online at http://wsc.nmbe.ch, version 18, accessed on June 11, 2018 [Internet].
Zhang, C., Sayyari, E., Mirarab, S. (Eds.), 2017. RECOMB International Workshop on Comparative Genomics.
Zonstein, S., Marusik, Y.M., 2013. On Levymanus, a remarkable new spider genus from Israel, with notes on the Chediminae (Araneae, Palpimanidae). Zookeys 27–45.
Zonstein, S.L., Marusik, Y.M., 2017. Descriptions of the two-eyed African spider genera *Chedimanops* gen. n. and *Hybosidella* gen. n. (Araneae, Palpimanidae, Chediminae). African Invertebr. 58, 23–47.
Zonstein, S.L., Marusik, Y.M., Omelko, M.M., 2016. Redescription of the type species of *Diaphorocellus* Simon, 1893 (Araneae, Palpimanidae, Chediminae). African Invertebr. 57, 93–103.