Molecular Phylogenetics and Evolution 100 (2016) 160–169

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenomic analyses reveal novel relationships among snake families

Jeffrey W. Streicher [a,b], John J. Wiens [a,*]

[a] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721-0088, USA
[b] Department of Life Sciences, The Natural History Museum, London SW7 5BD, UK

ARTICLE INFO

Article history:
Received 12 October 2015
Revised 25 March 2016
Accepted 11 April 2016
Available online 12 April 2016

Keywords:
Phylogenomics
Reptiles
Species-tree methods
Serpentes
Snakes
UCE

ABSTRACT

Snakes are a diverse and important group of vertebrates. However, relationships among the major groups of snakes have remained highly uncertain, with recent studies hypothesizing very different (and typically weakly supported) relationships. Here, we address family-level snake relationships with new phylogenomic data from 3776 nuclear loci from ultraconserved elements (1.40 million aligned base pairs, 52% missing data overall) sampled from 29 snake species that together represent almost all families, a dataset ~100 times larger than used in previous studies. We found relatively strong support from species-tree analyses (NJst) for most relationships, including three largely novel clades: (1) a clade uniting the boas, pythons and their relatives, (2) a clade placing cylindrophiids and uropeltids with this clade, and (3) a clade uniting bolyeriids (Round Island boas) with pythonids and their relatives (xenopeltids and loxocemids). Relationships among families of advanced snakes (caenophidians) were also strongly supported. The results show the potential for phylogenomic analyses to resolve difficult groups, but also show a surprising sensitivity of the analyses to the inclusion or exclusion of outgroups.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Snakes are a diverse and important group of vertebrates. They include ~3500 described species (Uetz et al., 2014), currently distributed among 24 families (taxonomy from Pyron et al., 2013). Snakes are particularly important to humans because they include many dangerously venomous species. These species are responsible for tens of thousands of human deaths per year (Kasturiratne et al., 2008). Yet, at the same time, snake venoms are also a valuable resource for medicine (Fox and Serrano, 2007). Additionally, snakes are a model system for many researchers across many fields, and have been the focus of comparative studies (i.e. among species) of molecular evolution, behavior, physiology, functional morphology, and ecology (e.g. Fry and Wüster, 2004; Gartner et al., 2009; Colston et al., 2010; Castoe et al., 2013; Vonk et al., 2013; Byrnes and Jayne, 2014; Senter et al., 2014; Bellini et al., 2015). These comparative studies generally require a phylogenetic framework.

The relationships among the major groups of snakes have proved to be surprisingly difficult to resolve. Among recent studies, very few relationships are universally agreed upon (Fig. 1A). This is surprising because recent studies have applied relatively large amounts of genetic and phenotypic data to these relationships. For example, Wiens et al. (2012) analyzed 44 nuclear loci for most snake families (Fig. 1B), but still found only weak support for some relationships (e.g. placement of boids, pythonids, bolyeriids, cylindrophiids, and uropeltids to each other). Reeder et al. (2015) added 2 loci, 691 morphological characters, and many fossil taxa to that dataset, but still found weak support for relationships among these same major snake clades (Fig. 1B). Pyron et al. (2013) analyzed a smaller number of mitochondrial and nuclear genes (12 total) but many extant snake species (1262 sp.), and found several relationships (Fig. 1C) that contradicted these two studies, with varying levels of support. For example, they weakly placed anomalepidids as sister to all other snakes (instead of leptotyphlopids and typhlopids), and bolyeriids (Round Island boas) as sister to a clade including calabariids, boids, cylindrophiids, uropeltids, xenopeltids, loxocemids, and pythonids. They also found some unusual relationships within advanced snakes (caenophidians), including the placement of acrochordids and xenodermatids as sister taxa (Fig. 1C) and placement of homalopsids with elapids and lamprophiids. An analysis by Zheng and Wiens (2016) combined the molecular datasets of Wiens et al. (2012) and Pyron et al. (2013), to yield a dataset of 52 genes with extensive taxon sampling (1262 species). The resulting maximum likelihood estimate for snake families (Fig. 1D) was similar to that of Wiens et al. (2012) and Reeder et al. (2015). However, the estimate of Zheng and Wiens (2016) was similar to that of Pyron et al. (2013) in weakly

* Corresponding author.
E-mail addresses: j.streicher@nhm.ac.uk (J.W. Streicher), wiensj@email.arizona.edu (J.J. Wiens).

http://dx.doi.org/10.1016/j.ympev.2016.04.015
1055-7903/© 2016 Elsevier Inc. All rights reserved.

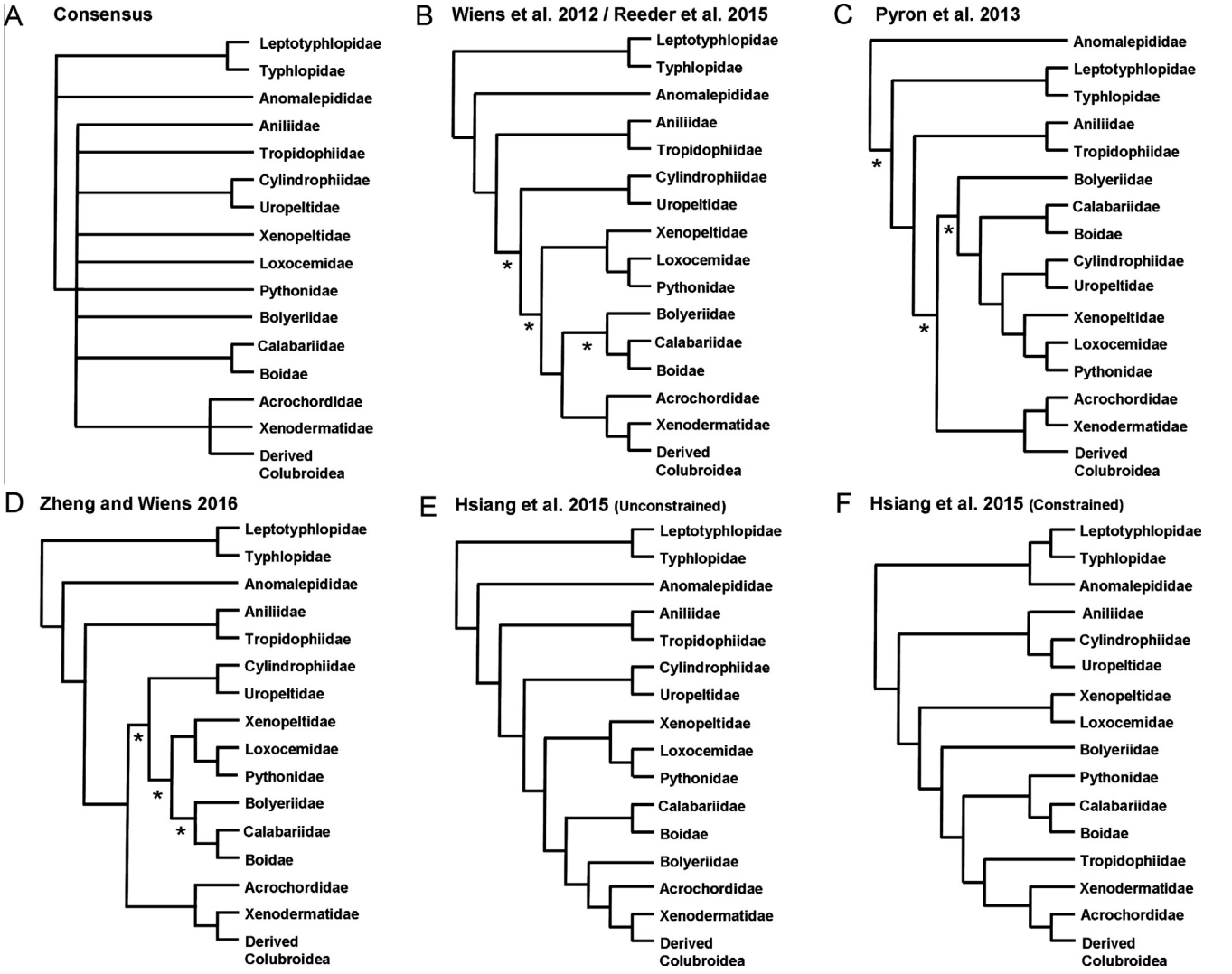

Fig. 1. Summary of recent hypotheses of higher-level snake phylogeny (B–F), including a strict consensus tree (A) of these hypotheses. Asterisks indicate relatively weakly supported nodes (support values less than 70%). The tree of Wiens et al. (2012) matches that of Reeder et al. (2015) for snake families. The trees in B, C, and D are based on concatenated maximum likelihood analyses. E and F are based on Bayesian analyses of Hsiang et al. (2015, their Figs. 2 and 4, respectively), where E is the molecular-only analysis and F is the constrained, combined analysis of molecular and morphological data. A few families were not included in all analyses, including Anomochilidae (related to or nested inside Cylindrophiidae; Gower et al., 2005; Pyron et al., 2013; Zheng and Wiens, 2016), Xenophiidae (most likely related to Bolyeriidae; Lawson et al., 2004; Zheng and Wiens, 2016), and Gerrhopilidae and Xenotyphlopidae (related to Typhlopidae; Vidal et al., 2007; Pyron et al., 2013; Zheng and Wiens, 2016).

supporting a clade including bolyeriids, calabariids, boids, cylindrophiids, uropeltids, xenopeltids, loxocemids, and pythonids (as sister to caenophidians). Hsiang et al. (2015) recently analyzed 21 nuclear loci and 766 morphological characters for a dataset including most snake families. They found strong support for relationships among many snake families in their separate analyses of the molecular data (Fig. 1E), relationships which were largely concordant with those from other recent molecular studies (but with some relationships at odds with previous molecular studies, such as placing bolyeriids as sister to caenophidians). In contrast, they found weak support for most snake relationships from their separate analyses of the phenotypic data. Their analyses of the combined phenotypic and molecular data were generally consistent with the molecular results, but with very weak support (except for placing anomalepidids with leptotyphlopids and typhlopids). Their preferred tree (Fig. 1F) was based on the combined molecular and phenotypic data, but with many relationships constrained to match the phenotype-based tree (despite the weak support for many of these relationships from the phenotypic data). This tree was strongly supported (seemingly because of the constraints) but very different from other recent analyses of snake relationships (Fig. 1F), including their unconstrained analyses of the same combined dataset. In summary, a strict consensus tree of snake relationships from recent analyses is largely unresolved (Fig. 1A).

In this study, we attempt to resolve higher-level snake relationships using new phylogenomic data and an explicit species-tree approach. Importantly, species-tree methods have not been used in previous studies of snake phylogeny at the deepest phylogenetic scales. We utilize ultra-conserved elements (UCEs) for sequence capture (e.g. Bejerano et al., 2004; Sandelin et al., 2004), and generate a molecular dataset that is ~100 times larger than that used in previous studies of this phylogenetic question. We utilize an explicit species-tree method (NJst; Liu and Yu, 2011) to estimate the phylogeny, given that species-tree methods may generally be more accurate than concatenated analyses for multi-locus data (e.g. Edwards et al., 2007; Leaché and Rannala, 2011). In particular, NJst may be more accurate than concatenated analysis when internal branches are relatively short (Liu and Yu, 2011), and short branches are known to be problematic for higher-level snake phylogeny, given their association with weak branch support and

conflicting gene trees (Wiens et al., 2008, 2012). NJst is especially useful because it allows inclusion of loci with missing outgroup data (greatly increasing the number of loci that can be included; e.g. Streicher et al., 2016), and can perform relatively rapid analyses of very large phylogenomic datasets (Liu and Yu, 2011). We do not incorporate the phenotypic data available for snakes, given that recent analyses have shown that these datasets generally have little impact in combined analyses with extensive molecular datasets (e.g. Reeder et al., 2015), and cannot be readily incorporated with species-tree analyses using currently available methods.

2. Materials and methods

2.1. Taxonomic sampling

Species sampled are listed in Table 1. Our ingroup sampling utilized many of the same species (and individuals) that were used in previous large-scale analyses of snake and squamate phylogenetics (e.g. Wiens et al., 2012; Reeder et al., 2015). Importantly, this sampling included representatives of almost all snake families, with multiple representatives of more species-rich families (e.g. Boidae, Colubridae, Pythonidae). Only a few families were not included, due to lack of usable tissue samples. These included Anomochilidae (which are most likely within or closely related to Cylindrophiidae; Gower et al., 2005; Pyron et al., 2013), Xenophidiidae (most likely closely related to Bolyeriidae, Lawson et al., 2004; Zheng and Wiens, 2016), and two families considered part of Typhlopidae until recently (Gerrhopilidae, Xenotyphlopidae; Vidal et al., 2007). Voucher specimens are listed in Table S1.

Recent analyses have shown strong support for a clade (Toxicofera) uniting snakes with anguimorphs and iguanians (e.g. Townsend et al., 2004; Vidal and Hedges, 2005; Wiens et al., 2010, 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016). Therefore, we included a total of five outgroup taxa, representing major clades within Iguania (Acrodonta: Agamidae, *Hydrosaurus*, data from Streicher et al., 2016; Pleurodonta: Dactyloidae, *Anolis*, data from Alföldi et al., 2011) and Anguimorpha (Anguidae, *Anniella*; Lanthanotidae, *Lanthanotus*). We also included a more distant outgroup from Gekkota (Carphodactylidae, *Saltuarius*), a clade relatively close to the squamate root (e.g. Townsend et al., 2004; Vidal and Hedges, 2005; Wiens et al., 2010, 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016). Preliminary analyses showed that including additional outgroup species (from Lacertoidea and Scincoidea) or excluding the sampled gekkotan had little impact on the resulting trees.

2.2. Targeted sequence capture

We targeted ultraconserved elements (UCEs) with sequence capture protocols largely following Faircloth et al. (2012; available at www.ultraconserved.org). UCEs are genomic regions that are highly conserved across divergent taxa (Bejerano et al., 2004; Sandelin et al., 2004). UCEs have been used for phylogenomic studies in several vertebrate groups (e.g. Faircloth et al., 2013; Sun et al., 2014; Crawford et al., 2015; Leaché et al., 2015; Streicher et al., 2016). To capture UCEs from snakes, we used a 5060-locus probe set (designed for tetrapods and available from www.ultraconserved.org). The UCEs targeted by these probes were identified by locating

Table 1
Species, families, number of reads, number of velvet contigs, number of UCEs sequenced, and NCBI Sequence Read Archive (SRA) accession numbers for squamate reptiles used in this study.

Taxon	Family	Reads[a]	Velvet contigs	UCEs	SRA accession
Acrochordus granulatus	Acrochordidae	1,723,934	8481	2681	SAMN04572709
Anilius scytale	Aniliidae	1,591,106	4645	2649	SAMN04572710
Aspidites melanocephalus	Pythonidae	472,141	3596	2767	SAMN04572711
Boa constrictor	Boidae	359,126	1822	1222	SAMN04572712
Boaedon fuliginosus	Lamprophiidae	2,286,465	7309	3042	SAMN04572808
Bothrops asper	Viperidae	1,488,338	5555	3364	SAMN04572713
Calabaria reinhardtii	Calabariidae	1,457,512	5579	3666	SAMN04572714
Casarea dussumieri	Bolyeriidae	366,524	2217	1616	SAMN04572715
Chilabothrus striatus	Boidae	3,483,858	8227	3586	SAMN04572718
Cylindrophis ruffus	Cylindrophiidae	221,695	1451	1151	SAMN04572716
Diadophis punctatus	Colubridae	1,351,991	5261	2906	SAMN04572717
Eryx colubrinus	Boidae	372,088	2548	2035	SAMN04572751
Exiliboa placata	Boidae	380,650*	2220	990	SAMN04572805
Homalopsis buccata	Homalopsidae	578,822	2501	1493	SAMN04572806
Lampropeltis getula	Colubridae	1,556,603	6137	3326	SAMN04572807
Lichanura trivirgata	Boidae	868,673	5186	3651	SAMN04572854
Liotyphlops albirostris	Anomalepididae	2,928,954	6402	2544	SAMN04572856
Loxocemus bicolor	Loxocemidae	493,421	2429	1717	SAMN04572857
Micrurus fulvius	Elapidae	10,942,113	4838	3370	SAMN04572895
Pareas hamptoni	Pareatidae	2,367,023	6592	2550	SAMN04572896
Python molurus	Pythonidae	N/A	N/A	3539	Castoe et al. (2013)
Rena humilis	Leptotyphlopidae	553,189	3724	2390	SAMN04572853
Trachyboa boulengeri	Tropidophiidae	5,929,427	15,782	2912	SAMN04572897
Tropidophis haetianus	Tropidophiidae	518,402	2411	1915	SAMN04572898
Typhlops jamaicensis	Typhlopidae	4,394,464	8755	2801	SAMN04572900
Ungaliophis continentalis	Boidae	1,544,090*	13,181	1251	SAMN04572922
Uropeltis melanogaster	Uropeltidae	9,895,587*	21,225	2403	SAMN04572924
Xenodermus javanicus	Xenodermatidae	1,402,066	7657	2138	SAMN04572925
Xenopeltis unicolor	Xenopeltidae	620,710*	4492	3266	SAMN04572927
Outgroup taxa					
Anolis carolinensis	Dactyloidae	N/A	N/A	4386	Alföldi et al. (2011)
Anniella pulchra	Anguidae	762,095	4266	3333	SAMN04572930
Hydrosaurus sp.	Agamidae	3,298,704	6779	2947	Streicher et al. (2016)
Lanthanotus borneensis	Lanthanotidae	2,171,897	6197	3073	SAMN04572931
Saltuarius cornutus	Carphodactylidae	4,420,846	13,199	3318	SAMN04572932

[a] Asterisks indicate pooled reads from more than a single sequencing effort.

60–100 consecutive bases having 92–100% sequence similarity across whole genome alignments from multiple amniote species (Faircloth et al., 2012). Using this pre-designed set of probes, we ordered a custom Sure Select XT target enrichment kit (Agilent).

DNA was extracted from tissue samples (previously preserved in ethanol, SDS-based lysis buffer, or RNA-later) using either Qiagen DNeasy kits (Qiagen Inc.) or SeraPure magnetic beads (Rohland and Reich, 2012). For DNA extraction using magnetic beads, we first digested a small amount of tissue in a solution of 180 μL of cell lysis buffer (100 mM NaCl, 100 mM Tris-Cl pH 8.0, 25 mM EDTA pH 8.0, 0.5% SDS) and 20 μL of proteinase K (20 mg/mL). Following this digestion we added 360 μL of bead solution (1.8X ratio of beads to sample) and on a magnetic plate washed the beads twice with 70% ethanol (per the standard protocol). We then eluted DNA in 30–100 μL of 10 mM Tris.

We started shotgun genomic library preparation with 100–200 ng of double-stranded DNA for each individual (concentrations determined via HS Assay Qubit fluorometer; Life Technologies). Each sample was randomly sheared using NEBNext dsDNA Fragmentase (New England Biolabs) at 37 °C for 25 min. We then used a master mix prep kit to repair ends, a-tail, and ligate custom adapters on each sample (NEBNext DNA library prep Master Mix Set; New England Biolabs). All samples were cleaned between treatments using SeraPure magnetic beads. For the ligation step, we ordered oligonucleotides to construct 48 uniquely barcoded adapters (Table S2). Adapters were constructed by mixing 20 μL of each oligonucleotide (200 uM each) with 10 μL of 10X annealing buffer (500 mM NaCl, 100 mM Tris–HCL, 10 mM EDTA), and 50 μL of water. We then incubated this mixture at 95 °C for 2 min on a heat block. Following this step we turned the heat block off and allowed the annealed adapters to reach room temperature over several hours. After ligation, we pooled samples into groups of 12–25 samples. We size-selected pooled ligated samples using a Sage Science Pippen Prep at a range of 438–538 bp. However, we acknowledge that the limited range of our size selection might have contributed to the incompleteness of the final dataset, given that most other UCE studies have used wider size selection strategies (e.g. Faircloth et al., 2014; 400–800 bp) or no size selection at all (e.g. Leaché et al., 2015). Following size-selection, we combined sample pools (between 25 and 48 individuals in each "master" pool) and performed a pre-capture enrichment PCR for 12 cycles. We used a high-fidelity Phusion polymerase (NEB) and Illumina® TruSeq primers. We confirmed successful library amplification and estimated library concentration using Bioanalyzer DNA 7500 chips (Agilent). We did not proceed to the sequence capture phase of the protocol with libraries that had less than 150 ng/uL of DNA.

Sequence capture was performed using Dynabeads® (M-270 Streptavidin; Life Technologies) and the Sure Select XT target enrichment kit to hybridize probes to UCE fragments in our shotgun library preparation. We performed a post-hybridization PCR for 18 cycles with Illumina TruSeq® primers. We confirmed that the capture library had been successfully amplified by visualizing PCR products on the Bioanalyzer. We sequenced capture libraries (either 25 or 48 individuals at a time) using an Illumina MiSeq 600-cycle sequencing paired-end run (i.e. 300 bp per paired read) at the University of Texas at Arlington genomics core facility (www.gcf.uta.edu). Three separate sequencing runs were used to generate the sequences used in this study. We processed demultiplexed sequence data to trim low quality ends and remove adapter contamination with the program *illumiprocessor* 2.0.2 (Faircloth et al., 2013; Bolger et al., 2014).

2.3. De-novo assembly, identification, and alignment of UCEs

We assembled all sequence data that passed quality filtering (read 1 + read 2 + singleton reads) for each species using Velvet 1.2.10 (Zerbino and Birney, 2008). We used a kmer length of 75 and a coverage cut-off of 10 in all Velvet 1.2.10 assemblies. We then moved resulting contigs files to a single folder and processed them with *phyluce* 2.0.0 (Faircloth et al., 2012). We identified those contigs that contained UCE sequences using the "fetch uce-contigs" command. We then assembled these together using the "fetch uce-counts" and "fetch uce-fastas" commands. Resulting UCE fasta files were aligned using the MUSCLE algorithm (Edgar, 2004) using default settings.

For five species, we only captured a handful of UCEs in the first two sequencing runs. Therefore, we re-ran five individuals in a third run with fewer multiplexed individuals (*Casarea*, *Exiliboa*, *Ungaliophis*, *Uropeltis*, and *Xenopeltis*). We then ran *Velvet* on all available data for these individuals to maximize the number of UCEs identified. In addition to the data generated for this study we also downloaded previously sequenced sets of UCEs for the species *Anolis carolinensis* and *Python molurus* (genomic data from Alföldi et al. (2011) [*Anolis*] and Castoe et al. (2013) [*Python*]; UCEs from Faircloth et al. (2012) and Faircloth (2016)). We also utilized our previously published data for *Hydrosaurus* (Streicher et al., 2016). The number of UCEs captured for each taxon is listed in Table 1.

From these alignments we filtered the UCE data to exclude any UCEs that did not have data for at least 50% of the taxa using the "align adjust" command (see below for justification for the 50% cutoff). This pipeline resulted in thousands of individual UCE alignments. These alignments were then used as the basis for our concatenated and species-tree analyses. Alignments used and trees estimated are available from the Dryad Digital Repository (doi:10.5061/dryad.4m402).

We note that some authors have excluded many loci in order to generate data matrices with few or no missing data cells (e.g. Leaché et al., 2014; Pyron et al., 2014). However, simulations and empirical analyses have shown that excluding characters and/or taxa solely because they contain missing data may reduce phylogenetic accuracy (Wiens and Morrill, 2011; Wiens and Tiu, 2012; Wagner et al., 2013; Huang and Knowles, 2014; Jiang et al., 2014). For our analyses here, eliminating missing data cells would require eliminating large number of loci, with little justification (other than the desire to eliminate missing data cells). Our recent work with squamate UCEs has suggested that support (and accuracy) appear to be maximized when allowing an intermediate level of missing data, specifically, when including loci that have up to 50% missing data among the sampled species (Streicher et al., 2016). We also performed limited analyses (see Section 3) that were more stringent about the inclusion of loci (e.g. including only those with no more than 25% missing taxa per locus, and then 30%). Using the criterion of 25% dramatically decreased the number of loci included (from 3776 to 241) and led to a somewhat different topology and weak support values (Fig. S1). Including only loci with up to 30% missing data reduced the number of loci somewhat (to 2421) but had no impact on the topology and increased support values at some nodes (Fig. S2).

2.4. Species-tree analyses

We used the program NJst (Liu and Yu, 2011) to perform species-tree analyses. This approach allows estimation of species trees when the gene trees do not contain all the relevant species and when there are large numbers of loci and taxa, unlike many other methods. To generate gene trees for the species tree analysis, we ran a separate maximum likelihood analysis on each UCE locus using the GTR + Γ model on each alignment. Maximum likelihood analyses were performed using RAxML 8.0 (Stamatakis, 2014). Given that using the incomplete matrix options in *phyluce* 2.0.0 can result in alignments with empty taxa (i.e. taxa containing only

missing data), before running RAxML we prepared individual UCE alignments by running a python script that removes empty taxa (written by Pashalia Kapali). We took the resulting RAxML bootstrap topologies and uploaded them to the Species Tree Web Server (STRAW; Shaw et al., 2013) in order to run NJst. We used the bootstrapping method of Seo (2008) to assess nodal support. This method uses a two-stage procedure, where UCE trees are resampled followed by a resampling of bootstrap pseudoreplicates within the sampled trees. Although interpretation of branch lengths from NJst is not straightforward (Shaw et al., 2013), recent analyses suggest that they are strongly correlated with those based on concatenated likelihood analyses (Streicher et al., 2016). This can also be seen qualitatively in our results (compare Figs. 2 and 3).

2.5. Concatenated analyses

Individual UCE alignments were concatenated using the convert function of *phyluce* 2.0.0. We used RAxML 8.0 (Stamatakis, 2014) to perform maximum likelihood analyses. A single GTRCAT model was applied across the entire concatenated dataset. We note that finding optimal partitions would have been challenging given the very large number of loci, and conventional data partitions (e.g. codon positions) are largely lacking in UCE data. Also, the GTRCAT model should account for much of the heterogeneity in rates among sites and loci (CAT is the RAxML fast approximation of the widely used Γ distribution for among-site rate variation; Yang, 1996). Given the limited number of taxa, we performed a single search per matrix. We ran these analyses on the Natural History Museum computing cluster. Nodal support was assessed via 100 bootstrap pseudoreplicates and by mapping bipartitions on the most-likely tree (i.e. RAxML.bipartitions file). We visualized resulting trees using FigTree 1.4.2 (available at http://tree.bio.ed.ac.uk/). Previous analyses suggest that use of GTRCAT does not appear to inflate bootstrap support values for datasets of this type (Streicher et al., 2016). For analyses that only contained ingroup taxa (snakes), we used midpoint rooting instead of designating an outgroup. Note that excluding outgroup taxa changed slightly the number of loci that were included (using the criterion of only including loci with no more than 50% missing taxa per locus).

3. Results

3.1. Species-tree analysis and comparison to other studies

The final dataset included 3776 loci and 1,398,192 base pairs, with 52% missing data overall. Our primary estimate is based on the species-tree analysis (Fig. 2). The results of this analysis are generally very strongly supported, with only a handful of clades with bootstrap support (bs) values less than 90%, and with most weakly supported clades within Boidae. Below, we describe this tree and compare the results to those of other recent analyses of higher-level snake phylogeny.

We support scolecophidians as a paraphyletic group at the base of snake phylogeny. Almost all recent molecular studies of snakes have also suggested that scolecophidians are paraphyletic (e.g. Wiens et al., 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016). Our results here are unusual in placing leptotyphlopids as sister to all other snakes, instead of placing leptotyphlopids and typhlopids together as sister taxa, but the clade uniting typhlopids with other snakes is only moderately well-supported (bs = 88%). There is strong support (bs = 97%) for placing anomalepidids as sister to non-scolecophidian snakes (alethinophidians). This placement of anomalepidids was found in most recent molecular studies (e.g. Wiens et al., 2012; Reeder et al.,

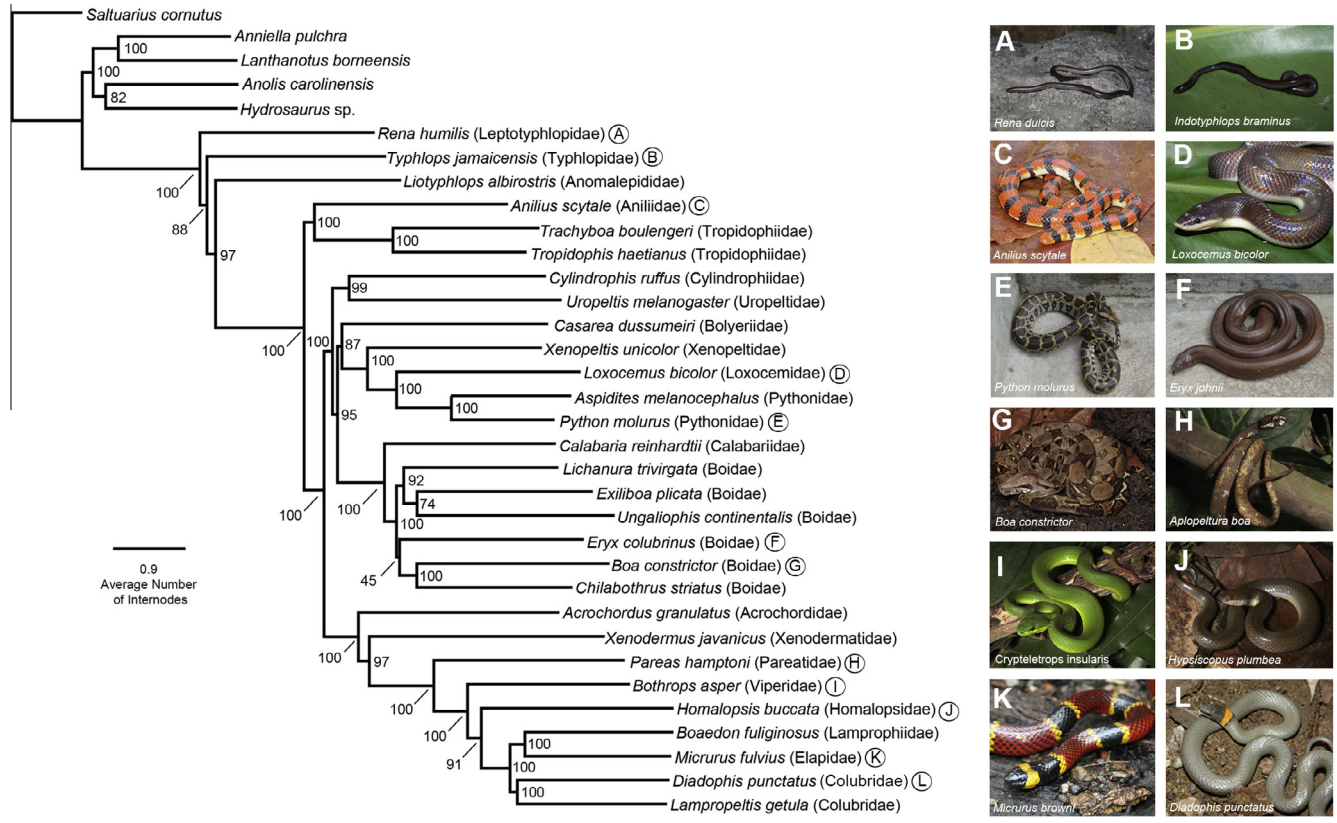

Fig. 2. Estimated phylogeny of major snake clades, based on species-tree analysis (NJst) of 3776 UCE loci and 1,398,192 aligned base pairs, with 52% missing data overall. Numbers at nodes are bootstrap support values. Snake images are from T. Gamble (C), E. Smith (G, I, J), J. Streicher (A, B, D, H, K), and J. Wiens (E, F, L).

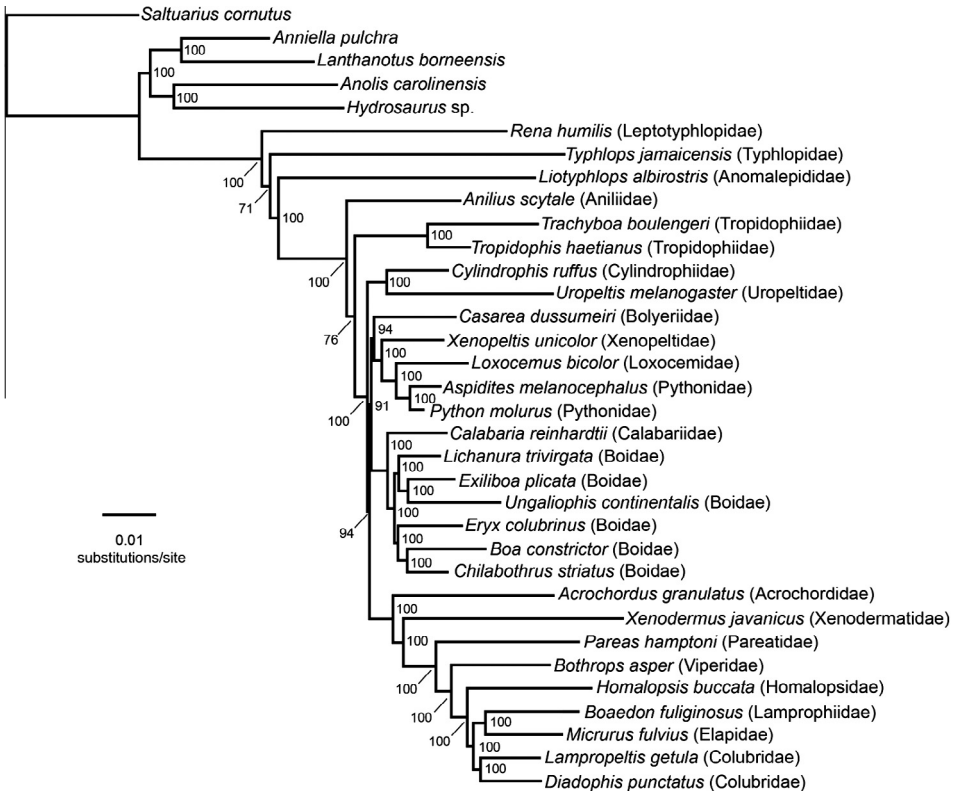

Fig. 3. Estimated tree from concatenated likelihood analysis (RAxML) of 3776 UCE loci (1,398,192 base pairs; 52% missing data), only including loci with no more than 50% missing taxa per locus. Numbers at nodes indicate bootstrap support values.

2015; Zheng and Wiens, 2016) but not Pyron et al. (2013) nor in the constrained combined-data trees of Hsiang et al. (2015).

Alethinophidians (snakes excluding scolecophidians) are strongly supported as monophyletic, and this traditionally recognized group has been supported by previous molecular, morphological, and combined analyses (Fig. 1; e.g. Wiens et al., 2012; Pyron et al., 2013; Hsiang et al., 2015; Reeder et al., 2015; Zheng and Wiens, 2016). Within alethinophidians, there is strong support for a clade uniting aniliids with tropidophiids, a result also found in most previous molecular studies but not in the constrained, combined-data tree of Hsiang et al. (2015). The clade containing the remaining alethinophidians is strongly supported, and has also been found in most previous molecular studies (e.g. Wiens et al., 2008, 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016).

We find three largely novel clades among the other alethinophidians (Fig. 2). The remaining alethinophidians are divided into two strongly supported clades, one containing the advanced snakes (Caenophidia) and the other containing the rest of the alethinophidians, including cylindrophiids, uropeltids, bolyeriids, xenopeltids, loxocemids, pythonids, calabariids, and boids. The latter clade is largely unique to the present study. For example, Wiens et al. (2012) and Reeder et al. (2015) showed the families within this group as being paraphyletic with respect to caenophidians (Fig. 1B), but with weak support for these relationships. However, Pyron et al. (2013) and Zheng and Wiens (2016) both found weak support for this clade. Within this clade, we found a strongly supported clade consisting of cylindrophiids and uropeltids (bs = 100%), which is then the sister group to a strongly supported group including the remaining families (bs = 95%). This latter clade is largely unique to our study (but is also weakly supported by Zheng and Wiens, 2016). Within the latter clade is a strongly supported clade consisting of boids and calabariids (bs = 100%) and a moderately well-supported clade (bs = 87%) uniting bolyeriids with the strongly supported clade of xenopeltids, loxocemids, and pythonids. The grouping of cylindrophiids with uropeltids and of boids with calabariids have each been widely supported in previous studies, as has the clade consisting of xenopeltids, loxocemids, and pythonids (e.g. Wiens et al., 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016). However, most previous studies have not placed these latter two clades together (i.e. boas and relatives and pythons and relatives), making this clade largely unique to our study (but see Zheng and Wiens, 2016). Furthermore, no other recent studies have placed bolyeriids with xenopeltids, loxocemids, and pythonids. For example, analyses by Wiens et al. (2012), Reeder et al. (2015), and Zheng and Wiens (2016) all placed bolyeriids instead as the sister group to the clade of calabariids and boiids (Fig. 1B and D), whereas Pyron et al. (2013) placed bolyeriids as sister to a clade including boiids, calabariids, cylindrophiids, uropeltids, xenopeltids, loxocemids, and pythonids (Fig. 1C). Finally, few recent studies have placed cylindrophiids and uropeltids as sister to a clade of boiids, calabariids, bolyeriids, xenopeltids, loxocemids, and pythonids (but see Zheng and Wiens, 2016).

Relationships within caenophidians were strongly supported and largely consistent with those of recent studies (e.g. Wiens et al., 2012; Reeder et al., 2015; Zheng and Wiens, 2016), showing the acrochordids as sister to all other caenophidians, followed successively by xenodermatids, pareatids, viperids, homalopsids, colubrids, and the clade of elapids and lamprophiids (Fig. 2). These relationships are inconsistent with those of Pyron et al. (2013), who placed xenodermatids with acrochordids, and homalopsids with the clade of lamprophiids and elapids. However, they are consistent with those of Pyron et al. (2014), who used 333 nuclear loci and species-tree methods to examine caenophidian relationships.

3.2. Concatenated analysis and impacts of incompleteness and rooting

Relationships estimated here by concatenated maximum likelihood analysis (Fig. 3) were largely similar to those estimated by the species-tree method (Fig. 2), but with two notable differences. First, the concatenated analysis placed aniliids as sister to all other alethinophidians (with moderate support, bs = 76%), rather than with tropidophiids as in most other recent studies (e.g. Wiens et al., 2012; Pyron et al., 2013; Reeder et al., 2015; Zheng and Wiens, 2016). Second, the concatenated analysis placed cylindrophiids and uropeltids as sister to a well-supported clade (bs = 91%) consisting of caenophidians and the clade of boids, calabariids, bolyeriids, xenopeltids, loxocemids, and pythonids. Cylindrophiids and uropeltids are placed with this latter clade in the NJst analyses.

Interestingly, the typical placement of aniliids (with tropidophiids) was restored in the concatenated analyses when the outgroup taxa were removed and only snakes are analyzed (Fig. 4A), utilizing midpoint rooting. Similarly, including loci with up to 25% missing taxa per UCE also restored the typical molecular placement of aniliids in the concatenated analysis, but with very low bootstrap support (Fig. S1). The low support is unsurprising given that this criterion for including loci dramatically decreased the sampled loci from 3776 to only 216. Allowing up to 30% missing taxa per UCE locus allowed inclusion of 2421 loci, and yielded similar results to the full concatenated analyses, and did not restore the typical molecular placement of aniliids (Fig. S2).

Removing outgroups from the concatenated analysis and using midpoint rooting (Fig. 4A) has other interesting effects beyond the placement of aniliids. First, leptotyphlopids are placed with typhlopids with strong support (bs = 100%), as suggested in previous molecular, morphological, and combined-data analyses (Fig. 1; Wiens et al., 2012; Pyron et al., 2013; Hsiang et al., 2015; Reeder et al., 2015; Zheng and Wiens, 2016). Second, the monophyly of scolecophidians is strongly supported (bs = 100%). Monophyly of scolecophidians has been strongly contradicted by recent molecular analyses, but is consistent with previous morphological analyses (Hsiang et al., 2015; Reeder et al., 2015), and some combined analyses (Hsiang et al., 2015) but not others (Reeder et al., 2015).

Removing outgroups (non-snakes) from the species-tree analysis and using midpoint rooting leads to placement of the root between caenophidians and all other snakes, a highly improbable root based on all other molecular and morphological analyses. We therefore show the NJst tree without outgroups as rooted following the midpoint-rooted concatenated analysis (Fig. 4B). This unrooted tree is identical to that with outgroups included, with even stronger support for some key nodes (e.g. support for the clade of bolyeriids, xenopeltids, loxocemids, and pythonids increases from 87% to 100%). Interestingly, the unrooted tree from NJst is also consistent with strong support for the leptotyphlopid–typhlopid clade and the monophyly of scolecophidians (bs = 100% each). However, the monophyly of these basal clades does ultimately depend on where the root is placed.

4. Discussion

Relationships among the major clades of snakes have been weakly supported and conflicting among recent studies (Fig. 1). Here, we use a large, novel phylogenomic dataset and an explicit species-tree method to better resolve these relationships, with a dataset including most snake families and 3776 loci. Our results provide strong support for most relationships among major clades of snakes. Furthermore, they strongly support three major clades that were largely absent from most previous studies: (1) a clade uniting the boas, pythons, and relatives (i.e. calabariids, boiids, bolyeriids, xenopeltids, loxocemids, and pythonids) with cylindrophiids and uropeltids as the sister group to caenophidians, (2) a clade uniting the boas, pythons and relatives, excluding the cylindrophiids and uropeltids, and (3) a clade uniting the bolyeriids with the xenopeltids, loxocemids, and pythonids. Below, we discuss the question of whether these relationships are truly resolved, and the unusual results from the concatenated analyses.

An important question in a study such as ours is: how do we know that these relationships are truly resolved? Several lines of evidence suggest that these relationships are correctly resolved, whereas a few others raise concerns about their veracity. On the positive side, we note first that these results are based on the

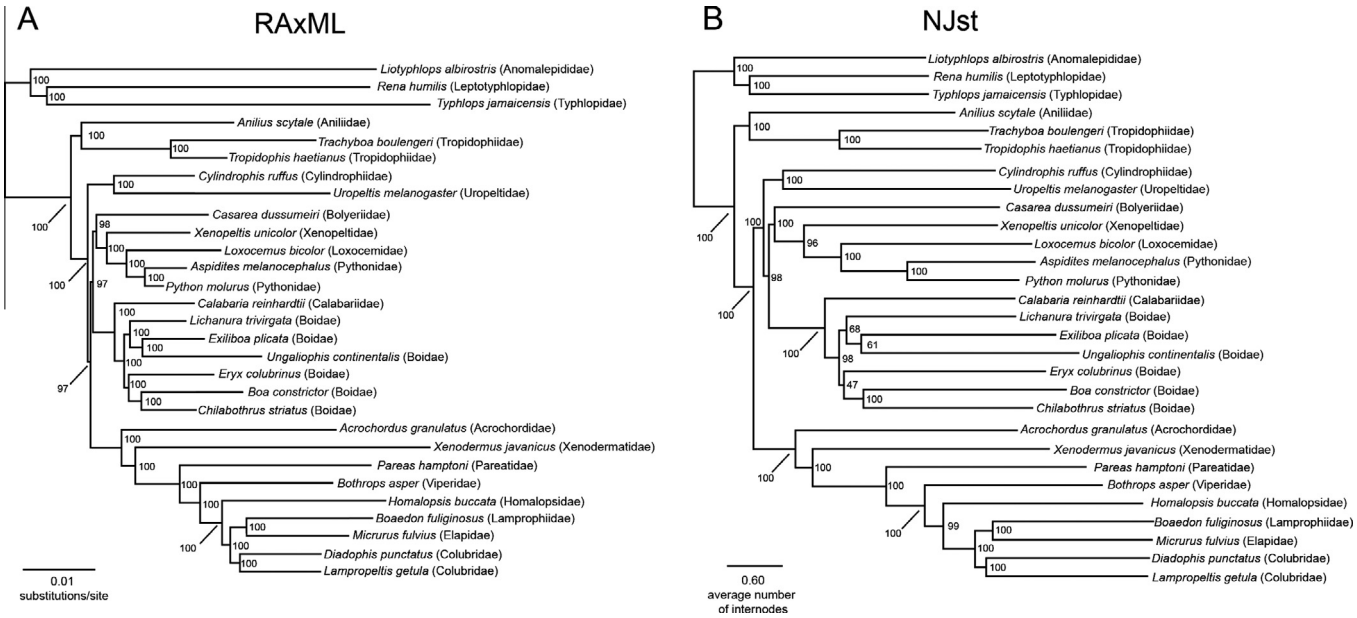

Fig. 4. Estimated trees from concatenated (A) and species-tree (B) analyses, from datasets with outgroup taxa removed. The datasets include 3335 UCE loci (1,287,909 base pairs), and only loci with no more than 50% missing taxa per locus are included (excluding outgroups changes the number of included loci by this criterion). (A) Estimated tree from concatenated likelihood analysis using RAxML, utilizing midpoint rooting. (B) Estimated phylogeny from species-tree analysis (NJst), rooted based on the midpoint-rooted concatenated analysis (midpoint rooting with NJst places the root between caenophidians and all other snakes). Numbers at nodes indicate bootstrap support values.

largest sample of loci (and base pairs) of any study of higher-level snake phylogeny so far, by two orders of magnitude (i.e. 46 loci vs. 3776). Second, we found that most relationships are strongly supported even by species-tree methods, which can give significantly lower support values than concatenated analyses (e.g. Streicher et al., 2016). Third, many of the relationships that we recover are consistent with strongly supported results of previous multi-locus molecular analyses (but using dozens of loci instead of thousands; e.g. Wiens et al., 2012; Pyron et al., 2013, 2014; Reeder et al., 2015; Zheng and Wiens, 2016), including relationships within caenophidians, and the grouping of boids with calabriids, of cylindrophiids with uropeltids, and of xenopeltids, loxocemids, and pythonids. Fourth, many of the novel relationships that we find here are strongly supported, and the conflict with previous studies involves relationships that were only weakly supported in previous studies. For example, the relationships among the cylindrophiid + uropeltid clade, the xenopeltid + loxocemid + pythonid clade, the boid + calabariid clade, and caenophidians were only weakly supported in previous studies (e.g. Fig. 1).

On the negative side, some aspects of the results are troubling. First, there are some disagreements between concatenated and species-tree results (Figs. 2 and 3). One disagreement is that some concatenated analyses support an unusual placement for aniliids (relative to other molecular analyses), whereas species-tree analyses support the more typical placement with tropidophiids. However, this result is only weakly supported and disappears in the concatenated analyses when outgroups are removed (Fig. 4) and in some analyses when less missing data are allowed (Fig. S1). Therefore, it seems clear that this result is artifactual. Nevertheless, this pattern does suggest that there is some misleading signal present in this dataset. The other disagreement is over the placement of the cylindrophiid-uropeltid clade (placed with boas and pythons in the species-tree analysis and as the sister to caenophidians, boas, pythons, and relatives in the concatenated analysis). The placement in the species-tree analysis is consistently strongly supported (bs = 100%), whereas the placement in the concatenated analysis is weaker (bs = 94%). Although disagreements between concatenated and species-tree analyses are potentially a cause for concern, simulation studies suggest that NJst should be more accurate than concatenated analyses for shorter internal branches (Liu and Yu, 2011), and it is on these shorter branches where we expect these methods to disagree (e.g. Lambert et al., 2015). It should also be noted that the concatenated and species-tree analyses agree on two of the major novel clades found in this study (i.e. the clade uniting boas, pythons, and relatives, and the placement of bolyeriids with pythonids, loxocemids, and xenopeltids).

The other troubling result is that there is a clade found in both the species-tree and concatenated analyses that is inconsistent with previous molecular and morphological studies. Specifically, both trees show leptotyphlopids as sister to all other snakes, whereas other recent molecular and combined analyses place leptotyphlopids with typhlopids with strong support (Fig. 1). This relationship is only moderately well supported by the species tree and concatenated analyses (bs = 88% and 76%, respectively). We suspect that this relationship is not correct, and therefore we do not consider it to be one of the major results of our study. Importantly, we find that removing the outgroup taxa (and utilizing midpoint rooting) restores the traditional placement of leptotyphlopids with typhlopids with strong support (bs = 100%), in the concatenated analyses (Fig. 4A). This analysis also strongly supports monophyly of scolecophidians (in contrast to most recent molecular analyses), and both of these results are consistent with the species-tree estimate without outgroups (Fig. 4B).

Overall, our results do show that phylogenomic analyses can yield misleading results, especially from concatenated analyses (given that both placements for tropidophiids cannot be correct). One potential explanation for this pattern is that the matrix used here contains considerable missing data. In this study, we included loci with up to 50% missing data each, and the matrix contained 52% missing data overall. However, we also performed analyses after excluding loci that had more than 25% missing data (Fig. S1). This dramatically reduced the number of UCE loci included, from 3776 to only 216. The trees resulting from the latter analyses show the typical placement of aniliids (with tropidophiids), but not the typical placement of leptotyphlopids, and many relationships among the major clades become only weakly supported. Excluding loci with up to 30% missing data allowed inclusion of more loci (2421) but yielded the same unusual placements for aniliids and leptotyphlopids (Fig. S2). Therefore, simply eliminating loci with more missing data does not seem to be an explanation or solution (see Section 3). Furthermore, those taxa with the most missing data (fewest UCEs, Table 1) are generally placed with strong support in their expected positions in the tree. Specifically, the four most incomplete taxa consist of three boids (Table 1), which are placed within a strongly supported Boidae (*Boa*, *Exiliboa*, *Ungaliophis*), and *Cylindrophis* (Cylindrophiidae), which is strongly supported as sister to Uropeltidae (as in all other recent studies; Fig. 1). Another potential explanation is that there is a bias in the capture of these ultraconserved loci among taxa, potentially related to the accelerated rates of molecular evolution that have been well documented in snakes (e.g. Castoe et al., 2013). However, the majority of relationships estimated here are consistent with estimates from multi-locus (non-UCE) data (Fig. 1), including monophyly of families. Therefore, this problem seems limited in scope, if it is present at all. Limited taxon sampling might also be an issue, especially for scolecophidians (leptotyphlopids, liotyphlopids, and typhlopids). However, many of the pivotal basal snake lineages are relatively species poor (e.g. aniliids, bolyeriids, calabariids, cylindrophiids, loxocemids, tropidophiids, xenopeltids), such that all extant genera have been included (and all extant species for aniliids, bolyeriids, calabariids and loxocemids). Thus, sampling additional species for these seven families may have little impact, since it would not subdivide long branches among families. Importantly, many of the relatively novel relationships that we find here have antecedents in those previous analyses with the most extensive taxon sampling (Pyron et al., 2013; Zheng and Wiens, 2016). In summary, our results illustrate the idea that some caution may still be needed in the interpretation of results from phylogenomic analyses, especially for concatenated analyses.

Finally, our results suggest that the preferred tree of Hsiang et al. (2015) should not be used for evolutionary studies of snakes (or used only with considerable caution). Our results strongly contradict many components of that tree. Moreover, many aspects of that tree were strongly contradicted by previous multi-locus studies (but with far fewer loci and characters) and by combined-data analyses including both molecular and phenotypic characters (including the unconstrained analyses of Hsiang et al. [2015]; Fig. 1E). Furthermore, many of these problematic aspects of the tree that were "constrained" by Hsiang et al. (2015) were not strongly supported by their phenotypic data to begin with. These problematic aspects include: (1) placing pythonids with boids and calabariids, instead of with xenopeltids and loxocemids, (2) placing aniliids with cylindrophiids and uropeltids instead of tropidophiids, (3) placing tropidophiids as sister to caenophidians instead of (with aniliids) as sister to all other alethinophidians, and (4) placing xenodermatids as sister to all other caenophidians, instead of placing acrochordids as sister to colubroids (all other caenophidians). These differences in trees may have important implications for inferences about snake biology. For example, the preferred tree of Hsiang et al. (2015) places most burrowing snake lineages near the root of the tree, with non-burrowing lineages

evolving only once subsequently. However, our results (and those of other recent molecular and combined-data analyses) imply multiple transitions between burrowing and non-burrowing lineages, such as within the clade of aniliids (burrowing) and tropidophiids (non-burrowing), and the clade of xenopeltids (burrowing), loxocemids (burrowing), and pythonids (non-burrowing). Given our results and those of previous studies, we strongly caution against using the preferred tree of Hsiang et al. (2015) in evolutionary studies.

In this paper, we report progress in resolving the major clades of snakes. Our species-tree analyses of 3776 loci reveal well-supported relationships not found in previous studies. They also provide additional support for relationships found in previous studies but with far fewer loci, including non-monophyly of scolecophidians and relationships within caenophidians. More generally our results illustrate the promise of phylogenomics and species-tree analyses to resolve difficult phylogenetic problems that have resisted resolution with smaller numbers of loci, although some caution is still needed.

Acknowledgements

We thank M. Fujita, J. Castoe, T. Castoe, and P. Foster for assistance with laboratory methods, sequencing, and computational resources. We thank P. Kapali for use of her python script for removing empty taxa from alignments. We thank E. Smith and T. Gamble for use of their photographs. We thank G. Mathews, E. Miller, and L. Barrientos for laboratory assistance. We thank T. Castoe for many helpful comments on the manuscript. Some photographs were obtained during trips sponsored by NSF DEB-0613802 to J. Campbell and DEB-1146324 to E. Smith and M. Harvey. We are very grateful to the many researchers and institutions who provided (or facilitated access to) tissues and DNA samples for use in this study: see Table S1 for listing of vouchers, specimens, institutions, and individuals (including C. Austin, D. Cannatella, J. Gauthier, S. B. Hedges, M. Kearney, M.S.Y. Lee, J. McGuire, J. Vindum, H. Voris). We especially thank T. Reeder for help in assembling the set of tissues used in the Deep Scaly project (and here). This work was supported by the University of Arizona.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.04.015.

References

Alföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., Russell, P., Lowe, C.B., Glor, R.E., Jaffe, J.D., Ray, D.A., Boissinot, S., Shedlock, A.M., Botka, C., Castoe, T.A., Colbourne, J.K., Fujita, M.K., Godinez Moreno, R., ten Hallers, B.F., Haussler, D., Heger, A., Heiman, D., Janes, D.E., Johnson, J., de Jong, P.J., Koriabine, M.Y., Lara, M., Novick, P.A., Organ, C.L., Peach, S.E., Poe, S., Pollock, D.D., de Queiroz, K., Sanger, T., Searle, S., Smith, J.D., Smith, Z., Swofford, R., Turner-Maier, J., Wade, J., Young, S., Zadissa, A., Edwards, S.V., Glenn, T.C., Schneider, C. J., Losos, J.B., Lander, M.B., Ponting, C.P., Lindblad-Toh, K., 2011. The genome of the green anole lizard and a comparative analysis with birds and mammals. Nature 477, 587–591.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved elements in the human genome. Science 304, 1321–1325.

Bellini, G.P., Giraudo, A.R., Arzamendia, V., Etchepare, E.G., 2015. Temperate snake community in South America: is diet determined by phylogeny or ecology? PLoS ONE 10, e0123237.

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

Byrnes, G., Jayne, B.C., 2014. Gripping during climbing of arboreal snakes may be safe but not economical. Biol. Lett. 10, 20140434.

Castoe, T.A., de Koning, A.P.J., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P., Degner, J.F., Daza, J.M., Gu, W., Reyes-Velasco, J., Shaney, K.J., Castoe, J.M., Fox, S.E., Poole, A.W., Polanco, D., Dobry, J., Vandewege, M.W., Li, Q., Schott, R., Kapusta, A., Minx, P., Feschotte, C., Uetz, P., Ray, D.A., Hoffman, F.G., Bogden, R., Smith, E.N., Chang, B.S.W., Vonk, F., Casewell, N.R., Henkel, C.V., Richardson, M.K., Mackessy, S.P., Bronikowski, A.M., Yandell, M., Warren, W.C., Secor, S.M., Pollock, D.D., 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc. Natl. Acad. Sci. USA 110, 20645–20650.

Colston, T.J., Costa, G.C., Vitt, L.J., 2010. Snake diets and the deep history hypothesis. Biol. J. Linn. Soc. 101, 476–486.

Crawford, N.G., Parham, J.F., Sellas, A.B., Faircloth, B.C., Glenn, T.C., Papenfuss, T.J., Henderson, J.B., Hansen, M.H., Simison, W.B., 2015. A phylogenomic analysis of turtles. Mol. Phylogenet. Evol. 83, 250–257.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.

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

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

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.

Faircloth, B.C., Sorenson, L., Santini, F., Alfaro, M.E., 2013. A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). PLoS ONE 8, e54848.

Faircloth, B.C., Branstetter, M.G., White, N.D., Brady, S.G., 2014. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol. Ecol. Res. 15, 489–501.

Fox, J.W., Serrano, S.M.T., 2007. Approaching the golden age of natural product pharmaceuticals from venom libraries: an overview of toxins and toxin-derivatives currently involved in therapeutic or diagnostic applications. Curr. Pharm. Des. 13, 2927–2934.

Fry, B.G., Wüster, W., 2004. Assembling an arsenal: origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. Mol. Biol. Evol. 21, 870–883.

Gartner, G.E.A., Hicks, J.W., Manzani, P.R., Andrade, D.V., Abe, A.S., Wang, T., Secor, S. M., Garland Jr., T., 2009. Phylogeny, ecology, and heart position in snakes. Physiol. Biochem. Zool. 83, 43–54.

Gower, D.J., Vidal, N., Spinks, J.N., McCarthy, C.J., 2005. The phylogenetic position of Anomochilidae (Reptilia: Serpentes): first evidence from DNA sequences. J. Zool. Syst. Evol. Res. 43, 315–320.

Huang, H., Knowles, L.L., 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Syst. Biol., syu046

Hsiang, A.Y., Field, D.J., Webster, T.H., Behlke, A.D.B., Davis, M.B., Racicot, R.A., Gauthier, J.A., 2015. The origin of snakes: revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. BMC Evol. Biol. 15, 87.

Jiang, W., Chen, S.-Y., Wang, H., Li, D.-Z., Wiens, J.J., 2014. Should genes with missing data be excluded from phylogenetic analyses? Mol. Phylogenet. Evol. 80, 308–318.

Kasturiratne, A., Wickremasinghe, A.R., de Silva, N., Gunawardena, N.K., Pathmeswaran, A., Premaratna, R., Savioli, L., Lalloo, D.G., de Silva, H.J., 2008. The global burden of snakebite: a literature analysis and modelling based on regional estimates of envenoming and deaths. PLoS Med. 5, e218.

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. 82, 146–155.

Lawson, R., Slowinski, J.B., Burbrink, F.T., 2004. A molecular approach to discerning the phylogenetic placement of the enigmatic snake *Xenophidion schaeferi* among the Alethinophidia. J. Zool. 263, 285–294.

Leaché, A.D., Rannala, B., 2011. The accuracy of species tree estimation under simulation: a comparison of methods. Syst. Biol. 60, 126–137.

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., 2014. 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.

Leaché, A.D., Chavez, A.S., Jones, L.N., Grummer, J.A., Gottscho, A.D., Linkem, C.W., 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biol. Evol. 7, 706–719.

Liu, L., Yu, L., 2011. Estimating species trees from unrooted gene trees. Syst. Biol. 60, 661–667.

Pyron, R.A., Burbrink, F.T., Wiens, J.J., 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evol. Biol. 13, 93.

Pyron, R.A., Hendry, C.R., Chou, V.M., Lemmon, E.M., Lemmon, A.R., Burbrink, F.T., 2014. Effectiveness of phylogenomic data and coalescent species-tree methods for resolving difficult nodes in the phylogeny of advanced snakes (Serpentes: Caenophidia). Mol. Phylogenet. Evol. 81, 221–231.

Reeder, T.W., Townsend, T.M., Mulcahy, D.G., Noonan, B.P., Wood Jr., P.L., Sites Jr., J. W., Wiens, J.J., 2015. Integrated analyses resolve conflicts over squamate reptile phylogeny and reveal unexpected placements for fossil taxa. PLoS ONE 10, e0118199.

Rohland, N., Reich, D., 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22, 939–946.

Sandelin, A., Bailey, P., Bruce, S., Engström, P.G., Klos, J.M., Wasserman, W.W., Ericson, J., Lenhard, B., 2004. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. BMC Genom. 5, 99.
Senter, P., Harris, S.M., Kent, D.L., 2014. Phylogeny of courtship and male–male combat behavior in snakes. PLoS ONE 9, e107528.
Seo, T.K., 2008. Calculating bootstrap probabilities of phylogeny using multilocus sequence data. Mol. Biol. Evol. 25, 960–971.
Shaw, T., Ruan, Z., Glenn, T., Liu, L., 2013. STRAW: Species TRee Analysis Web Server. Nucleic Acids Res. 41, W230–W241.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1315.
Streicher, J.W., Schulte II, J.A., Wiens, J.J., 2016. How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. Syst. Biol. 65, 128–145.
Sun, K., Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2014. The evolution of peafowl and other taxa with ocelli (eyespots): a phylogenomic approach. Proc. R. Soc. Lond. B Biol. Sci. 281, 20140.
Townsend, T., Larson, A., Louis, E.J., Macey, J.R., 2004. Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. Syst. Biol. 53, 735–757.
Uetz, P., Goll, J., Hallermann, J., 2014. The Reptile Database Available: <www.reptile-database.org>. Accessed August, 2014.
Vidal, N., Azvolinsky, A., Cruaud, C., Hedges, S.B., 2007. Origin of tropical American burrowing reptiles by transatlantic rafting. Biol. Lett. 4, 115–118.
Vidal, N., Hedges, S.B., 2005. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. C. R. Biol. 328, 1000–1008.
Vonk, F.J., Casewell, N.R., Henkel, C.V., Heimberg, A., Jansen, H.J., McCleary, R.J.R., Kerkkamp, H.M.E., Vos, R., Guerreiro, I., Calvete, J.J., Wüster, W., Woods, A.E., Logan, J.M., Harrison, R.A., Castoe, T.A., de Koning, A.P.J., Pollock, D.D., Renjifo, C., Currier, R.B., Salgado, D., Pla, D., Sanz, L., Hyder, A.S., Ribeiro, J.M.C., Arntzen, J. W., van den Thillart, G.E.E.J.M., Boetzer, M., Pirovano, W., Dirks, R.P., Spaink, H. P., Duboule, D., McGlinn, E., Kini, R.M., Richardson, M.K., 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proc. Natl. Acad. Sci. USA 10, 20651–20656.
Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22, 787–798.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst. Biol. 60, 719–731.
Wiens, J.J., Tiu, J., 2012. Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. PLoS ONE 7, 42925.
Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D., Sites Jr., J.W., Townsend, T.M., Reeder, T.W., 2008. Branch length, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. Syst. Biol. 57, 420–431.
Wiens, J.J., Kuczynski, C.A., Townsend, T., Reeder, T.W., Mulcahy, D.G., Sites Jr., J.W., 2010. Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the placement of fossil taxa. Syst. Biol. 59, 674–688.
Wiens, J.J., Hutter, C.R., Mulcahy, D.G., Noonan, B.P., Townsend, T.M., Sites Jr., J.W., Reeder, T.W., 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. Biol. Lett. 8, 1043–1046.
Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analysis. Trends Ecol. Evol. 11, 367–372.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assemble using de Bruijn graphs. Genome Res. 18, 821–829.
Zheng, Y., Wiens, J.J., 2016. Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. Mol. Phylogenet. Evol. 94, 537–557.