



Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*)

R. Alexander Pyron^{a,b,*}, Kyle A. O'Connell^{a,b,c}, Emily Moriarty Lemmon^d, Alan R. Lemmon^e, David A. Beamer^f

^a Department of Biological Sciences, The George Washington University, Washington, DC 20052, USA

^b Division of Amphibians and Reptiles, Department of Vertebrate Zoology, National Museum of Natural History Smithsonian Institution, Washington, DC 20560, USA

^c Global Genome Initiative, National Museum of Natural History Smithsonian Institution, Washington, DC 20560, USA

^d Department of Biological Science, Florida State University, Tallahassee, FL 32306-4295, USA

^e Department of Scientific Computing, Florida State University, Tallahassee, FL 32306-4120, USA

^f Department of Natural Sciences, Nash Community College, Rocky Mount, NC, 27804, USA

ARTICLE INFO

Keywords:

Desmognathus
Candidate species
Hybridization
Networks

ABSTRACT

Gene flow between evolutionarily distinct lineages is increasingly recognized as a common occurrence. Such processes distort our ability to diagnose and delimit species, as well as confound attempts to estimate phylogenetic relationships. A conspicuous example is Dusky Salamanders (*Desmognathus*), a common model-system for ecology, evolution, and behavior. Only 22 species are described, 7 in the last 40 years. However, mitochondrial datasets indicate the presence of up to 45 “candidate species” and multiple paraphyletic taxa presenting a complex history of reticulation. Some authors have even suggested that the search for species boundaries in the group may be in vain. Here, we analyze nuclear and mitochondrial data containing 161 individuals from at least 49 distinct evolutionary lineages that we treat as candidate species. Concatenated and species-tree methods do not estimate fully resolved relationships among these taxa. Comparing topologies and applying methods for estimating phylogenetic networks, we find strong support for numerous instances of hybridization throughout the history of the group. We suggest that these processes may be more common than previously thought across the phylogeography-phylogenetics continuum, and that while the search for species boundaries in *Desmognathus* may not be in vain, it will be complicated by factors such as crypsis, parallelism, and gene-flow.

1. Introduction

The phylogeography-phylogenetics continuum encompasses a number of related processes that complicate efforts to reconstruct evolutionary histories and species limits (Edwards et al., 2016). Nuclear and mitochondrial trees may differ strongly, and gene genealogies may be discordant from one another (Leaché, 2009; Toews and Brelsford, 2012; Bonnet et al., 2017), with horizontal gene-transfer, incomplete lineage sorting, and paralogy affecting these patterns in myriad ways (Holder et al., 2001; Edwards, 2009). The confluence of these processes may drive conflicting signals across gene trees, species trees, networks, and morphological species boundaries. Hybridization is increasingly recognized as a common process even among distantly related lineages, with confounding effects on phylogenetic inference and difficulties for detecting reticulation (see reviews in Burbrink and Gehara, 2018; Near, 2019).

This scenario is observed in Dusky Salamanders (*Desmognathus*), for which multiple data sources have provided confusing and contradictory results over time (see, for instance, Karlin and Guttman, 1981; Titus and Larson, 1996; Rissler and Taylor, 2003; Jackson, 2005; Kozak et al., 2005; Tilley et al., 2013; Jones and Weisrock, 2018). Dusky salamanders are among the most conspicuous and abundant vertebrates in riparian ecosystems of the eastern United States (Conant et al., 2016). Early molecular studies indicated that patterns of genetic diversity within known populations were more complex than previously imagined (Tilley et al., 1978; Tilley and Mahoney, 1996; Mead et al., 2001; Bonett, 2002; Camp et al., 2002).

A broadly sampled mitochondrial phylogeny confirmed this suspicion across the genus (Kozak et al., 2005; KEA05 hereafter), revealing that lineage diversity was at least twice as high as previously believed. Furthermore, many species were not monophyletic, while others had haplotypes that did not match their morphological designations, and

* Corresponding author at: Department of Biological Sciences, The George Washington University, Washington, DC 20052, USA.

E-mail address: rapyron@colubroid.org (R.A. Pyron).

<https://doi.org/10.1016/j.ympev.2020.106751>

Received 8 September 2019; Received in revised form 2 January 2020; Accepted 28 January 2020

Available online 03 February 2020

1055-7903/© 2020 Elsevier Inc. All rights reserved.

subsequent studies have generally corroborated these findings (Jones et al., 2006; Beamer and Lamb, 2008; Tilley et al., 2008; Jones and Weisrock, 2018). The degree of complex genetic interaction between populations of putatively distinct taxa is great enough to have led some authors to suggest that the search for species limits may be in “vain” (Tilley et al., 2013).

Recently, Beamer (2015) and Beamer and Lamb (2020; “BL20” hereafter) took a systematic approach to describing this population-level genetic diversity using mitochondrial data. They used an “ecodrainage” approach, sampling individuals from all recognized, extant species known to occur within each sampling region, comprising the intersection of each Level IV Ecoregion (terrestrial ecological regions of the United States; Omernik and Griffith, 2014) and independent river drainage for a total of 179 localities across the range of *Desmognathus*. From 536 specimens of 21 nominal species, they described a total of 45 geographically cohesive mitochondrial subclades, representing a first-pass assessment of lineage diversity. Their results offer a promising perspective on species delimitation, define the spatial extent of candidate species, highlight the major areas of taxonomic incongruence, and suggest that the most problematic cases of recent genetic admixture are limited in geographic extent.

In contrast to previous studies relying heavily on a few mitochondrial markers, a phylogenomic approach can achieve several broad outcomes that address the limitations of traditional Sanger-based datasets (Hare, 2001; Toews and Brelsford, 2012). First, we can assess the genomic distinctiveness of mitochondrial clades to identify candidate species from the phylogeographic divisions (e.g., Pyron et al., 2016). Second, we can identify instances of admixture between lineages defined by nuclear and mitochondrial data, which we expect to be prevalent in taxa with a known and complex history of reticulation (Mallet et al., 2016). Third, we can evaluate relationships among candidate species using species trees and networks, for which discordance with mitochondrial trees may illuminate instances of horizontal gene-transfer or incomplete lineage sorting (Leaché, 2010; Ruane et al., 2014; Burbrink and Gehara, 2018).

Here, we use a densely sampled phylogenomic dataset of long, conserved anchored hybrid enrichment (AHE) loci (Lemmon et al., 2012) to assess the monophyly of candidate species, their phylogenetic relationships in a species-tree context, and genomic admixture using phylogenetic networks. We find remarkable congruence between mitochondrial and nuclear data in diagnosing at least 49 cryptic candidate species hidden within the 22–24 currently recognized, extant taxa. In contrast, topologies inferred from mitochondrial, species-tree, and network analyses all differ substantially in estimating relationships among candidate species. Most incongruence can be explained by rapid radiation and incomplete lineage sorting, recent exchange of mitochondrial haplotypes, or ancient hybridization. Future species-delimitation analyses in an integrative taxonomic framework combining both morphological and molecular data will allow us to assess the taxonomic validity of these candidate species.

2. Materials and methods

2.1. Lineage diagnosis and taxon sampling

It is still premature to employ a fully operationalized species concept to diagnose and delimit species in *Desmognathus*; this will require additional data such as morphology and behavior, as well as reconciling 200 years of varying criteria used to describe the existing species (see discussion in BL20). However, diagnostic criteria are still needed to recognize distinct lineages as candidate species from the various datasets. We rely on genealogical exclusivity under the Phylogenetic Species Concept (see Baum and Donoghue, 1995), using the existing delimited candidate species from previous studies (e.g., KEA05; BL20) as a baseline estimate.

Using the named lineages from KEA05 and BL20 as a starting point,

we diagnose candidate species as the least inclusive genealogically exclusive, ecogeographically distinct lineage of previously described morphospecies. When all sampled populations of a morphospecies are monophyletic and do not form ecogeographic subclades, that species is congruent as currently recognized. When a morphospecies is not monophyletic, the various topologically distinct subclades are recognized and named as in KEA05. When some or all populations of a morphospecies do form a monophyletic group, but also form genealogically exclusive, ecogeographically lineages, they are recognized and named as in BL20.

Additional details for all data, analyses, and results are given in the [Supplementary Information](#) (SI). Our mitochondrial dataset may thus reveal additional new lineages (e.g., *marmoratus* H; see below), and our nuclear data may either corroborate existing mitochondrial lineages, yield new lineages that are not evidenced in the mitochondrial data (e.g., *fuscus* E), or combine multiple mitochondrial lineages into more inclusive candidate species (e.g., *conanti* B/D). We do not employ any criterion of genetic distinctiveness within genealogically exclusive lineages. Thus, some congruent morphospecies (e.g., *valentinei*) nevertheless contain deep divergences that may merit recognition as candidate species in the future (see BL20).

We sampled 1–10 individuals from the 43 candidate species inferred by BL20 from mitochondrial data. We attempted to span the geographic and morphological range of the lineages, as BL20 noted that some haplotypes occurred in morphologically and geographically distinct populations. To these 43 clades we added samples from nine additional mitochondrial candidate species that we identified in preliminary mitochondrial analyses since the publication of BL20. This sampling yields a total of 52 preliminary genealogically exclusive lineages supported by mitochondrial data (labeled as such in [Fig. 1](#)) in our analysis to be tested with respect to a phylogenomic dataset. In all analyses, we included a single individual of *Phaeognathus hubrichti* as the outgroup. We include a total of 160 *Desmognathus* and one *Phaeognathus*. The current sampling gives a baseline estimate of mitochondrial diversity at the ecodrainage level within *Desmognathus* (see KEA05 and BL20), although additional candidate species may remain to be discovered (see below).

2.2. Mitochondrial analysis

For the mitochondrial gene-tree analysis, we used a subset of the data analyzed by BL20, comprising a fragment of the genes encoding NADH dehydrogenase subunit 2 (*ND2*); the tRNAs for Tryptophan, Alanine, and Asparagine; the origin of light-strand replication; the tRNAs for Cysteine and Tyrosine; and a fragment of Cytochrome Oxidase subunit 1 (*COI*). We supplemented this with additional Sanger sequencing of the *COI* region, and with mitochondrial assemblies generated in Geneious 11.0.4 (Biomatters Ltd.) by mapping bycatch-reads from the anchored loci (see below) to a reference mitochondrial genome (AY728227; *fuscus* B) under the default settings. The sequences were aligned using MAFFT (Katoh and Standley, 2013) in Geneious under the default parameters. Protein-coding regions were translated and checked by eye to ensure proper reading-frame and absence of stop codons. The final alignment length was 1990 bp and is 79% complete (available at Dryad repository doi.org/https://doi.org/10.5061/dryad.34tmpg4g1; see Genbank accessions for *COI* in Appendix A).

Following BL20, we analyzed the matrix using Maximum Likelihood (ML), partitioned by gene and codon position, with a single partition for the non-coding regions. We used RAxML 8.2.9 (Stamatakis, 2006), which employs a GTR + Γ (GTRGAMMA) model for all partitions. We created 1000 non-parametric bootstrap replicates, from which 200 independent ML searches were started every 5th replicate. For the best-scoring tree, the proportions of the 1000 replicates including each branch were plotted as node support.

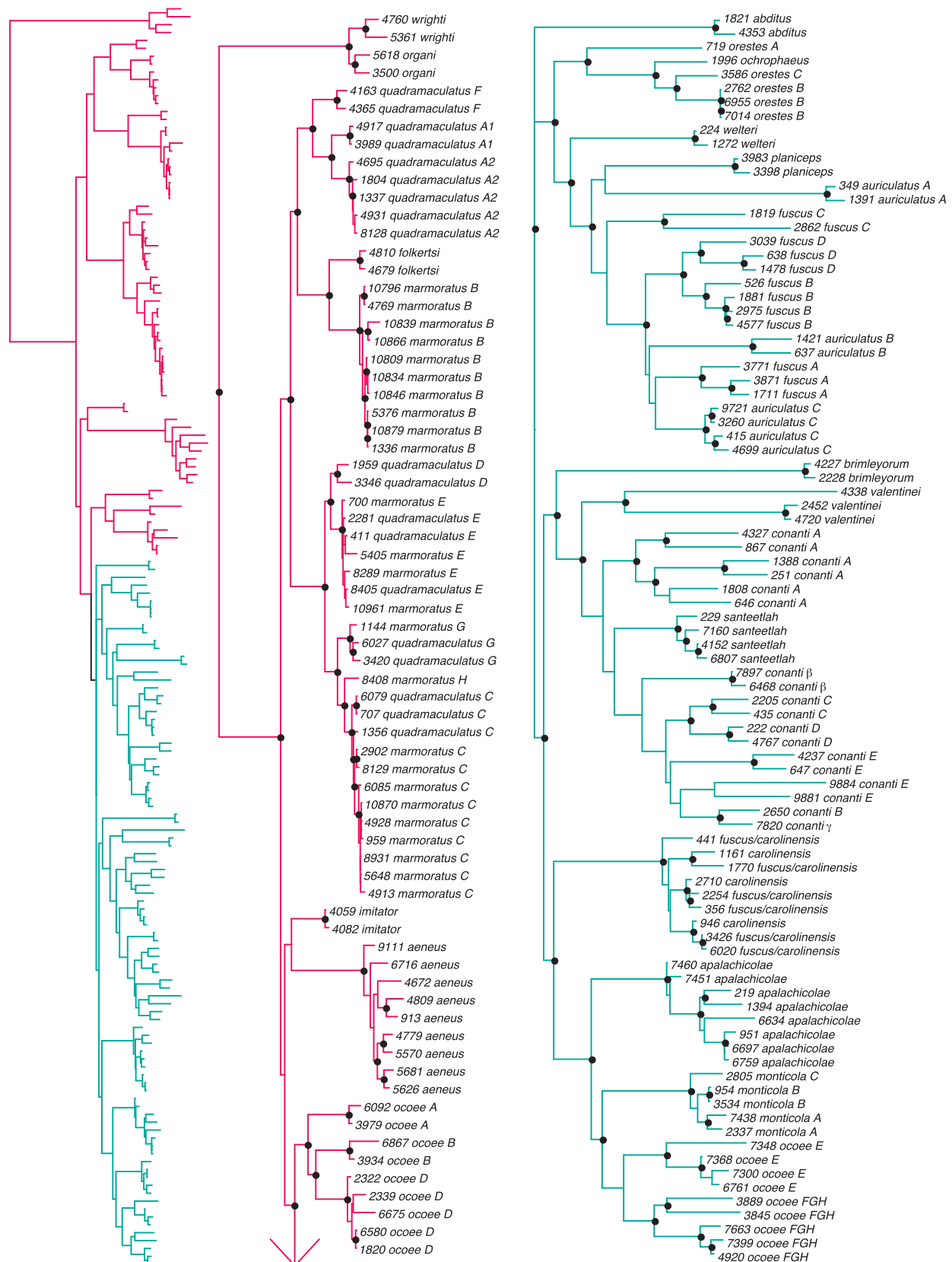


Fig. 1. Results from concatenated ML analysis of 1995 bp of mitochondrial DNA for 160 specimens (outgroup *Phaeognathus* not shown), representing 52 mitochondrial clades *sensu* Beamer and Lamb (2020), modified as noted (see Materials and Methods). Trees were estimated with 1000 non-parametric bootstrap replicates and 200 random ML searches. Black circles indicate nodes supported at BS > 70%.

2.3. Anchored hybrid enrichment

Data were collected following Lemmon et al. (2012) at the Center for Anchored Phylogenomics at Florida State University (www.anchoredphylogeny.com). Each genomic sample was sonicated to a fragment size of ~150–350 bp using a Covaris E220 Focused-ultrasonicator with Covaris microTUBES. Library preparation and indexing were performed on a Beckman-Coulter Biomek FXp liquid-handling robot following a protocol modified from Meyer and Kircher (2010), but with a size-selection step after blunt-end repair using SPRIselect beads (Beckman-Coulter Inc.; 0.9x ratio of bead to sample volume). After pooling 12–16 indexed samples at equal quantities, enrichments were performed on each multi-sample pool using an Agilent Custom SureSelect kit (Agilent Technologies), developed for amphibians by Hime et al. (2020). To develop this resource, Hime et al. (2020) targeted ~366 anchor loci (averaging 1090 bp each) derived from the vertebrate AHE set of Lemmon et al. (2012) and mined genomic resources of five salamander species (*Desmognathus fuscus*, *Ambystoma mexicanum*, *Notophthalmus viridescens*, *Cryptobranchus alleganiensis*, and *Ensatina eschscholtzii*) in addition to seven frog and one caecilian species (see Hime et al., 2020). After enrichment, the 10 enrichment pools were combined in equal quantities for sequencing on three PE150 Illumina HiSeq2000 lanes. Sequencing was performed in the Translational Science Laboratory in the College of Medicine at Florida State University.

To generate alignments, raw reads passing the Illumina CASAVA (v1.8) high chastity filter were demultiplexed based on 8 bp indices. Overlapping read pairs were subsequently merged using the methods of Rokyta et al. (2012). Downstream read processing followed the methods of Prum et al. (2015) and Hamilton et al. (2016), but with *Desmognathus fuscus*, *Cryptobranchus alleganiensis*, and *Ambystoma mexicanum* as references during the quasi-de novo assembly. Orthology was assessed using a neighbor-joining clustering approach that utilizes a common-k-mer distance matrix (see Hamilton et al., 2016 for details). Sets of orthologous sequences were then aligned using MAFFT v7.023b (Katoh and Standley, 2013), and trimmed/masked to remove problematic regions. Alleles were not phased; heterozygosity was instead treated with ambiguity codes within individuals. The final dataset consisted of 381 loci, for a total of 554,321 bp, and was 89.2% complete for 161 individuals.

2.4. Concatenated and gene-tree analysis

Similar to the mitochondrial analysis, we first analyzed the nuclear matrix using Maximum Likelihood (ML), partitioned by locus. We used RAxMLv8.2.9 (Stamatakis, 2006), which employs a separate GTRGAMMA model for each partition. We created 1000 non-parametric bootstrap replicates, from which 200 independent ML searches were started every 5th replicate. For the best-scoring tree, the 1000 optimizations were plotted as node support. Secondary to this concatenated analysis, we analyzed each of the 381 genes separately, for the purpose of downstream species-tree analysis. We used the same ML analysis parameters described above, though with only 100 bootstrap replicates per gene. We expect low resolution for most individual gene-trees (Kuo and Avise, 2005), but these can be combined for a rich source of information regarding gene-tree discordance to infer a well-resolved species tree (Edwards, 2009). Incorporating gene-tree error has been shown to contribute significantly to the resolution of the Tree of Life, though some nodes may remain unresolved (Pyron et al., 2014).

2.5. Species-tree analyses

Multiple species-tree methods exist based on the multi-species coalescent (Liu et al., 2009), many of which operate on summaries of gene-tree estimates and show strong performance across a range of empirical conditions (Mirarab et al., 2016). Of these, ASTRID is faster and at least as accurate as competitors such as NJst and ASTRAL under

most empirical conditions, particularly with larger datasets (hundreds of taxa and genes) and low missing data (Vachaspati and Warnow, 2015). Thus, we preferred ASTRID for our estimate of the multi-locus species tree. We first estimated the species tree using ASTRID 1.4, which produces a matrix of gene-tree distances, and uses neighbor-joining to produce the species tree.

After estimating a topology using ASTRID, we followed the recommendations of the program developers to use ASTRAL-II 5.0.1 (Mirarab et al., 2014) to estimate branch lengths in coalescent units on the ASTRID topology. To estimate support, we used the recommended coalescent-based measure that assesses the local posterior probability for each branch based on the quartet of surrounding branches (Sayari and Mirarab, 2016). The same algorithm is then also used to estimate branch lengths in coalescent units. We treat this as our best estimate of the species tree. As an additional assessment incorporating gene-tree uncertainty, we used the 381 sets of 112 bootstrap replicates (as 16 computing nodes each performed 7 replicates per gene) as the input set for ASTRID 1.4 and ASTRAL-II 5.0.1. We then estimated 112 ASTRID species trees (one from each set of bootstrap replicates) and took the strict consensus. For branch-length and support values, we used the coalescent-based support measure from the 381 ML gene-trees, mapped on this consensus topology of the multi-locus bootstrap replicates.

2.6. Network analyses

A bifurcating tree is increasingly recognized as a poor descriptor of many parts of the diagrammatic history of life, where horizontal connections are present among multiple branches in reconstructed phylogenies (see review in Burbrink and Gehara, 2018; MacGuigan and Near, 2019). Such processes may have severe impacts on phylogenetic estimation (Solís-Lemus et al., 2016). To reflect these connections, numerous methods exist for estimating phylogenetic networks (Huson and Bryant, 2005; Than et al., 2008; Pickrell and Pritchard, 2012), though most are limited by computational intensity to a few taxa and reticulations. We investigated the presence and impact of these processes using two distinct approaches.

First, we used the program TreeMix 1.13 (Pickrell and Pritchard, 2012) to estimate a topology while modeling admixture as migration events between branches. This is a relatively simple algorithm that uses SNP data converted to bi-allelic frequency distributions within populations (see SI for details) and estimates population splits (and thus a species tree) for those data. Subsequently, migrations are added to the population graph, increasing the goodness-of-fit of the model to the observed frequency data under a Gaussian approximation of drift. For the sliding window of linkage disequilibrium, we tested $k = 500$ and 1000 as recommended by the authors. As reticulations drastically increase graph complexity, we limited all network analyses to a maximum of $\sqrt{n}$, giving a maximum of 7 for the 49-species analysis.

Second, a recently developed method (SNaQ; Solís-Lemus and Ané, 2016) implemented in the Julia package PhyloNetworks (Solís-Lemus et al., 2017) scales up to at least 24 taxa and 5 reticulation events in a tractable amount of computational time (i.e., days to weeks). The method uses a maximum pseudolikelihood estimator applied to the quartet concordance factors (CF) of 4-taxon trees under the coalescent model, incorporating incomplete lineage sorting and extended to include reticulation events. The observed CF from the estimated gene trees is then used to estimate a semi-directed species network with estimated reticulation events and γ -values indicating the proportion of ancestral contribution to the hybrid genome.

We used PhyloNetworks to read in the 381 RAxML gene-trees and to estimate observed concordance factors, with all individuals per clade mapped as alleles to species. Given the computational intractability of the full dataset, we broke the tree into three sections for analysis as recommended by the developers of the algorithm (C. Ané, pers. comm.). We note that this strategy may potentially obscure reticulation events between lineages not included in all three analyses; these may be

revealed in future studies using more computationally advanced methods. We used the ASTRAL topology trimmed to include the focal clades of interest as the starting topology, and tested values for h (number of reticulations) from 0 to 5, assessing maximum support using a slope heuristic for the increase in likelihood plotted against h (Solís-Lemus and Ané, 2016). We ran 10 independent runs per h -value to ensure convergence on a global optimum.

2.7. Reanalysis of Jones and Weisrock (2018)

Jones and Weisrock (2018; “JW18” hereafter) presented a phylogenomic dataset of *quadramaculatus* and *marmoratus* from throughout their range. They revealed non-monophyly of the group (see Jackson, 2005), as lineages from east of the French Broad River were more closely related to *carolinensis*, the only other species included. More astonishingly, they reported that morphological *quadramaculatus* and *marmoratus* did not form monophyletic groups within their two main clades but were instead interdigitated. Thus, the Nantahala (western) and Pisgah (eastern) clades were said to be two geographically and ecologically distinct, cryptic species, each exhibiting polymorphism for the blackbelly or shovelnose phenotype.

This result is at odds with the mitochondrial datasets of KEA05, Jones et al. (2006), and BL20, and the nuclear and mitochondrial dataset of Jackson (2005), all of whom sampled the same areas. All of those studies showed that morphological *marmoratus* and *quadramaculatus* share mitochondrial haplotypes east of the French Broad River, mirrored by the nuclear dataset of JW18 in the Pisgah clade. However, all three studies also showed that morphological *marmoratus* from western North Carolina and northern Georgia (west of the French Broad River, in the range of JW18’s Nantahala clade) formed two genealogically exclusive, monophyletic lineages on either side of the Eastern Continental Divide, and that these lineages did not share mitochondrial haplotypes with sympatric morphological *quadramaculatus*.

Furthermore, all previous studies showed that the western North Carolina lineage of morphological *marmoratus* (called *marmoratus* C in KEA05 and BL20) was more closely related to morphological *marmoratus* and *quadramaculatus* from east of the French Broad River (i.e., *marmoratus* C is part of the Pisgah clade) than it is to morphological *marmoratus* from northern Georgia (*marmoratus* B in KEA05 and BL20; part of the Nantahala clade). Thus, mitochondrial data show that the Pisgah and Nantahala clades are not geographically exclusive, in contrast to the nuclear data presented by JW18, and also contain no evidence of hybridization between the two species west of the French Broad River, as shown by JW18. Therefore, we re-analyzed JW18’s data to determine the root cause of this discordance.

Full details are given in the SI; in short, we mapped their raw ddRAD reads to the AHE loci used here, to locate overlap where AHE loci also contained restriction sites. We identified 36 AHE loci that contained restriction sites (with ~140 bp flanking regions), for a total overlap ranging from 70 to 827 bp of ddRAD loci per AHE locus. We trimmed and concatenated these data for a total of 219 terminals (161 AHE + 58 ddRAD), and up to 6653 bp of sequence data per terminal. We analyzed this partitioned, concatenated dataset using ML in IQ-TREE (Nguyen et al., 2015) due to the lower computational burden compared to the RAxML analyses of the full dataset. We then compared this topology with our main analyses to determine the amount of congruence between our results and those of JW18, and the likely sources of any discordance.

3. Results

3.1. Phylogenetic inference

The mitochondrial (Fig. 1), concatenated nuclear (Fig. 2), and species-tree nuclear (Fig. 3) analyses all estimated strongly supported topologies with the major candidate species evident. As expected, the

mitochondrial analysis yielded strong congruence with previous analyses using similar datasets (e.g. KEA05; BL20). In contrast, the multi-locus nuclear dataset shows robust discordance with the long-established mitochondrial patterns (Fig. 4). While portions of the trees are congruent or similar with respect to some major groupings, there are significant topological disagreements regarding the placement of species such as *abditus*, *carolinensis*, *fuscus*, *marmoratus*, *ocoe*, and *quadramaculatus*. In particular, the *folkertsi*, *marmoratus*, and *quadramaculatus* lineages do not form a monophyletic group in any nuclear analyses, as in JW18. However, despite topological discordance regarding the placement of various clades, the mitochondrial and nuclear datasets are in fact highly concordant with respect to the identity and cohesiveness of numerous candidate species.

3.2. Candidate species

Most mitochondrial subclades from BL20 received 100% bootstrap support in analyses of the nuclear datasets and in agreement with our reduced mitochondrial dataset (Figs. 1–4). Some individual clades interdigitated and were combined into larger candidate species, while a few new nuclear clades were revealed by the anchored data, for a total of 49 candidate species from our original sampling of 52 mitochondrial subclades.

We find that 14 currently recognized species are either equivalent to a candidate species or comprise a monophyletic group thereof. These are *abditus*, *aeneus*, *apalachicola* (sensu Beamer and Lamb, 2008), *brimleyorum*, *folkertsi*, *imitator*, *monticola*, *ochrophaeus* (sensu Tilley and Mahoney, 1996), *organi*, *planiceps*, *santeetlah*, *valentinei*, *welteri*, and *wrighti*. Thus, 14 of the 22–24 currently recognized species appear to be robust and validly defined at present. This is not to say that these species do not contain phylogeographic divisions that could also merit recognition as full, distinct species. In particular, *aeneus* and *valentinei* have a large amount of intraspecific genetic diversity (Figs. 1–4), and *monticola* comprises three distinct candidate species A, B, and C (Figs. 1–4; see BL20). Finally, some of these species (*folkertsi*, *ochrophaeus*, and *santeetlah*) are nested within *quadramaculatus*/*marmoratus* (part; see Jones and Weisrock, 2018), *orestes*, and *conanti* (part), respectively.

There are also 16 instances in which mitochondrial subclades of currently recognized species do not form a monophyletic group, and thus represent independent candidate species. These are *auriculatus* A, *auriculatus* B, *conanti* β, *conanti* C, *conanti* E (weakly paraphyletic in the mitochondrial dataset, but strongly supported in the nuclear analyses), *conanti* γ, *fuscus* A, *fuscus* B, *fuscus* D, *marmoratus* B, *ocoe* A, *ocoe* B, *ocoe* D, *ocoe* E, *quadramaculatus* A1 and A2, and *quadramaculatus* F. We note that the β and γ lineages of Tilley et al. (2013) from the Blue Ridge represent two distinct candidate species, in contrast to Tilley et al. (2013), who regarded these as introgressed populations representing “failed” species. They instead have unique mitochondrial haplotypes and distinct nuclear genomes and appear to represent locally endemic species that are simply morphologically congruent with *conanti* in the broad sense.

Additionally, there are five major instances in which a single nuclear lineage comprises two or more mitochondrial lineages that are interdigitated (not genealogically exclusive) in the nuclear analyses. These are *conanti* B/D, *marmoratus* E/H, *ocoe* F/G/H (as noted above), *orestes* A/C, and *quadramaculatus* D/E. We consider each of these to represent candidate species, agglomerating the divergent mitochondrial haplotypes into more inclusive lineages.

Finally, there are 5 major instances of nuclear lineages sharing haplotypes among mitochondrial subclades, involving 10 total mitochondrial candidate species. These are: *fuscus* C (contains all *fuscus*/*carolinensis*), *fuscus* E (contains some haplotypes from *auriculatus* C), *conanti* F (contains some haplotypes from *conanti* A), and *marmoratus*/*quadramaculatus* C and G, each of which are distinct, independent candidate species in the nuclear analyses. Thus, shared haplotypes

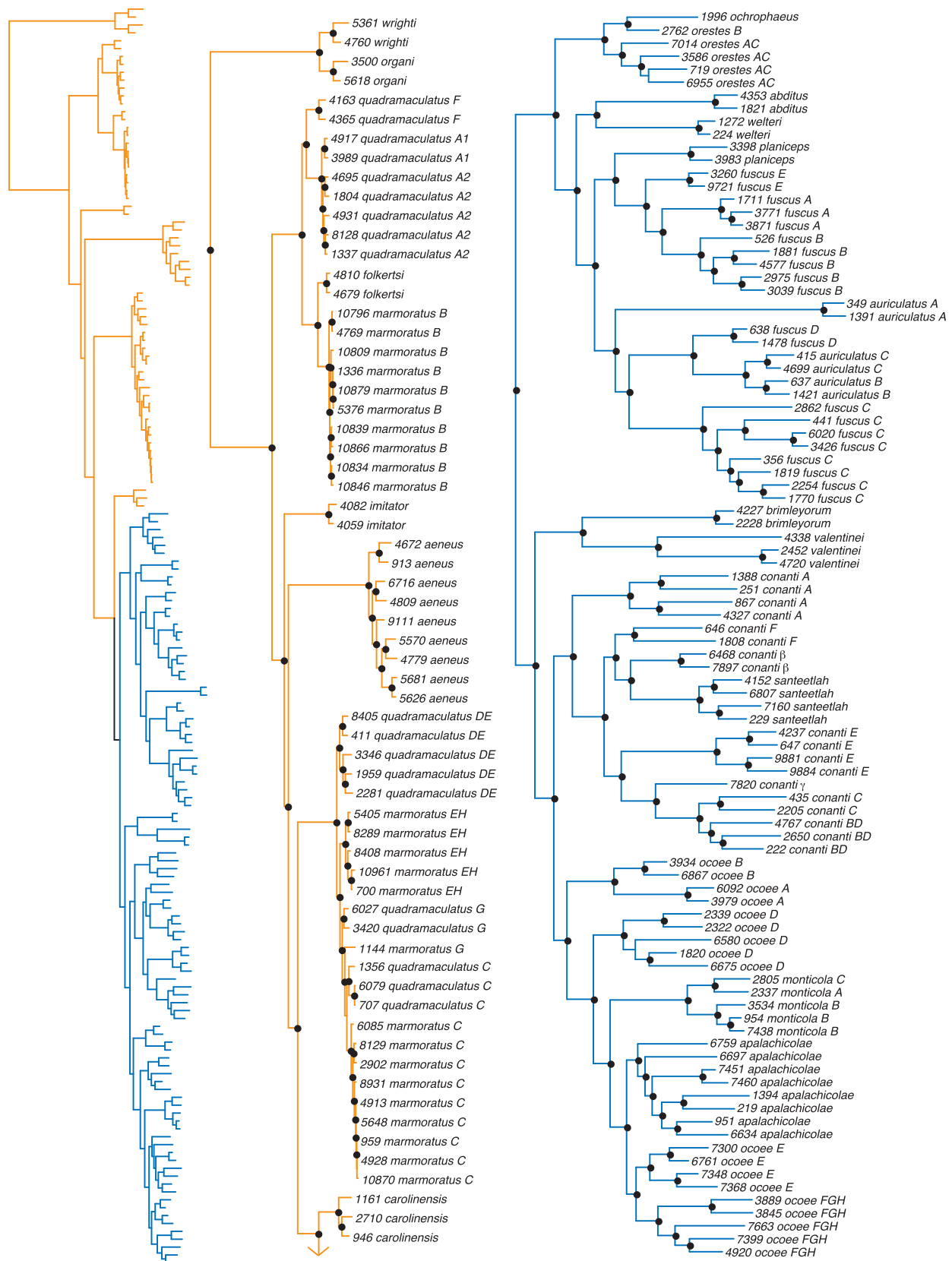


Fig. 2. Results from concatenated ML analysis of 554,321 bp from 381 loci for 160 specimens (outgroup *Phaeognathus* not shown), representing 49 nuclear clades as described above (see Materials and Methods). Trees were estimated with 1000 non-parametric bootstrap replicates and 200 random ML searches. Black circles indicate nodes supported at BS > 70%.

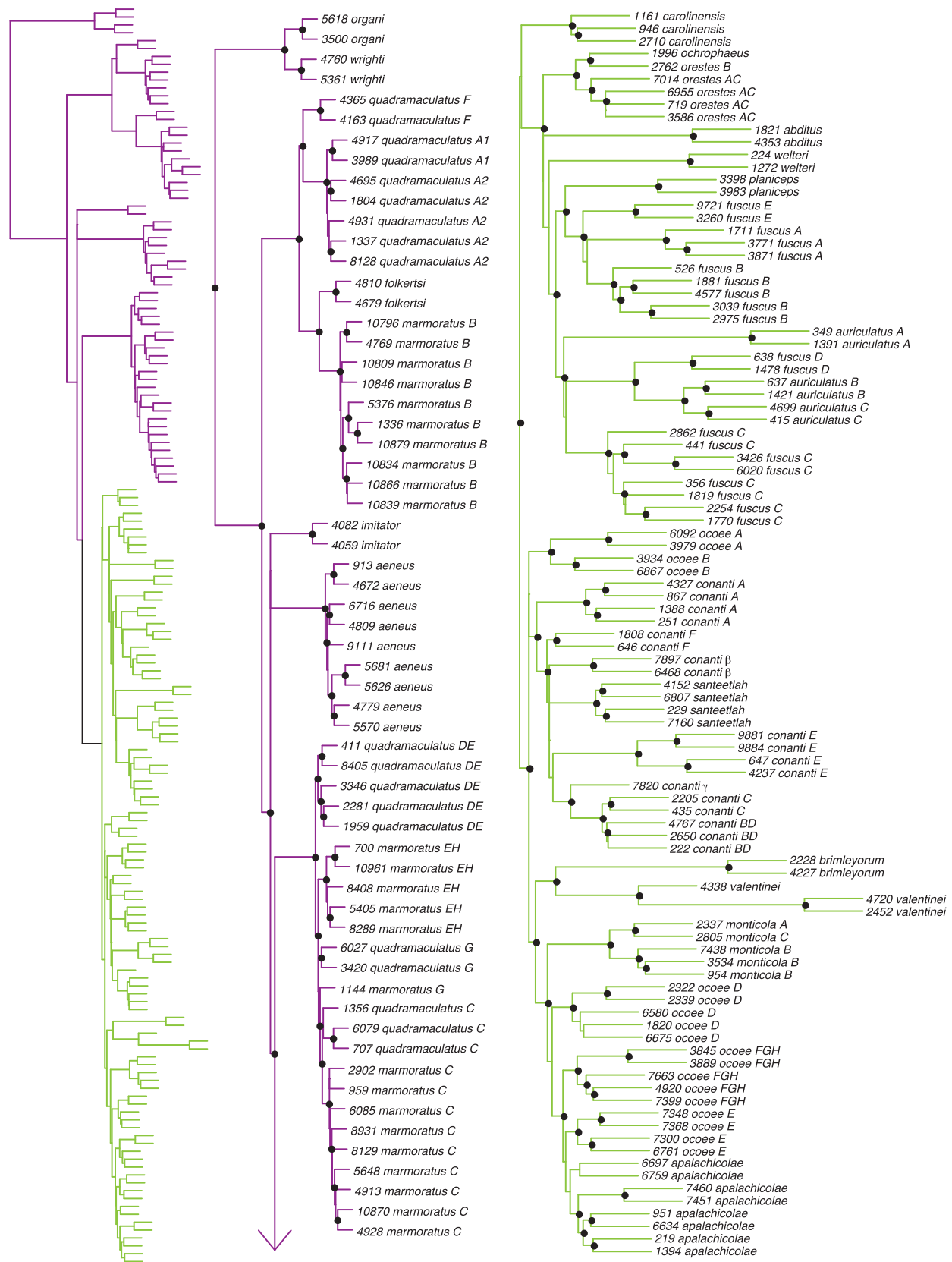
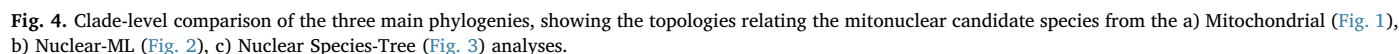


Fig. 3. Results from species-tree analysis of 381 gene trees from 160 specimens (outgroup *Phaeognathus* not shown), representing 49 nuclear clades as described above (see Materials and Methods). The topology was estimated using the ASTRID algorithm (Vachaspati and Warnow, 2015), while branch lengths and local posterior probabilities (Mirarab et al., 2014) were estimated using the ASTRAL algorithm (Sayari and Mirarab, 2016). Black circles indicate local posterior probabilities > 95%.



presumably indicative of recent introgression are observed between both closely related as well as relatively distant lineages, some of which are not currently in close geographic proximity.

3.3. Topological discordance

Strongly supported differences between the mitochondrial and nuclear data (Figs. 1–4) are related to three major patterns. The first are apparent instances of recent admixture resulting in similar mitochondrial haplotypes shared between nuclear-based lineages. The second are weakly supported disagreements about the relationships among lineages within well-defined species groups as defined below, which we do not address in-depth here. The third are major topological discordances related to the placement and monophyly of the blackbelly (*quadramaculatus*) and shovelnose (*marmoratus*) lineages, as well as the placement of *carolinensis*, *abditus*, and *ocoee* A, B, and D. Placement of these lineages varies sharply between the mitochondrial and nuclear trees in a way that defies simple explanations of rapid radiation and incomplete lineage sorting, and is likely due to deep-time reticulation.

As in KEA05 and BL20, all *folkerti*, *marmoratus*, and *quadramaculatus* lineages (the blackbelly and shovelnoses) form a monophyletic group in the mitochondrial analyses, though the latter two species are paraphyletic (Fig. 1). In all nuclear analyses (concatenated, species-tree, and network; Figs. 2, 3), two separate clades are recovered: one containing *folkerti*, *marmoratus* B, *quadramaculatus* F, and *quadramaculatus* A1 and A2 (the Nantahala clade of JW18); and the other comprising *quadramaculatus* D/E, *marmoratus* E/H, *quadramaculatus* C, *marmoratus* C, *quadramaculatus* G, and *marmoratus* G (the Pisgah clade). These two clades do not form a monophyletic group (Fig. 4). We suggest that this must be explained by a deep-time reticulation, though no support is found for this hypothesis in our network analyses (see below).

Curiously, relationships within each of the two groups are almost identical between the mitochondrial and nuclear datasets. Thus, some populations of the *marmoratus* and *quadramaculatus* C, D, and E lineages share related mitochondrial haplotypes in sympatry, as in KEA05, Jackson (2005), Jones et al. (2006), BL20, and our mitochondrial dataset (Fig. 1), but this is not reflected by genealogical interdigitation in our nuclear dataset, in contrast to JW18.

Second, in the mitochondrial dataset, *carolinensis* forms a clade with *fuscus*/*carolinensis*, which is more closely related to the *monticola* and *conanti* lineages than to the *fuscus* lineages (Fig. 1). In the concatenated nuclear dataset, *fuscus* C is distinct from *carolinensis*, which is the sister taxon to a lineage containing all *Desmognathus* except *organi*, *wrighti*, *folkerti*, *marmoratus*, *quadramaculatus*, *imitator*, and *aeneus* (Fig. 2). In the species-tree analysis, *carolinensis* is more closely related to the *fuscus* lineages than it is to the *conanti* or *monticola* lineages (Fig. 3).

Third, *abditus* is the sister taxon to a group containing all *conanti*, *fuscus*, and *monticola* (and other) lineages in the mitochondrial analysis (Fig. 1). In the nuclear analyses (Figs. 2, 3), it forms a clade with *auriculatus*, *fuscus*, *planiceps*, and *welteri*. However, this relationship is weakly supported.

Fourth, *ocoee* A, B, and D form a clade in the mitochondrial analysis that is strongly supported as the sister taxon to all *Desmognathus* except *organi*, *wrighti*, *folkerti*, *marmoratus*, *quadramaculatus*, *imitator*, and *aeneus* (Fig. 1). In the nuclear analyses (Figs. 2, 3), *ocoee* D is distinct from *ocoee* A + B, and each of those clades is more closely related to *monticola* and *conanti* lineages than they are to *fuscus*. This is a far more deeply nested position than in the mitochondrial analysis (Figs. 1–4). We find that this is due to a deep-time reticulation (see below).

3.4. Network analyses

The Treemix analysis with $k = 1000$ and $m = 3$ provided the best goodness-of-fit with the fewest parameters (Fig. 5), and accounts for several of the major genealogical discordances noted above. The

topology estimated for the population graph is highly similar to the ASTRAL tree, with three migration events that explain (at least partially) the mito-nuclear discordance regarding the placement of *carolinensis*, *fuscus* C, and *ocoee* A, B, and D (Fig. 4). Essentially the same results were estimated in the $k = 500/m = 2$ –4 analyses (see SI text for further details), but the $k = 1000$ analysis implies a larger window for linkage disequilibrium, as recommended by the developers (Pickrell and Pritchard, 2012), and better accounts for residual variation.

The analyses supported a scenario where an ancestral lineage (represented by the internal red arrow) hybridized with *ocoee* A, which occupies the same nested position with the *conanti* lineages as in the other nuclear datasets. That ancestral lineage then either went extinct or was completely absorbed by *ocoee* A. The second and third migration events are from the *carolinensis* lineage (again occupying roughly the same position as in the species tree, as the sister taxon to the *fuscus* lineages and relatives) to *conanti* F, and from *conanti* F to *fuscus* C. While the involvement of *conanti* F in the *fuscus*/*carolinensis* hybridization is not immediately evident from our mitochondrial sampling, *conanti* F is a newly discovered nuclear clade here and shares mitochondrial haplotypes with *conanti* A, having thus been involved in at least some hybridization events or rapid radiations. This population has been troublesome for species assignment (e.g., Pope, 1924), with some workers recognizing it as *carolinensis* and others assigning it to *fuscus*. This intermediate morphology may be due to past introgression between *fuscus* C, *carolinensis*, and *conanti* F.

The SNaQ analyses proved intractable for the full dataset, so we analyzed three major sets of lineages separately. The estimated topology of each subclade is extremely similar to the ASTRID topology, and each analysis supported a single reticulation. In the first (Fig. 6a), *fuscus* and relatives, the lineage comprising *auriculatus* A + *fuscus* C is the sister taxon to the clade comprising *fuscus* D and *auriculatus* B + C, but is estimated to have 32% ancestry from *carolinensis*, resolving one of the major instances of obvious mitochondrial introgression in this group (see BL20). The absence of *carolinensis* mitochondria in *auriculatus* A is puzzling if this result is correct, but *auriculatus* A has undergone drastic declines since the 1970's (Graham et al., 2010), and thus our samples may not reflect its full historical genetic makeup.

In the clade comprising *conanti* and relatives (Fig. 6b), the lineage comprising *conanti* C + BD is the sister taxon to *conanti* γ, but derives 21% of its ancestry from *ocoee* FGH, again partially explaining one of the major instances of discordance and leaky species boundaries in *Desmognathus* (see Tilley et al., 2013). Finally, among the root lineages (Fig. 6c), the lineage comprising *imitator* + *aeneus* derives 26% of its ancestry from the lineage subtending *fuscus* and relatives. This resolves the topological discordance and weak support for the placement of these species among the various mitochondrial and nuclear analyses (Figs. 1–4).

3.5. Reanalysis of JW18

Our combined reanalysis of JW18's ddRAD data with our AHE loci shows strong overall congruence (see SI Fig. S2), suggesting that portions of the ddRAD data contain similar phylogenetic signal to the AHE loci, and that JW18's topological results are mirrored by our own as noted above. Very similar results from a combined nuclear and mitochondrial dataset were also reported by Jackson (2005). In particular, the combined ddRAD + AHE dataset is congruent in recovering non-monophyly of the various *marmoratus* and *quadramaculatus* lineages; their Nantahala clade represents *marmoratus* B, *folkerti*, and *quadramaculatus* A and F, while their Pisgah clade represents the *marmoratus* and *quadramaculatus* C, D, E, G, and H lineages. Additionally, one of their Nantahala "*quadramaculatus*" specimens is actually a *folkerti* that was mislabeled (site 25; J. Wooten, pers. comm.); this species was not addressed by JW18.

In contrast to their results, the distribution of the Pisgah-associated *marmoratus* C mitochondrial lineage recovered by Jackson (2005),

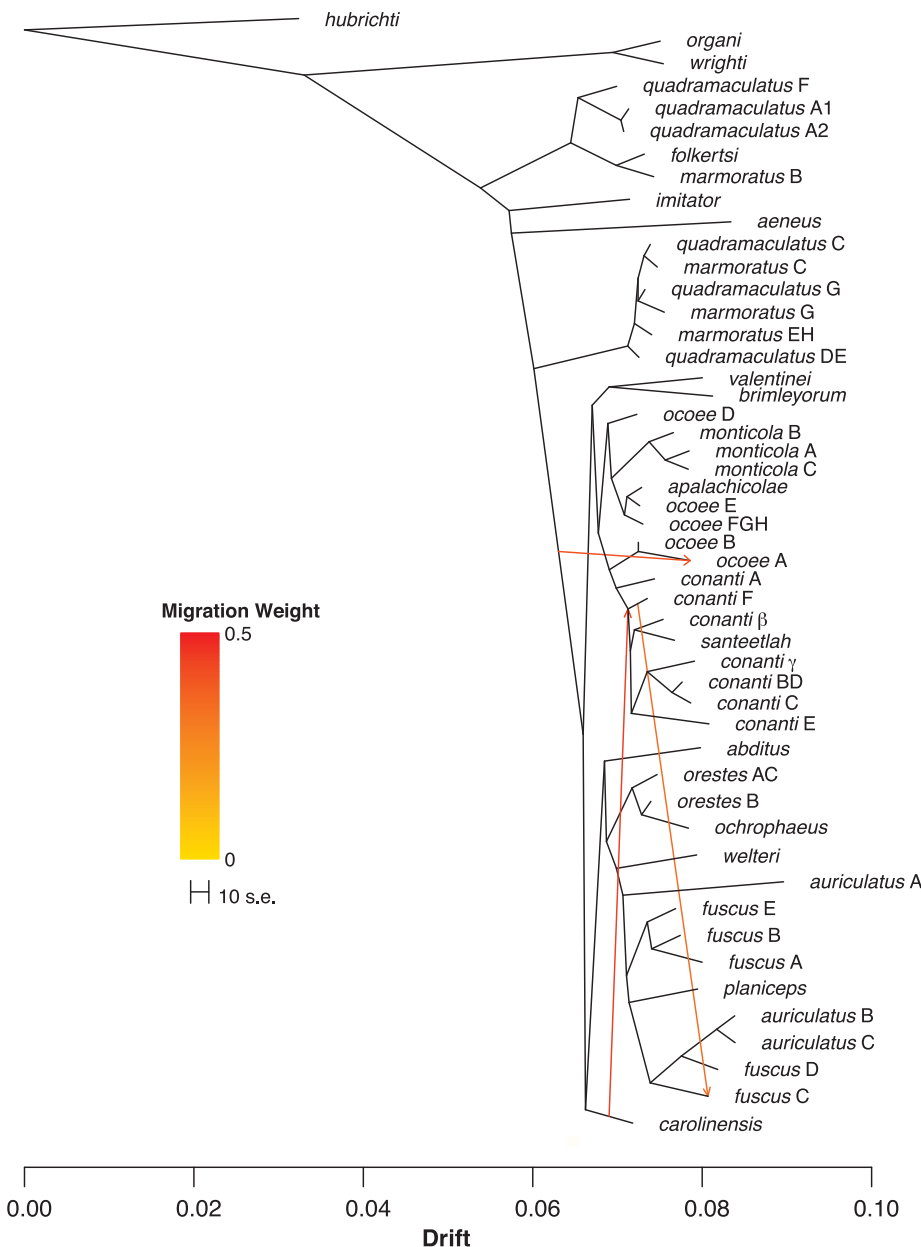


Fig. 5. TreeMix analysis with $k = 1000$ and $m = 3$, along with residual plot (inset), showing general concordance with the species-tree topology, but with support for migrations between an ancestral lineage and *ocoee* A, and between *carolinensis*, *conanti* F, and *fuscus* C; patterns that are observed (at least in part) in the discordance between the mitochondrial and nuclear datasets (Figs. 1–4). The migration event from an internal branch to *ocoee* A represents a putative ancestral species diverging from that topological position and persisting long enough to hybridize with *ocoee* A. Subsequently, it either went extinct or was completely subsumed by extant populations of *ocoee* A.

KEA05, Jones et al. (2006), BL20, and here is corroborated by this reanalysis, and extends deep into the geographic range of the Nantahala lineage (Fig. S3), but these populations were apparently not sampled by JW18. Instead, our 9 samples of morphological *marmoratus* from this region (North Carolina west of the Asheville Basin) are all *marmoratus* C, characterized by unique mitochondrial and nuclear genomic constitutions distinct from co-distributed *quadramaculatus* A (SI Figs. 2–3).

In contrast, neither Jackson (2005), KEA05, Jones et al. (2006), BL20, nor the current study sampled any morphological *marmoratus* with *quadramaculatus* A mitochondrial or nuclear genomes, the pattern reported by JW18. Their putative samples of western NC *marmoratus* all cluster with *quadramaculatus* A in the reanalysis (Fig. S2). For their results to be valid, we would have to posit that JW18 sampled populations of morphological *marmoratus* C from west of the French Broad River with *quadramaculatus* A genomes that no other researchers including Jackson (2005), KEA05, Jones et al. (2006), BL20, or the present authors sampled, and that JW18 simultaneously failed to sample *marmoratus* C at any of these same sites where they are known to occur. Instead, we suggest that JW18 did not actually sample any

morphological *marmoratus* in the range of the Nantahala clade, but instead misidentified immature or aberrant *quadramaculatus* A at those sites. This is bolstered by the fact that two of their putative *carolinensis* are actually misidentified *monticola* B (see SI Fig. S2). Unfortunately, JW18 report not taking any voucher specimens or photographs, making it impossible to confirm or deny this explanation.

4. Discussion

Overall, we find a high degree of concordance between the mitochondrial and nuclear datasets and recover at least 49 candidate species in *Desmognathus* (Fig. 4). Previous work (Anderson and Tilley, 2003; Tilley et al., 2008; Kratovil, 2017) suggests that there may be more. The topologies estimated from the concatenated mitochondrial data, concatenated nuclear data, and species-tree analysis of the nuclear data all differ strongly, revealing the historical signatures of admixture, incomplete lineage sorting, and rapid radiation, suggesting further avenues for targeted research (e.g. *fuscus*, *conanti*, and relatives). All of these processes seemingly trace to the permeable nature of species

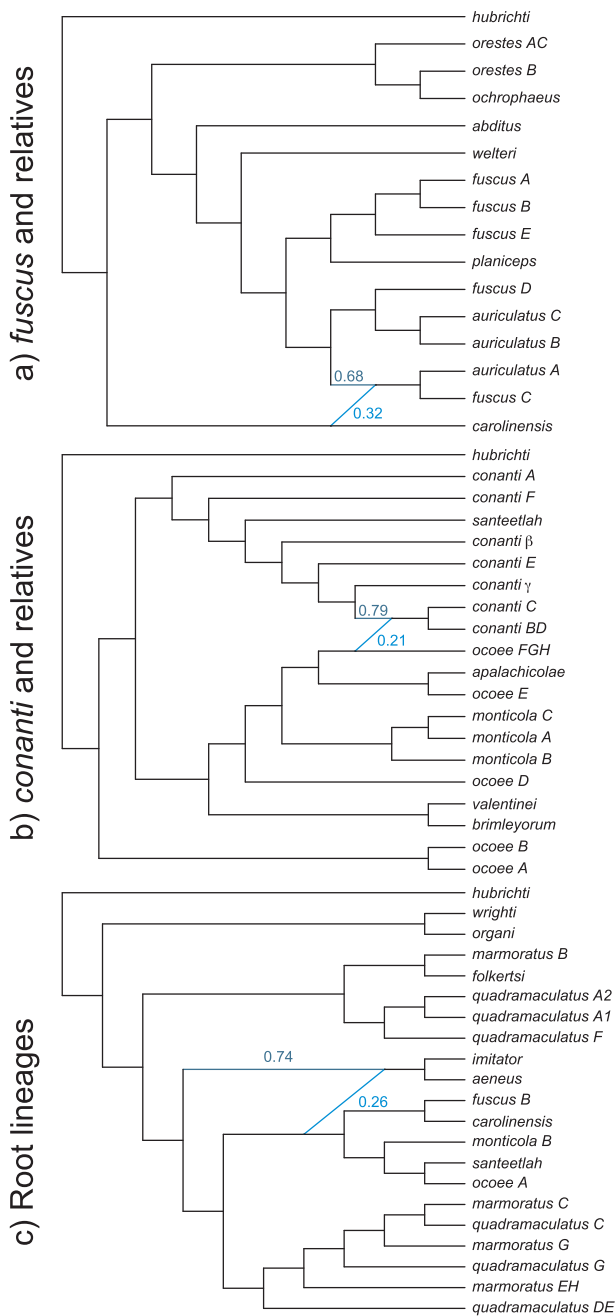


Fig. 6. SNaQ analyses of three subclades: *fuscus* and relatives (a), *conanti* and relatives (b), and basal-most lineages (c). Reticulations among lineages here reveal ancestral hybridization events not captured by the TreeMix analyses (Fig. 5), further resolving some observed mitonuclear discordance in extant species (Figs. 1–4).

boundaries (Harrison and Larson, 2014) and result in a variety of patterns across the phylogeography-phylogenetics continuum (Edwards et al., 2016) from the influence of both deep-time reticulation and recent introgression (Martin et al., 2013; Burbrink and Gehara, 2018).

4.1. Candidate species

Promisingly, at least 14 species are coherent and well defined as currently recognized (see Fig. 5). In contrast, a number of species including *auriculatus*, *conanti*, *fuscus*, *marmoratus*, *ocoe*, and *quadramaculatus* represent complicated and as-yet unresolved species complexes, each containing multiple topologically disparate species-level

lineages. Nevertheless, our results show promise regarding the potential for resolving species limits across the genus in future analyses, as nearly all candidate species defined here represent geographically and genetically divergent populations that are reciprocally monophyletic on molecular phylogenetic trees. There is also a moderate degree of concordance in terms of estimated topological relationships between these clades within and among analyses of both nuclear and mitochondrial datasets (Fig. 4).

However, as noted by numerous previous authors (e.g., Tilley et al., 2013; Jones and Weisrock, 2018), these relationships are obscured in many parts of the tree by mitonuclear discordance and deep-time reticulation. Many of these instances are resolved by our network analyses, providing at least partial explanations for genealogical discordance regarding the placement of *aeneus*, *carolinensis*, *conanti* C, B/D, F, and A, *fuscus* C, *imitator*, and *ocoe* A, B, and D. In contrast, none of the network analyses provided satisfactory resolutions for the mitonuclear discordances observed in *marmoratus* and *quadramaculatus* (see Jackson, 2005; JW18). Nor do they explain other observed instances of mitochondrial genome-capture in clades such as *fuscus* E, of which the populations sampled here have mitochondrial haplotypes from *auriculatus* C, a lineage that occurs over 200 km away (see BL20).

We suspect that this is due to a combination of issues, such as limited sampling, computational inadequacy, and overwhelming (or curiously absent) signal. To the former point, our tree of only 50 species is estimated to contain at least 6 reticulations, straining our ability to compute and interpret them simultaneously and coherently (Figs. 4–6; see Holder et al., 2001). To the latter, mitochondrial genome capture may occur with little or no nuclear introgression (Good et al., 2015), and selection pressure on hybrids is poorly understood (Vijay et al., 2016). We anticipate that future datasets focusing on particular lineages while sampling more individuals will be able to parse out patterns in a more tractable manner. Regardless, the results presented here offer a first-pass assessment of the broad-scale history of reticulation in this group.

4.2. Species limits

Tilley et al. (2013) offered one of the first population-level datasets for *Desmognathus* sampling multiple species across a wide geographic area, and we corroborate their findings of extreme morphological cryptic and complex genealogical histories that defy easy classification. In particular, they noted the extensive gene flow between distantly related species such as *carolinensis* and *conanti*, not just close relatives, which we here illustrate is a common range- and genus-wide feature of *Desmognathus* species. Our data offer a promising insight: that these patterns can be unified under a framework wherein the diversity of *Desmognathus* is divided into ~49 candidate species that are geographically and genetically coherent once recognizable instances of admixture are taken into account. Thus, the search for species limits is complicated, but may not be in “vain” (Tilley et al., 2013). Rather, treating “species” such as *fuscus*, *conanti*, and *ocoe* as valid taxa *a priori* yields intractably complex interpretations when the results of phylogeographic analyses are viewed through that lens (as discussed by Tilley et al., 2013), but this is alleviated under the scheme of KEA05 and BL20, and corroborated here.

Rather than “failed” species, we suggest that most of the candidate species assessed here are instead “good” species, albeit ones that often exchange genes with other lineages in sympatry on the margins of their ranges. Such patterns of speciation with gene flow are increasingly recognized as common features of phylogeographic divergence (Nosil, 2008) and are already well-known in amphibians (Bogart et al., 2007; Niemiller et al., 2008). Increasing the geographic sampling of the populations identified here, as well as additional phylogenomic processing such as allele phasing (Andermann et al., 2018) will allow us to address the magnitude, direction, and prevalence of admixture among related or geographically proximate clades. Incorporating morphological,

behavioral, ecological, and other data will permit species delimitation in an integrative taxonomic framework (Dayrat, 2005; Padial et al., 2010). Together, these integrative data will provide robust estimates of species-level taxonomy, genetic diversity, historical and recent gene-flow, and relationships among lineages, whether best represented as a bifurcating tree or a reticulating network.

4.3. Cryptic diversity and JW18

The detailed study of JW18 provided the first well-documented evidence based on phylogenomic sampling that the various lineages of *quadramaculatus* and *marmoratus* did not form a monophyletic group as they did in previous mitochondrial analyses (e.g., Kozak et al., 2005). We corroborate their results here and show through the combined re-analysis of our AHE loci and their ddRAD reads that this same topological signal extends across both datasets.

In contrast, we present strong evidence that at least some of their specimens of *carolinensis* and *marmoratus* were misidentified in the field (see Results and SI), casting doubt on their conclusions regarding the spatial extent of the Pisgah clade, which we show extends well into the geographic range of the Nantahala clade. This geographic pattern is also supported by the combined nuclear and mitochondrial dataset of Jackson (2005) and the mitochondrial results of Jones et al. (2006). We also dispute their conclusions regarding the non-monophyly of the various lineages of morphological *quadramaculatus* and *marmoratus*, all of which form monophyletic groups by phenotype in our analyses.

However, given the clearly rampant gene-flow between many of these lineages as evidenced by shared mitochondrial haplotypes in the Pisgah clade, we concur with JW18 that patterns of diversification in the group have been extremely complex, and that many populations might represent hybrids. We regret that the root causes of this discordance cannot be addressed directly, as JW18 did not retain voucher specimens or photographs of these highly cryptic species. This underscores the importance of maintaining robust natural-history collections, particularly for organisms as difficult to identify in the field as *Desmognathus* salamanders. This is especially important given the strong role that such collections will play in untangling the genomic roots of phenotypic evolution in the group (see Lamichhane et al., 2019).

5. Conclusions

Taxonomic diversity, species limits, and phylogenetic relationships in *Desmognathus* are complex and unclear, as revealed by recent mitochondrial and allozyme datasets. Here, we address these issues using mitochondrial and phylogenomic datasets sampling nearly all known lineages of *Desmognathus*, often including multiple populations. We show the presence of at least 49 candidate species masquerading under the current 22-species taxonomy, many of which will likely prove to be valid, cryptic species in future analyses. However, discerning relationships among these species is confounded by their complex history of admixture across space and time. Both recent mitochondrial introgression and deep-time reticulation have occurred frequently, between populations that may be distant or proximate, spatially or phylogenetically. The increasingly fine-scale resolution of population and phylogenomic sampling will allow these historical processes to be unraveled in future analyses.

CRedit authorship contribution statement

R. Alexander Pyron: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Kyle A. O'Connell:** Formal analysis, Methodology, Writing - review & editing. **Emily Moriarty Lemmon:** Methodology, Project administration, Writing - review & editing. **Alan R. Lemmon:** Formal analysis,

Methodology, Project administration, Writing - review & editing. **David A. Beamer:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgments

This work was funded in part by U.S. NSF grants DBI-0905765, DEB-1441719, and DEB-1655737 to RAP; and DEB-1656111 to DAB. We are grateful to M. Kortyna, S. Holland, A. Bigelow, and K. Birch from FSU's Center for Anchored Phylogenomics for assistance with molecular data collection and bioinformatics analyses, to A. Wynn (NMNH) and J. Trueblood and H. Vega-Bernal (Nash CC) for assistance with specimens, and the Bell lab and K. de Queiroz (NMNH) for their critical feedback on this manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2020.106751>.

References

- Andermann, T., Fernandes, A.M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., Aleixo, A., Faircloth, B.C., Antonelli, A., 2018. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Syst. Biol.* 68, 32–46.
- Anderson, J.A., Tilley, S.G., 2003. Systematics of the *Desmognathus ochrophaeus* complex in the Cumberland Plateau of Tennessee. *Herp. Monogr.* 17, 75–110.
- Baum, D.A., Donoghue, M.J., 1995. Choosing among alternative “phylogenetic” species concepts. *Syst. Bot.* 20, 560–573.
- Beamer, D.A., Lamb, T., 2008. Dusky salamanders (*Desmognathus*, Plethodontidae) from the Coastal Plain: multiple independent lineages and their bearing on the molecular phylogeny of the genus. *Mol. Phylo. Evol.* 47, 143–153.
- Beamer, D.A., 2015. Rectifying Limitations On Species Delineation In Dusky Salamanders: Lineage Detection Using An Ecoregion-Drainage Sampling Regime (PhD). Eastern Carolina University, Greenville, N.C.
- Beamer, D.A., Lamb, T., 2020. Towards rectifying limitations on species delineation in dusky salamanders (*Desmognathus*: Plethodontidae): an ecoregion-drainage sampling grid reveals additional cryptic clades. *Zootaxa*. <https://doi.org/10.11646/zootaxa.4734.1.1>. (in press).
- Bogart, J.P., Bi, K., Fu, J., Noble, D.W.A., Niedzwiecki, J., 2007. Unisexual salamanders (genus *Ambystoma*) present a new reproductive mode for eukaryotes. *Genome* 50 (2), 119–136. <https://doi.org/10.1139/G06-152>.
- Bonett, R.M., 2002. Analysis of the Contact Zone between the Dusky Salamanders *Desmognathus fuscus fuscus* and *Desmognathus fuscus conanti* (Caudata: Plethodontidae). *Copeia* 2002 (2), 344–355. [https://doi.org/10.1643/0045-8511\(2002\)002\[0344:AOTCZB\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2002)002[0344:AOTCZB]2.0.CO;2).
- Bonnet, T., Leblois, R., Rousset, F., Crochet, P.A., 2017. A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. *Evolution* 71, 2140–2158.
- Burbink, F.T., Gehara, M., 2018. The biogeography of deep time phylogenetic reticulation. *Syst. Biol.* 67, 743–755.
- Camp, C.D., Tilley, S.G., Austin Jr., R.M., Marshall, J.L., 2002. A new species of black-bellied salamander (genus *Desmognathus*) from the Appalachian Mountains of northern Georgia. *Herpetologica* 58, 471–484.
- Conant, R., Collins, J.T., Powell, R., 2016. Peterson Field Guide to Reptiles and Amphibians of Eastern and Central North America. Houghton Mifflin Harcourt, Boston.
- Dayrat, B., 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.* 85, 407–417.
- Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? *Evolution* 63, 1–19.
- Edwards, S.V., Potter, S., Schmitt, J., Bragg, J.G., Moritz, C., 2016. Reticulation, divergence, and the phylogeography-phylogenetics continuum. *Proc. Natl. Acad. Sci. USA* 113, 8025–8032.
- Good, J.M., Vanderpool, D., Keeble, S., Bi, K., 2015. Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution* 69, 1961–1972.
- Graham, S.P., Timpe, E.K., Laurencio, L.R., 2010. Status and possible decline of the southern dusky salamander (*Desmognathus auriculatus*) in Georgia and Alabama, USA. *Herpetol. Conserv. Biol.* 5, 360–373.
- 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.
- Hare, M.P., 2001. Prospects for nuclear gene phylogeography. *Trends Ecol. Evol.* 16, 700–706.
- Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. *J. Hered.* 105, 795–809.

- Hime, P.M., Lemmon, A.R., Lemmon, E.C.M., Scott-Prendini, E., Brown, J.M., Thomson, R.C., Kratochvil, J.D., Noonan, B.P., Pyron, R.A., Peloso, P.L.V., Kortyna, M.L., Keogh, J.S., Donnellan, S.C., Mueller, R.L., Raxworthy, C.J., Kunte, K., Ron, S., Das, S., Gaitonde, N., Green, D.M., Labisko, J., Che, J., Weisrock, D.W., 2020. Phylogenomics uncovers ancient gene-tree discordance in the Amphibian Tree of Life. *Syst. Biol.*, USYB-2019-084 Submitted for publication.
- Holder, M.T., Anderson, J.A., Holloway, A.K., 2001. Difficulties in detecting hybridization. *Syst. Biol.* 50, 978–982.
- Huson, D.H., Bryant, D., 2005. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267.
- Jackson, N.D., 2005. Phylogenetic history, morphological parallelism, and speciation in a complex of Appalachian salamanders (Genus: *Desmognathus*). MS Thesis. Brigham Young University, Provo, Utah.
- Jones, M.T., Voss, S.R., Ptacek, M.B., Weisrock, D.W., Tonkyn, D.W., 2006. River drainages and phylogeography: an evolutionary significant lineage of shovel-nosed salamander (*Desmognathus marmoratus*) in the southern Appalachians. *Mol. Phylo. Evol.* 38, 280–287.
- Jones, K.S., Weisrock, D.W., 2018. Genomic data reject the hypothesis of sympatric ecological speciation in a clade of *Desmognathus* salamanders. *Evolution* 72, 2378–2393.
- Karlin, A.A., Guttman, S.I., 1981. Hybridization between *Desmognathus fuscus* and *Desmognathus ochrophaeus* (Amphibia: Urodela: Plethodontidae) in northeastern Ohio and northwestern Pennsylvania. *Copeia* 1981, 371–377.
- Katoh, K., Standley, D.M., 2013. MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kozak, K.H., Larson, A.A., Bonett, R.M., Harmon, L.J., 2005. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: *Desmognathus*). *Evolution* 59, 2000–2016.
- Kratochvil, J.D., 2017. Mitochondrial and nuclear patterns of conflict and concordance at the gene, genome, and behavioral scales in *Desmognathus* salamanders. PhD Thesis. University of Kentucky, Lexington.
- Kuo, C.-H., Avise, J.C., 2005. Phylogeographic breaks in low-dispersal species: the emergence of concordance across gene trees. *Genetica* 124, 179–186.
- Lamichanay, S., Card, D.C., Grayson, P., Tonini, J.F.R., Bravo, G.A., Näpflin, K., Termignoni-Garcia, F., Torres, C., Burbrink, F., Clarke, J.A., Sackton, T.B., Edwards, S.V., 2019. Integrating natural history collections and comparative genomics to study the genetic architecture of convergent evolution. *Phil. Trans. R. Soc. B: Biol. Sci.* 374, 20180248.
- Leaché, A.D., 2009. Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). *Syst. Biol.* 58, 547–559.
- Leaché, A.D., 2010. Species trees for spiny lizards (Genus *Sceloporus*): Identifying points of concordance and conflict between nuclear and mitochondrial data. *Mol. Phylo. Evol.* 54, 162–171.
- Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61, 727–744.
- Liu, L., Yu, L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. *Mol. Phylo. Evol.* 53, 320–328.
- Mallet, J., Besansky, N., Hahn, M.W., 2016. How reticulated are species? *BioEssays* 38, 140–149.
- Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., Jiggins, C.D., 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome. Res.* 23, 1817–1828.
- Mead, L.S., Tilley, S.G., Katz, L.A., 2001. Genetic structure of the Blue Ridge dusky salamander (*Desmognathus orestes*): inferences from allozymes, mitochondrial DNA, and behavior. *Evolution* 55, 2287–2302.
- Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30, i541–i548.
- Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* 2010 (6), pdb.prot5448. <https://doi.org/10.1101/pdb.prot5448>.
- Mirarab, S., Bayzid, M.S., Warnow, T., 2016. Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Syst. Biol.* 65, 366–380.
- Near, T.J., 2019. Phylogenomic signatures of ancient introgression in a rogue lineage of darters (Teleostei: Percidae). *Syst. Biol.* 68 (2), 329–346. <https://doi.org/10.1093/sysbio/syy074>.
- Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Niemiller, M.L., Fitzpatrick, B.M., Miller, B.T., 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Mol. Ecol.* 17, 2258–2275.
- Nosil, P., 2008. Speciation with gene flow could be common. *Mol. Ecol.* 17, 2103–2106.
- Omernik, J.M., Griffith, G.E., 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Env. Man.* 54, 1249–1266.
- Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. *Front. Zool.* 7, 16.
- Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Gen.* 8, e1002967.
- Pope, C.H., 1924. Notes on North Carolina salamanders with especial reference to the egg-laying habits of *Leurognathus* and *Desmognathus*. *Am. Mus. Nov.* 153, 1–15.
- Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., Lemmon, A.R., 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526, 569–573.
- 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. Phylo. Evol.* 81, 221–231.
- Pyron, R.A., Hsieh, F.W., Lemmon, A.R., Lemmon, E.M., Hendry, C.R., 2016. Integrating phylogenomic and morphological data to assess candidate species-delimitation models in Brown and Red-bellied snakes (*Storeria*). *Zool. J. Linn. Soc.* 177, 937–949.
- Rissler, L.J., Taylor, D.R., 2003. The phylogenetics of *Desmognathus* salamander populations across the southern Appalachians. *Mol. Phylo. Evol.* 27, 197–211.
- Rokyta, D.R., Lemmon, A.R., Margres, M.J., Aronow, K., 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genom.* 13, 312.
- Ruane, S., Bryson Jr., R.W., Pyron, R.A., Burbrink, F.T., 2014. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. *Syst. Biol.* 63, 231–250.
- Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.* 33, 1654–1668.
- Solis-Lemus, C., Ané, C., 2016. Inferring phylogenetic networks with maximum pseudo-likelihood under incomplete lineage sorting. *PLoS Gen.* 12, e1005896.
- Solis-Lemus, C., Yang, M., Ané, C., 2016. Inconsistency of species tree methods under gene flow. *Syst. Biol.* 65, 843–851.
- Solis-Lemus, C., Bastide, P., Ané, C., 2017. PhyloNetworks: a package for phylogenetic networks. *Mol. Biol. Evol.* 34, 3292–3298.
- Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.
- Than, C., Ruths, D., Nakhleh, L., 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinform.* 9, 322.
- Tilley, S.G., Merritt, R.B., Wu, B., Highton, R., 1978. Genetic differentiation in salamanders of *Desmognathus ochrophaeus* complex (Plethodontidae). *Evolution* 32, 93–115.
- Tilley, S.G., Mahoney, M.J., 1996. Patterns of genetic differentiation in salamanders of the *Desmognathus ochrophaeus* complex (Amphibia: Plethodontidae). *Herp. Monogr.* 10, 1–42.
- Tilley, S.G., Eriksen, R.L., Katz, L.A., 2008. Systematics of dusky salamanders, *Desmognathus* (Caudata: Plethodontidae), in the mountain and Piedmont regions of Virginia and North Carolina. *USA. Zool. J. Linn. Soc.* 152, 115–130.
- Tilley, S.G., Bernardo, J., Katz, L.A., Lopez, L., Roll, J.D., Eriksen, R.L., Kratochvil, J., Bittner, N.K.J., Crandall, K.A., 2013. Failed species, innumerate forms, and the vain search for species limits: cryptic diversity in dusky salamanders (*Desmognathus*) of eastern Tennessee. *Ecol. Evol.* 3, 2547–2567.
- Titus, T.A., Larson, A., 1996. Molecular phylogenetics of *desmognathine* salamanders (Caudata: Plethodontidae): a reevaluation of evolution in ecology, life history, and morphology. *Syst. Biol.* 45, 451–472.
- Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930.
- Vachaspati, P., Warnow, T., 2015. ASTRID: accurate species trees from internode distances. *BMC Genom.* 16, S3.
- Vijay, N., Bossu, C.M., Poelstra, J.W., Weissensteiner, M.H., Suh, A., Kryukov, A.P., Wolf, J.B., 2016. Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Comm.* 7, 13195.