Molecular Phylogenetics and Evolution 206 (2025) 108309

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Disentangling a genome-wide mosaic of conflicting phylogenetic signals in Western Rattlesnakes

Justin M. Bernstein [a], Yannick Z. Francioli [a], Drew R. Schield [b], Richard H. Adams [c], Blair W. Perry [d], Keaka Farleigh [b], Cara F. Smith [e], Jesse M. Meik [f], Stephen P. Mackessy [g], Todd A. Castoe [a,*]

[a] Department of Biology, University of Texas at Arlington, Arlington, TX 76019, USA
[b] Department of Biology, University of Virginia, Charlottesville, VA 22903, USA
[c] Department of Entomology and Plant Pathology, University of Arkansas Agricultural Experimental Station, University of Arkansas, Fayetteville, AR 72701, USA
[d] Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA 95064, USA
[e] Department of Biochemistry and Molecular Genetics, 12801 East 17th Avenue, University of Colorado Denver, Aurora, CO 80045, USA
[f] Department of Biological Sciences, Tarleton State University, Stephenville, TX 76402, USA
[g] School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639, USA

ARTICLE INFO

Keywords:
Crotalus
Genome biology
Introgression
Phylogeny
Recombination
Reticulation
Site-specific likelihood

ABSTRACT

Species tree inference is often assumed to be more accurate as datasets increase in size, with whole genomes representing the best-case-scenario for estimating a single, most-likely speciation history with high confidence. However, genomes may harbor a complex mixture of evolutionary histories among loci, which amplifies the opportunity for model misspecification and impacts phylogenetic inference. Accordingly, multiple distinct and well-supported phylogenetic trees are often recovered from genome-scale data, and approaches for biologically interpreting these distinct signatures are a major challenge for evolutionary biology in the age of genomics. Here, we analyze 32 whole genomes of nine taxa and two outgroups from the Western Rattlesnake species complex. Using concordance factors, topology weighting, and concatenated and species tree analyses with a chromosome-level reference genome, we characterize the distribution of phylogenetic signal across the genomic landscape. We find that concatenated and species tree analyses of autosomes, the Z (sex) chromosome, and mitochondrial genome yield distinct, yet strongly supported phylogenies. Analyses of site-specific likelihoods show additional patterns consistent with rampant model misspecification, a likely consequence of several evolutionary processes. Together, our results suggest that a combination of historic and recent introgression, along with natural selection, recombination rate variation, and cytonuclear co-evolution of nuclear-encoded mitochondrial genes, underlie genome-wide variation in phylogenetic signal. Our results highlight both the power and complexity of interpreting whole genomes in a phylogenetic context and illustrate how patterns of phylogenetic discordance can reveal the impacts of different evolutionary processes that contribute to genome-wide variation in phylogenetic signal.

1. Introduction

The rapid accumulation of genome-scale datasets holds great promise for improved phylogenetic inferences and a better-resolved Tree of Life. However, while it may be expected that more data will lead to more clear and decisive phylogenetic estimates, we often find extensive conflicting phylogenetic signal across sites, genes, genomes, and analytical methods (Roycroft et al., 2020; Zhao et al., 2023). Indeed, genome-scale datasets often exhibit substantial phylogenetic conflict that manifests as differing but well-supported topologies among different genomic regions (Schrempf and Szöllősi, 2020) or even between distinct nucleotide sites within a locus (Castoe et al. 2009; Steenwyk et al. 2023). Accordingly, analyses of large genome-scale datasets bring new challenges for phylogenetic inference. An understanding of how statistical artifacts, various evolutionary processes, and aspects of genome biology may interact to produce conflicting

* Corresponding author.
E-mail address: todd.castoe@uta.edu (T.A. Castoe).

https://doi.org/10.1016/j.ympev.2025.108309
Received 31 October 2024; Received in revised form 4 February 2025; Accepted 8 February 2025
Available online 10 February 2025
1055-7903/© 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

phylogenetic signals across the genome is key to generating robust and replicable phylogenetic inference and for gaining biological insight from these distinct evolutionary signals.

Differing estimates of the underlying phylogenetic tree across distinct genomic regions or sites can result from many factors and processes, including incomplete lineage sorting (ILS; Meleshko et al., 2021; Parins-Fukuchi et al., 2021; Singhal et al., 2021; Smith et al., 2015; Zhao et al., 2023), natural selection (Adams et al., 2018; Castoe et al., 2009; Edwards, 2009), migration and introgression (Lin et al. 2019; Reilly et al. 2019; Sanderson et al., 2023), and population structure (Oliver, 2013). Each of these elements of the overall diversification history of a species group can violate models used in phylogenetic inference (Castoe et al., 2009) and result in alternative topologies that are often deemed "incorrect" and discarded. ILS is one of the most widespread and well-studied sources of discordance between gene trees and the overall species trees (Liu and Edwards, 2009; Mendes and Hahn, 2018), leading to the use of multispecies coalescent (MSC)-based methods that take into account ongoing population genetic processes that generate ILS (Mirarab et al., 2016). The use of heuristic summary statistic strategies and the MSC (Liu et al., 2010; Zhang et al., 2018) for large phylogenomic datasets (Faircloth et al., 2012; Lemmon et al., 2012) has seen widespread adoption throughout the field of systematics. These approaches seek to address discordance produced by ILS, introgression, and population structure (Cloutier et al. 2019; Jiang et al. 2020; Jiao et al. 2021), but high levels of gene tree estimation error on numerous but short loci (e.g., ultraconserved elements or anchored hybrid enrichment loci) can significantly hinder our ability to detect reliable phylogenetic signal and the phenomena that shape it (Blom et al., 2016; Roch and Warnow, 2015). As datasets become larger, the impacts of model violation and misspecification may also be amplified, further increasing potential phylogenetic conflict as a consequence of gene tree estimation error and poor model fit; each of these must also be disentangled from true, distinct evolutionary signals (Gatesy and Springer, 2014; Roch and Warnow, 2015).

Despite these inherent challenges, genome-scale data can enhance our understanding of complex evolutionary processes that have shaped genomic diversity across lineages and reveal how these processes may interact with aspects of genome biology (e.g., the recombination landscape). Accordingly, understanding heterogeneity in phylogenetic signal across the genome can advance our understanding of processes that have shaped extant genomic diversity. Identified patterns of heterogeneity can, in turn, help to establish predictive frameworks that can aid in disentangling complex speciation scenarios such as those involving both introgression and natural selection.

Here, we address the challenge of interpreting whole genome data in a phylogenetic context using the Western Rattlesnake species complex (genus *Crotalus*). These rattlesnakes represent a collection of three species groups, each consisting of multiple lineages, whose range collectively spans much of western North America, and includes the Prairie Rattlesnake (*C. viridis*), Western Rattlesnake (*C. oreganus*), and Mojave Rattlesnake (*C. scutulatus*). This species complex displays high levels of variation in their morphology, coloration, venom composition, and natural history, and multiple lineages are known to hybridize (Ashton, 2001; Mackessy, 2010; Myers, 2021; Schield et al., 2019; Schield et al., 2018; Smith et al., 2023; Strickland et al., 2018). While previous studies have suggested several competing hypotheses for the taxonomy of this group (Ashton, 2001; Davis et al., 2016; Klauber, 1972), this complex is currently recognized as containing seven nominal species: *C. viridis*, *C. oreganus*, *C. cerberus*, *C. concolor*, *C. lutosus*, *C. helleri*, and *C. scutulatus* (Uetz et al., 2023). In addition to the inherent importance of the taxonomy, the extensive variation in morphological and venom traits within this group make it an excellent model for comparative biology. However, such efforts are hindered by a lack of agreement regarding the phylogeny of the group (Ashton and Queiroz, 2001; Campbell and Lamar, 2004; Davis et al., 2016; Douglas et al., 2002; Klauber, 1972; Myers et al., 2024; Pook et al., 2000).

Here, we use whole genome resequencing data from 30 individuals representing all major lineages of the Western Rattlesnake species complex and a chromosome-level reference genome for the Prairie Rattlesnake (*Crotalus viridis*; Schield et al. 2019a) to investigate genome-wide variation in phylogenomic signal. We take a genome-informed view of phylogenetic reconstruction by considering the recombination landscape and exploring the distribution of phylogenetic signals across the mitochondrial genome, nuclear macrochromosomes, nuclear microchromosomes, and sex (Z) chromosomes. We use this dataset to i) understand how phylogenetic estimates vary within and among chromosomes, ii) identify processes that contribute to interspecific genomic variation while taking recombination into account, and iii) determine the most likely speciation scenario of rattlesnakes in the Western Rattlesnake complex. Our analyses yield multiple strongly-supported trees estimated from specific genomic regions and analyses, and we investigate the distribution of conflicting signals across sites and regions of the genome to disentangle potential underlying causes of these patterns.

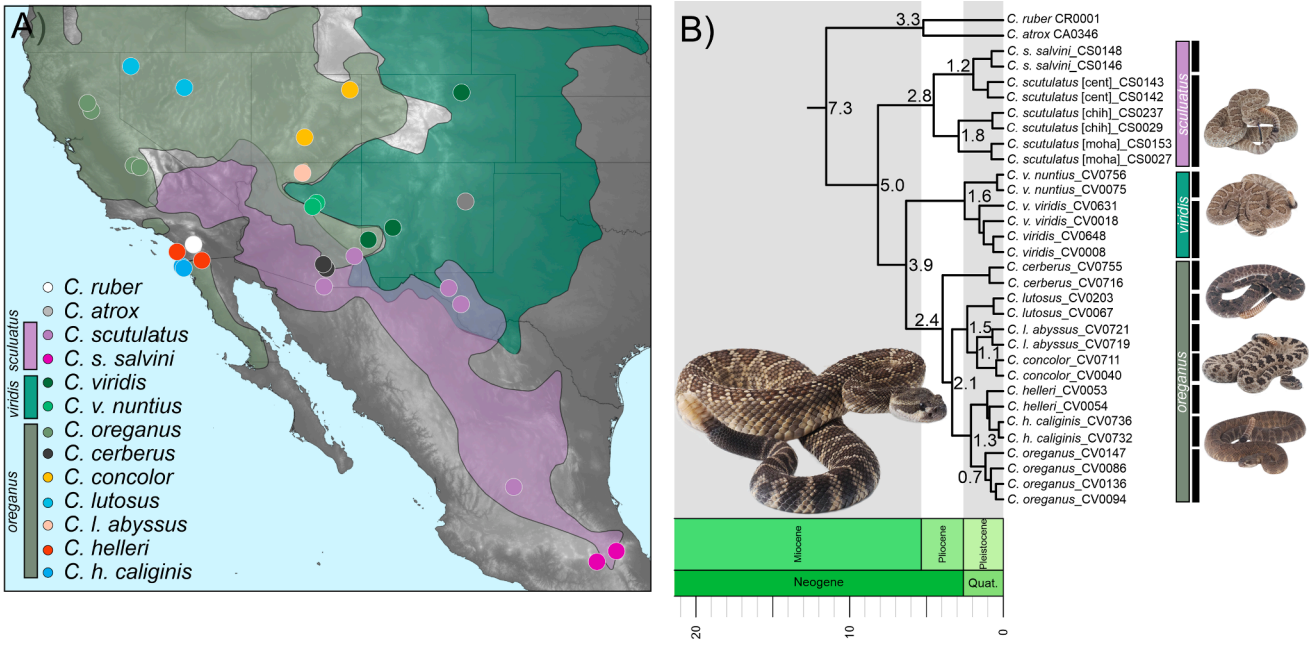
2. Materials & methods

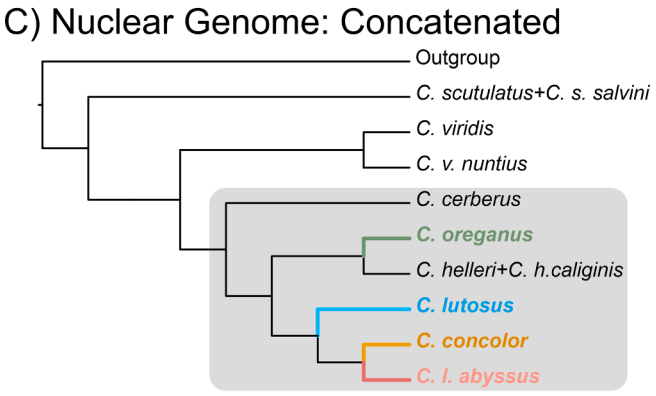
2.1. Sample collection and whole genome resequencing

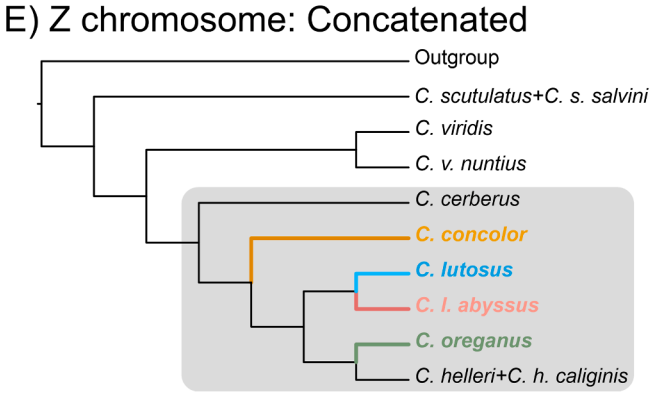
We obtained two samples for each major lineage of the Western Rattlesnake species complex for whole genome sequencing, sampling individuals from nearby localities for each lineage where possible (Supplementary Table S1; Fig. 1A). The Northern Pacific Rattlesnake (*Crotalus oreganus oreganus*) and the Prairie Rattlesnake (*C. viridis viridis*) span large geographic distributions and population substructure (Schield et al., 2019; Schield et al., 2022). We thus sampled a total of four individuals for each of these lineages (i.e., two individuals from different regions with evidence of distinct population genetic structure; specifically, *C. v. viridis* were sampled from Colorado and New Mexico, and *C. o. oreganus* were sampled from populations north and south of the San Francisco Bay area of California). Within the Mojave Rattlesnake (*C. scutulatus*) complex, we sampled pairs of individuals for each of the four major lineages identified in (Schield et al., 2018). We also sampled one individual each of C. *atrox* and *C. ruber* as outgroup taxa, for a total sampling of 32 individuals (Supplementary Table S1; Fig. 1A). Our taxonomy and designation of species and subspecies here is based on the current recognition in The Reptile Database (Uetz et al., 2023). We extracted genomic DNA from liver or blood tissue for each specimen, which were stored in DNA lysis buffer or snap frozen using liquid nitrogen and stored at −80℃. DNA was extracted using standard phenol–chloroform-isoamyl extraction and precipitation, and we quantified purified DNA products using a Qubit fluorometer. We prepared genome sequencing libraries from purified DNA using KAPA HyperPlus and Illumina Nextera DNA Flex kits, and sequenced genomic libraries on Illumina NovaSeq 6000 S4 lanes with 150 bp paired-end reads. Raw sequencing reads were adapter trimmed and processed using the Illumina BaseSpace distributions of FastQC v1.0.0 and Fastq Toolkit v2.2.0 (BaseSpace Labs). We trimmed bases on 5′ and 3′ ends of reads with quality scores < 20 and discarded any reads with a final length < 36 bp, or that had an average quality score < 30.

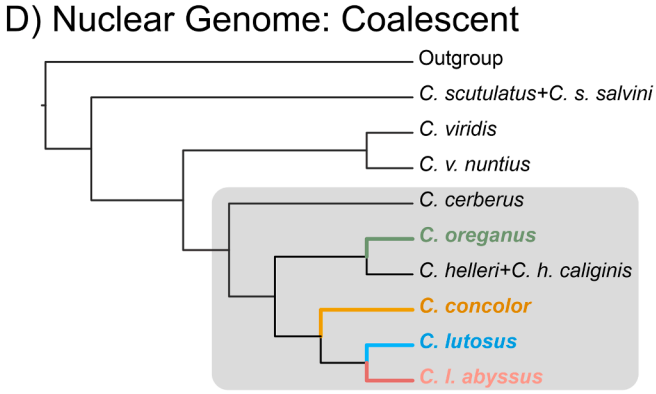
2.2. Reference genome, annotation, and recombination rate estimation

We used the Prairie Rattlesnake (*C. viridis*) genome assembly (CroVir3.0; [Schield et al. 2019a]) as the reference for our whole genome resequencing data and used the genome annotation for downstream dataset filtering and parsing. The assembly and annotation include chromosome-length scaffolds corresponding to the 18 chromosomes in the rattlesnake karyotype ($2n = 36$) and 17,352 annotated protein-coding genes. Repeat elements were also previously annotated using RepeatMasker (Smit et al., 2015), as described in Pasquesi et al. (2018) and Schield et al. (2019a). Population-scaled mean recombination rate estimates for *C. viridis* were taken directly from Schield et al. (2020).









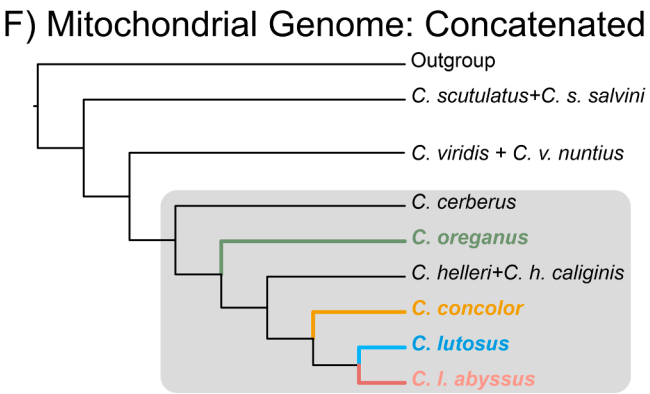


Fig. 1. Phylogeny of the CVOS group. A) Map showing the sampling of species and subspecies used in this study (points = individuals). Taxa are named based on currently supported species status. The ranges of the three species groups (*C. scutulatus*, *C. viridis, and C. oreganus*) are shown by shaded polygons and represented by the vertical bars in the key. B) Chronogram based of the concatenated maximum likelihood (ML) tree of CVOS. Colored bars next to tips represent the broader species groups; black bars designate lineages by species and subspecies. Timescale is shown below phylogeny (Holocene not shown). All nodes are strongly supported. Numbers in front of nodes (or above/below due to space constraints) represent divergence dates in millions of years. C–F) Simplified phylogenies from four datasets and analyses: concatenated ML with all chromosomes (C), species tree with all chromosomes (D), concatenated ML tree of the Z chromosome (E), and concatenated ML tree of all heavy-strand-encoded mitochondrial genes. The *C. oreganus* species group is shown in a grey bubble, with colored branches (corresponding to the map) showing lineages with alternative topologies across trees. Trees in C–F show topology only; branch lengths are not informative.

2.3. Variant calling and filtering

We trimmed and filtered all genomic reads and aligned them to the *C. viridis* reference genome using the BWA v0.7.1 mem algorithm (Li and Durbin, 2009) with default settings, and marked duplicate mappings for downstream removal. We quantified read mapping statistics using samtools v1.9 stats and flagstats (Li et al., 2009). On average, 96.51 % of reads mapped to the reference genome among ingroup taxa; proportions of mapped reads were slightly lower for outgroup taxa (93.64 % and 94.69 % in *C. atrox* and *C. ruber*, respectively). Read mapping resulted in an average coverage of 37.6 × per sample (range = 10.5–138.4 × ; Supplementary Table S1). We used the GATK v4.0.8.1 HaplotypeCaller (McKenna et al., 2010) to call indels and single nucleotide polymorphisms (SNPs) within individuals, and to generate genomic variant call files (gVCFs), which contain information for all sites in the genome. We then called variants across all 32 individuals in our sampling using the GATK v3.8.1.0 GenotypeGVCFs module, specifying the '-allSites' option to output information at each genomic position among individuals. We used v3.8.1.0 for this step because it is capable of analyzing multiple gVCFs simultaneously without having to first merge gVCFs.

We performed post-variant discovery filtering steps using bcftools v1.9 (Li et al., 2009), in order to remove spurious effects and artifacts of repetitive regions, indels, and low quality base calls. First, we recoded all positions (variant and invariant) with read depth below 5× (i.e., DP < 5) or that had genotype quality less than 30 (i.e., GQ < 30) as missing genotypes per individual using the bcftools filter module. To avoid the potential effects of indel misalignment, and to also maintain associations between variant information and features of the *C. viridis* genome annotation, we recoded indels as ambiguous genotypes ('N's). After these filtering steps, we generated a consensus genome sequence for each individual using bcftools consensus, specifying IUPAC ambiguity codes for heterozygous sites (−I), 'N's for missing genotypes (−M N), and including SNP positions and invariant sites retained after previous filtering steps (−i 'type="snp" | alt="."'). Here, we also masked all positions that overlapped with repeat elements in the genome as ambiguous genotypes, based on the *C. viridis* reference genome repeat element annotation.

2.4. Preparation of genome-scale datasets

We parsed sequence alignments from our filtered genomic dataset to examine variation in genome-wide estimates of the evolutionary history of the species complex. We first quantified the mean and maximum missing data (i.e., ambiguous genotypes after upstream filtering steps) among samples for all nonoverlapping 10 kb windows of the genome. We then identified the set of alignments with mean and maximum 'missingness' below 35 % and 60 %, respectively. This included 31,093 10 kb regions with no missing taxa, which sampled representative regions from all chromosomes. Raw sequencing reads are available on the Sequence Read Archive (SRA) under Projects PRJNA1150930 and PRJNA1137891; all sample information can be found in Supplementary Table S1.

2.5. Concatenated and species tree analyses

We used both concatenation and coalescent-based species tree methods to obtain estimates of the nuclear phylogeny for the Western Rattlesnake species complex. First, we concatenated all 31,093 alignments of 10 kb length in the phylogenomic dataset described above using 'catsequences' (https://github.com/ChrisCreevey/). The final concatenated alignment had a total length of 310,000,193 bp. We inferred a maximum likelihood tree using IQ-TREE v2.2.6 (Minh et al., 2020b), using the ModelFinder (Kalyaanamoorthy et al., 2017) option to select the best partition scheme and molecular substitution model for each partition. We calculated support for nodes in the ML phylogeny by performing 1000 ultrafast bootstrap (UFB) replicates (Hoang et al., 2018); relationships with ≥ 95 are considered strongly supported.

Second, we inferred a coalescent-based phylogeny using a two-step procedure. Phylogenies were inferred for each 10 kb alignment using a custom shell script (https://github.com/jbernst) to iterate through all 10 kb alignments in parallel batches. Here, we refer to the phylogenies from 10 kb windows as 'gene trees.' These 31,093 estimated gene trees were used as input to estimate the species tree using ASTRAL-III (Zhang et al., 2018). We assessed gene tree discordance among nodes in the inferred species tree topology by annotating quartet support for each bipartition using the '-t 8′ option in ASTRAL-III. Local posterior probabilities ≥ 0.95 were considered strongly supported relationships. ASTRAL was run using a mapping file that allowed for each species and subspecies to represent its own lineage, with the exceptions of *C. s. salvini* being grouped within *C. scutulatus* and *C. h. caliginis* grouped within *C. helleri* (see *Results*).

We also partitioned the estimated gene trees into three different sets for phylogenetic reconstructions: i) the entire gene tree set (all 31,093), ii) 17 subsets of autosome-specific gene trees (i.e., one set of gene trees from each autosome), and iii) the set of gene trees (453 trees) estimated from the Z (sex) chromosome. We obtained a concatenated and coalescent species tree for each of these sets of gene trees. Estimated gene and species trees were visualized using FIGTREE v1.4.4 (Rambaut, 2012). Additionally, for some analyses in our study we reference chromosomes by 'type', which refers to the following: macrochromosomes, microchromosomes, and the Z (sex) chromosome.

2.6. Mitochondrial genome assembly and analysis

We assembled near-complete mitochondrial genomes from whole genome shotgun reads for each individual using MITObim v1.9.1 (Hahn et al., 2013), which leverages the assembly tool MIRA v4.0.1 (Chevreux et al., 1999). For each sample, we performed a reference-guided assembly in MITObim, using a mitochondrial genome for *C. concolor* as a reference (reference genome GenBank accession: PV035282; GenBank accession numbers of individual mitochondrial genes for all samples are available in Supplementary Table S3). MITObim was run using a maximum of 10 iterations of baiting and mapping for each sample by mapping paired reads determined to be derived from the mitochondrial genome using default parameters. We annotated our newly-assembled mitochondrial genomes using the MITOS2 (v2.1.9) on Galaxy (Arab et al. 2017; Donath et al., 2019), which provided sequence coordinates for the entire complement of ribosomal RNA, transfer RNA, and protein coding genes for each individual. We performed phylogenetic analyses on a concatenated alignment of all 12 heavy-strand-encoded mitochondrial protein coding genes. We extracted gene sequences from the assembled mitochondrial genomes and performed alignments using Geneious R11.1.5 with the Geneious Algorithm under default parameters. After inspecting each alignment and performing minor manual edits to account for gaps, we concatenated each alignment. We performed phylogenetic reconstruction on the concatenated mitochondrial gene alignment using the same workflow as was used for nuclear data.

2.7. Statistical analysis of tree discordances and tree space

We used IQ-TREE2 (Minh et al., 2020b) to calculate genealogical concordance using two metrics: gene and site concordance factors (gCF and sCF, respectively; Minh et al. 2020a; Mo et al. 2023). To understand discordance at both the species and population levels, we used our concatenated tree as a reference, which was largely congruent to the coalescent-based tree (*see Results*, Fig. 1). gCF and sCF values represent the percentage of genes or sites, respectively, in an alignment that support a particular branch in the reference tree. Together, these metrics provide information on the percentage of gene trees and sites supporting the nuclear concatenated tree. We tested for significant differences in gCF and sCF values between chromosome types using an analysis of

variance (ANOVA) and a Tukey's Honestly Significant Difference (HSD) test in R.

Because we inferred multiple distinct but well-supported phylogenies from different types of analyses and chromosome sets (see *Results*), we were motivated to further test if particular chromosomes were associated with distinct gene tree signals. For this, we analyzed the distribution of the inferred phylogenies in tree space in the R package TreeSpace v1.1.4.3 (Jombart et al., 2017). We computed pairwise distance matrices between all inferred gene trees based on the Robinson-Foulds metrics. We applied multidimensional scaling (MDS; principal coordinates analysis; PCoA) to these pairwise distances to visualize gene tree distributions within and across different chromosomes. After MDS analyses, hierarchical clustering was performed to identify evidence of gene tree groups using Ward's method. We also designated three PCoA groupings to visualize if topologies of gene trees from chromosome types were separated in tree space. Due to the size of our dataset, we used a maximum of 500 random gene trees per chromosome for a total dataset of 7106 trees for tree space analysis.

To quantify phylogenetic similarity between coalescent tree inferences, we calculated probabilistic species tree distances (pSTD; Adams and Castoe 2020) on 19 coalescent trees estimated using ASTRAL, including a species tree estimated from individual chromosomes (18 trees) and a species tree constructed using gene trees from all chromosomes. We calculated Hellinger, Kullback-Leibler, and Jensen-Shannon distances using the R package *pSTDistanceR* v0.0.0.9000 (Adams and Castoe, 2020), which uses the program HYBRID-COAL v0.2.1 (Zhu and Degnan, 2016) to extract gene tree topology probabilities under both the multispecies coalescent (without hybridization) and the network multispecies coalescent (with hybridization).

2.8. Divergence date estimation

To maximize the amount of data used in our divergence dating while keeping computation times feasible, we used the treePL v1.0 (Smith and O'Meara, 2012) penalized likelihood method. The use of treePL for divergence time estimation requires an input tree based on a given number of nucleotide sites in an alignment and calibration points with minimum and maximum dates. We use our concatenated tree instead of the ASTRAL tree because the former is based on nucleotide sequence data with most tips collapsed into lineages, while the latter is based on already-constructed gene trees.

We performed treePL analysis until convergence and used optimal run parameters with the 'thorough' and 'prime' commands, respectively. We performed this process five times to ensure consistency of results. To identify the best smoothing parameter for our analysis, which affects the rate variation penalty across the tree, we used random subsampling replicate cross validation (RSRCV). This approach samples multiple terminal nodes with replacement and calculates the rates and dates of the tree with terminal nodes removed. Then, the average error is sampled over the nodes (Smith and O'Meara, 2012). The smoothing parameter with the lowest error (=0.1) was used in the analysis. We used three dates for our calibrations. The first was the divergence of *C. atrox* and *C. ruber* following Castoe et al. (2007); we obtained the 95 % confidence interval of this node (normal distribution, 3.2 million years [Myr] offset, standard deviation [SD] of 1; Castoe et al. 2007). The second calibration was using the Late Miocene fossils on the ancestor of *C. scutulatus*, *C. viridis*, and *C. oreganus* (Holman, 2000). We follow Schield et al. (2019b), which set a log normal prior for this ancestor with an offset of 5 Myr and a mean of 0.01 and SD of 0.6 based on the fossil dates. We set our third calibration on the most recent common ancestor of all taxa in the tree. Here, we set the minimum date as the median estimated divergence time obtained in Schield et al. (2019b) (7.32 Myr), with no maximum bound.

2.9. Tests for introgression using reticulation networks

Prior research has shown evidence of ancient and contemporary hybridization among multiple lineages in this group, particularly at the boundaries of where species ranges come into contact or overlap (Glenn and Straight, 1990; Myers, 2021; Nikolakis et al., 2022; Pook et al., 2000; Schield et al., 2019; Zancolli et al., 2016). To test if the Western Rattlesnake phylogeny is better explained by reticulate evolution than by a strictly bifurcating tree, and to see if different chromosomes yield different reticulate topologies, we ran PhyloNet v3.8.2 (Wen et al., 2018). Using the gene trees used in phylogenetic reconstruction as input, we ran PhyloNet on gene trees from each chromosome separately, and for all chromosomes together. We used the InferNetwork_MPL approach with 10 iterations per run and ran an analysis with these parameters, testing for a maximum of 1 to 5 reticulations. We considered the iteration with the highest total log probability as the most well-supported inference from each analysis. Although PhyloNet has been shown to be accurate for species network inference (Hibbins and Hahn, 2022), it includes a correction for increased model complexity, which may scale poorly for larger, more complex datasets (Elworth et al., 2019; Wen et al., 2018). Thus, we conservatively use this analysis primarily to test whether different chromosomes qualitatively support reticulation vs bifurcating topologies rather than interpreting the results as definitive estimates of the number of reticulation events.

To further test for introgression and reticulating phylogenies, we ran network inference under a network multispecies coalescent framework using NANUQ implemented in the MSCQuartets v2.0 R package (Rhodes et al., 2021). NANUQ relies on a two-hypothesis testing framework using an α and β parameter; smaller values of α and β require stronger evidence of reticulations and any resolution of the network, respectively. We ran several iterations of this analysis with different α and β parameters using all chromosomes, only macrochromosomes, only microchromosomes, and the Z chromosome. We chose the optimal α and β parameters (conservatively choosing low values) for each run based on visual inspection of simplex plots, as recommended (Allman et al. 2019; see *Results* for α and β parameters). Because most topological discordance observed in the resulting trees (see *Results*) were within the *C. oreganus* group (*C. cerberus*, *C. oreganus*, *C. concolor*, *C. lutosus*, *C. l. abyssus*, *C. helleri*, *C. h. caliginis*), we also ran the four NANUQ analyses only on the *C. oreganus* group to test for introgression within this clade, specifically the Intermontane Clade (ImC: *C. concolor*, *C. lutosus*, and *C. l. abyssus*). The resulting quartet counts were used to construct a network quartet distance matrix between taxa. This matrix was then used to create a splits graph under the Neighbor-Net algorithm in SplitsTree4 (Huson and Bryant, 2006).

2.10. Topological investigation of phylogenies

To investigate gene tree variation across regions of the genome, we performed topology weighting by iterative sampling of sub-trees using Twisst (Martin & Belleghem, 2017). Twisst requires species groupings to identify the frequency of all possible topologies in the genome based on the defined groups. Thus, to limit the number of unrooted topologies in our analysis, we defined groups and performed analyses for two different datasets. The first dataset included *C. scutulatus*, *C. viridis*, and *C. oreganus* (with *C. atrox* + *C. ruber,* together, as an outgroup); this run collapsed *C. s. salvini* into *C. scutulatus*, *C. v. nuntius* into *C. viridis*, and all other ingroup samples into *C. oreganus*; we refer to this as the 'species group' dataset. The second dataset included only members of the *C. oreganus* species group (*C. cerberus, C. oreganus*, *C. helleri*, *C. h. caliginis*, *C. concolor*, *C. lutosus*, and *C. l. abyssus*) to address discordance between the mitochondrial tree and nuclear tree that is specific to this group (see *Results*). Using *C. cerberus* as an outgroup, we ran Twisst on this dataset with three defined ingroups: *C. oreganus*, *C. helleri* + *C. h. caliginis*, and the Intermontane Clade (*C. concolor*, *C. lutosus*, and *C. l. abyssus*, as one group). We refer to this as the '*oreganus* clade' dataset.

In sum, our grouping schemes for topology weighting analysis each contained three ingroup clades, resulting in three topologies that were weighted across the genome for each Twisst run. The species group dataset yielded a tree that matches the species tree, and two alternative trees. The *oreganus* clade dataset yielded three trees, one that matches the mitochondrial topology, another that matches the species tree topology, and a third alternative topology.

To explore the relationship between variation in tree topologies and recombination rate, we summarized the proportion of topology weights for both the species group and *oreganus* clade datasets across the range of mean recombination rate in the genome, split into 10 bins of equal size in recombination rate range and number of gene trees in each bin (with each successive bin from 1 to 10 increasing in recombination rate). Separately, we conducted additional analyses to test whether nuclear-encoded mitochondrial genes were enriched for higher topology weights for the alternative mitochondrial topology. To do this, we used Twisst to obtain the proportions of topology weights supporting the mitochondrial tree at all 10 kb regions that overlap a nuclear-encoded mitochondrial gene (n = 12 genes, 41 10 kb windows). We then compared topology weights at these regions with 41 randomly selected 10 kb regions in the genome that do not overlap with nuclear-encoded mitochondrial genes. A Fisher's exact test was used to test for enrichment for higher mitochondrial tree weights in the set of 10 kb windows that overlap with nuclear-encoded mitochondrial genes.

2.11. Testing alternative tree support using site-specific likelihoods

Evolutionary processes such as natural selection or convergence can lead to model misspecification/violation and, consequently, to extreme site-specific likelihood differences for competing trees, along with extreme site-specific support for alternative (non-species tree) topologies (Castoe et al., 2009). To further explore genome-wide variation in support for alternative topologies, we estimated site-specific log-likelihoods (SSLs) in IQ-TREE2. Three sets of site-specific likelihoods were calculated for each chromosome by comparing each alignment to three topologies: 1) nuclear (i.e., based on all chromosomes), 2) mitochondrial, and 3) Z chromosome. We calculated the difference in SSLs (ΔSSL) between the nuclear and mitochondrial topologies, and nuclear and Z topologies, respectively, to identify support for the nuclear versus alternative (mitochondria or Z chromosome) topologies genome-wide. We calculated ΔSSL so that positive and negative ΔSSL values indicate support for the nuclear topology and alternative topology, respectively. ΔSSL patterns were then summarized by 1) relative frequency of sites supporting an alternative tree, and by 2) the overall sum of likelihoods over our 10 kb genomic windows/alignments. We plotted the proportion of sites within 10 kb windows that support the nuclear (positive values) versus alternative (negative values) topology using the R package karyoploteR v1.28.0 (Gel and Serra, 2017) and custom R scripts (https://github.com/jbernst). For analyses that involve nuclear-encoded mitochondrial genes, we focused on the set of nuclear-encoded genes that represent subunits of the mitochondrial oxidative phosphorylation complex I–V pathway (Oxidative Phosphorylation; OXPHOS hereafter), based on the KEGG database (Kanehisa, 2019, 2000; Kanehisa et al., 2023). For the nuclear versus mitochondrial comparison of site likelihoods, we also identified how many sites containing extreme values of ΔSSL were found within nuclear-encoded mitochondrial genes.

3. Results

3.1. Topological heterogeneity across chromosomal trees

Our divergence dating analysis suggests that the Western Rattlesnake clade diverged from the outgroup (*C. atrox* and *C. ruber*) ~ 7.3 million years ago (mya), with the *C. scutulatus* group diverging from the *C. viridis* and *C. oreganus* crown group ~ 5.0 mya (Fig. 1B). We estimated that the three species groups (*C. scutulatus*, *C. viridis*, and *C. oreganus*) began to diversify ~ 2.8, ~1.6, and ~ 2.4 mya, respectively (Fig. 1B, Supplementary Fig. S1). The nuclear concatenated phylogeny estimated from 310,000,193 bp across all chromosomes was recovered with strong support at all nodes (Fig. 1C). This topology largely agreed with the coalescent-based tree obtained using ASTRAL (Fig. 1D), with the exception of *C. concolor* and *C. lutosus*, which switch placement between the two trees. In both trees, *C. scutulatus* is sister to the *C. viridis* and *C. oreganus* groups. *Crotalus s. salvini* is recovered within *C. scutulatus*, and *C. v. nuntius* is sister to *C. viridis*. Within the *C. oreganus* group, *C. cerberus* is sister to all other species, and the Intermontane Clade (*C. lutosus*, *C. concolor*, and *C. l. abyssus*) is most closely related to *C. oreganus* and *C. helleri* + *C. h. caliginis* (Fig. 1C, D). Concatenated trees for individual chromosomes were broadly congruent with the tree estimated from all chromosomes, with *C. scutulatus* sister to *C. viridis* and the *C. oreganus* species groups (Supplementary Fig. S2). Overall, trees derived from microchromosomes had more poorly supported nodes (UFB < 95; local posterior probabilities ≤ 0.95) compared to those derived from loci on macrochromosomes and the Z chromosome (Supplementary Fig. S2).

When compared to the all-chromosome concatenated tree, variation in reconstructed evolution histories from specific chromosomes is most apparent by the lack of resolution of the *C. oreganus* group, especially in microchromosomes (Supplementary Fig. S2). In particular, the greatest topological heterogeneity is observed in the Intermontane Clade, where *C. concolor* is either sister to *C. lutosus* and *C. l. abyssus* or embedded within *C. l. abyssus*, depending on the chromosomal region surveyed (Supplementary Fig. S2). We estimated a distinct tree from the Z chromosome compared to trees derived from autosomal loci, with *C. concolor* being sister to the remaining *C. oreganus* group (Fig. 1E versus 1C, D), rather than being in a clade with the other two Intermontane taxa (*C. lutosus* and *C. l. abyssus*).

Similar to concatenated inferences, the coalescent-based species (ASTRAL) trees estimated from microchromosomes vary in support and topology more than the macrochromosomes (Supplementary Fig. S3). Macrochromosomal species trees all have strongly supported relationships. However, unlike the concatenated trees, we find that all macrochromosomes, except macrochromosome 4, support *C. concolor* as sister to the other Intermontane species. This topology was also strongly supported in the coalescent tree based on all chromosomes (Fig. 1D). Microchromosomes and the Z chromosome yield varying support for the evolutionary history of the Intermontane Clade, specifically the placement of *C. lutosus* and *C. concolor*, with microchromosome species trees yielding the lowest node support (Supplementary Fig. S3). We recovered a different topology from the mitochondrial genome dataset, with *C. v. nuntius* nested within *C. viridis*, and *C. helleri* and *C. h. caliginis* strongly supported as sister to the group containing *C. concolor*, *C. lutosus*, *C. l. abyssus* rather than sister to *C. oreganus* (Fig. 1F; Supplementary Fig. S4).

3.2. Genomic variation in concordance factors and tree space

Large datasets often provide strong support values for resolution of nodes, which may obscure evidence for variable topologies across sampled loci (Kumar et al., 2012; Minh et al., 2020a). To assess variation in topology support across the genome, we first analyzed site (sCF) and gene (gCF) concordance factors across our sampled gene tree alignments. Although the concatenated nuclear tree using all chromosomes has UFB = 100 at all nodes, the majority of gene trees do not support this topology (low gCF values; Supplementary Fig. S5). We observed higher sCF values compared to gCF values for macro-, micro-, and the Z chromosome (Fig. 2A, B). Considering all chromosomes, the number of nodes that had gCFs > 50 % ranged from 4 to 5, and the number of nodes that had sCFs > 50 % ranged from 20 to 22 (Supplementary Fig. S5). Microchromosomes showed higher gene concordance overall (ANOVA p-value = 0.025; Tukey's HSD p-value = 0.018). We found no significant differences in gCF values between the Z chromosome and macrochromosomes or microchromosomes, and no significant difference in

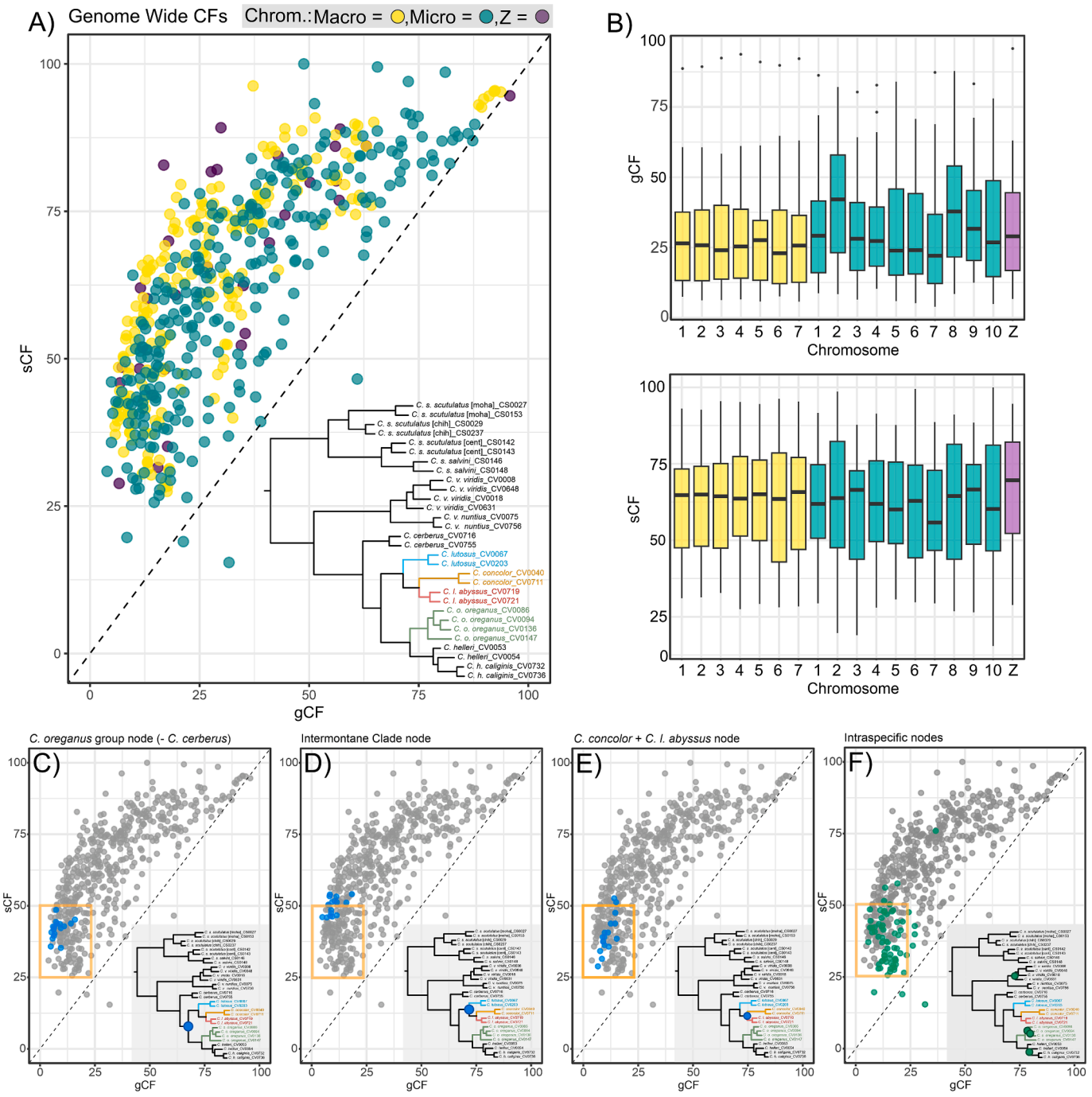


Fig. 2. Gene (gCF) and site (sCF) concordance factors calculated using the CVOS concatenated phylogeny with all chromosomes as a reference tree. A) gCF and sCF plot, with points colored by chromosome type. Each point represents a node in the concatenated tree for each chromosome (29 points per chromosome). Inset tree shows the reference concatenated tree from Fig. 1B. Axes are in percent. B) Boxplots showing the distribution of gCF (top) and sCF (bottom) for each chromosome. Colors of box plots match the chromosome type key in panel A. C–F) gCF/sCF plot from Panel A, highlighting the node at the base of the *C. oreganus* group lineages subtending *C. cerberus* (C), the node at the base of the Intermontane Clade (D), the node representing the split between *C. concolor* and *C. lutosus abyssus* (E), and population-level nodes (F). All inset trees in panels C–F represent the reference tree, with a circle at respective nodes that are highlighted in the panel, and are colored to match the tree in panel A.

sCF values between any of the chromosome types (ANOVA p-value > 0.05). Depending on the specific microchromosome, 4–9 nodes had gCFs > 50 %, and 19–23 nodes had sCF values > 50 %. The Z chromosome had the highest median sCF, as well as 7 and 22 nodes with gCFs > 50 % and sCFs > 50 %, respectively. Additionally, microchromosomes showed a greater range of gCFs and sCFs compared to macrochromosomes (Fig. 2B), suggesting a greater degree of topological variation. When considering all chromosomes, we found that the nodes corresponding to the *C. oreganus* group (excluding *C. cerberus*; Fig. 2C), the Intermontane Clade (Fig. 2D), and the sister relationship between *C. concolor* and *C. l. abyssus* (Fig. 2E) had the lowest gCF values, similar to those observed for intraspecific nodes (Fig. 2F). Our results highlight that chromosomes vary substantially in their degree of discordance (i.e., topological variation). We find that sCF values tended to be higher than gCF values (Fig. 2B). We note, however, that sCF and gCF values are difficult to directly compare. Higher sCF values versus gCF values should

not be directly interpreted as higher discordances across sites versus loci, because these two values are calculated differently and may be influenced differently by characteristics of datasets (Lanfear and Hahn, 2024). We also find that all chromosomes show the greatest discordance for relationships within the *C. oreganus* group, based on particularly low sCF and gCF values for these nodes (Fig. 2C–F).

Our PCoA reveals substantial overlap in tree space across gene trees from distinct chromosomes and chromosome types (Supplementary Fig. S6). We also found high levels of variation in some of these trees across chromosome types, with many gene trees from the Z chromosome occupying the outer margins of observed tree space (Supplementary Fig. S6). Our analysis of coalescent trees estimated from data from distinct chromosomal datasets using probabilistic distances resulted in four groups of phylogenetic similarity, two of which contain both macrochromosomes and microchromosomes (Fig. 3). Two of these groups are comprised only of the Z chromosome or microchromosome 9. The third group includes microchromosomes 3, 4, 5, 7, 10, and macrochromosome 5 (average Hellinger distance [dH] = 0.056), with the fourth group containing all remaining macrochromosome and microchromosome trees and the species tree (all chromosomes) (average dH = 0.015). The coalescent trees inferred for microchromosome 9 and the Z chromosome were substantially different from other trees (average dH of mi9 and Z from all other trees = 0.211 and 0.327) in MDS dimensional space (Fig. 3). These chromosomes each form distinct groups and occupy a unique region in tree space.

3.3. Topological weighting across chromosomes

Topology weighting of loci sampled from across the genome revealed that, when collapsing subtrees into the broader species groups, a speciation scenario in which *C. scutulatus* is sister to the remaining members of the Western Rattlesnake clade (i.e., the *C. viridis* + *C. oreganus* groups) is the dominant signal in the genome (blue tree; Fig. 4A; Supplementary Fig. S7). This dominant topology is also consistent with each phylogenetic analysis using data from all chromosomes (i.e., nuclear data for the concatenated and species tree analyses).

Our Twisst analysis of three alternative topologies for the *C. oreganus* clade identified the topology of the concatenated and coalescent nuclear trees (the Intermontane Clade sister to *C. oreganus* and *C. helleri* + *C. h. caliginis*) as the most common evolutionary history observed genome-wide (Fig. 4A). We observed that windows in the *C. oreganus* clade dataset that recovered the nuclear tree topology were congruent with windows in the species group dataset that also recover the nuclear topology (blue tree and light red tree; Fig. 4A). However, we also find that microchromosomes 9 and 10, and the Z chromosome, contain regions with the highest weights for the alternative mitochondrial topology (Fig. 4A; Supplementary Figs. S7, S8).

Overall, we find that regions associated with lower recombination rates (e.g., centromeres, regions of the Z chromosome) were also associated with higher support for trees that reflect the coalescent (species) tree (Fig. 4A, B). Some chromosomes (macrochromosomes 1, 2, and 3) have regions of low and high recombination rates that correspond with an alternative topology (Fig. 4A). The Z chromosome in both Twisst analyses show greater weights for alternative topologies in the recent stratum and pseudoautosomal region (PAR; Fig. 4B). We also find a similar trend genome-wide, with increasing recombination rate being associated with decreasing weight (frequency) for the species tree in the species group dataset (Fig. 4C). However, this trend is not observed in Twisst analyses of the *C. oreganus* clade dataset on the Z chromosome (Fig. 4B) or genome-wide (Fig. 4C).

Considering evidence in the nuclear genome for an alternative

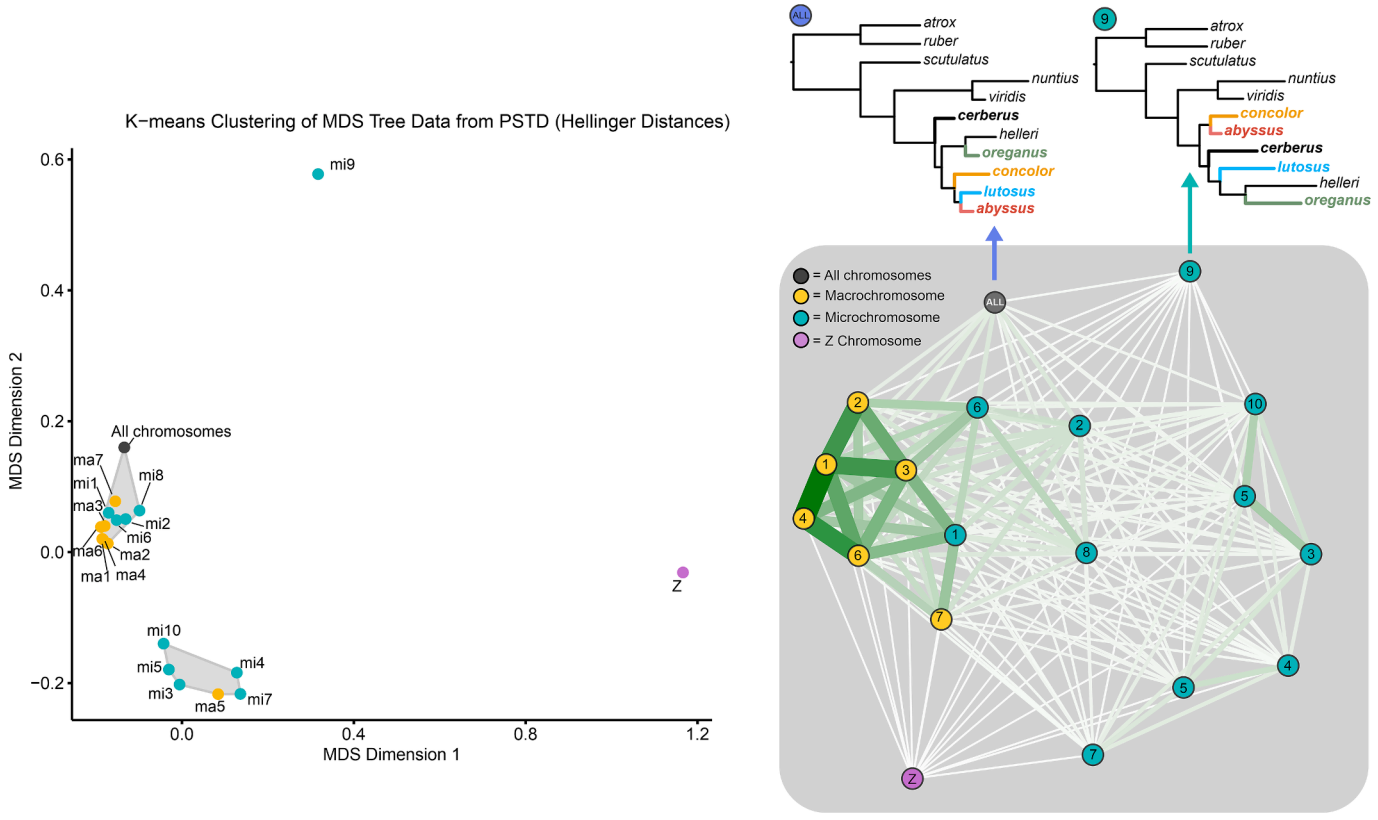
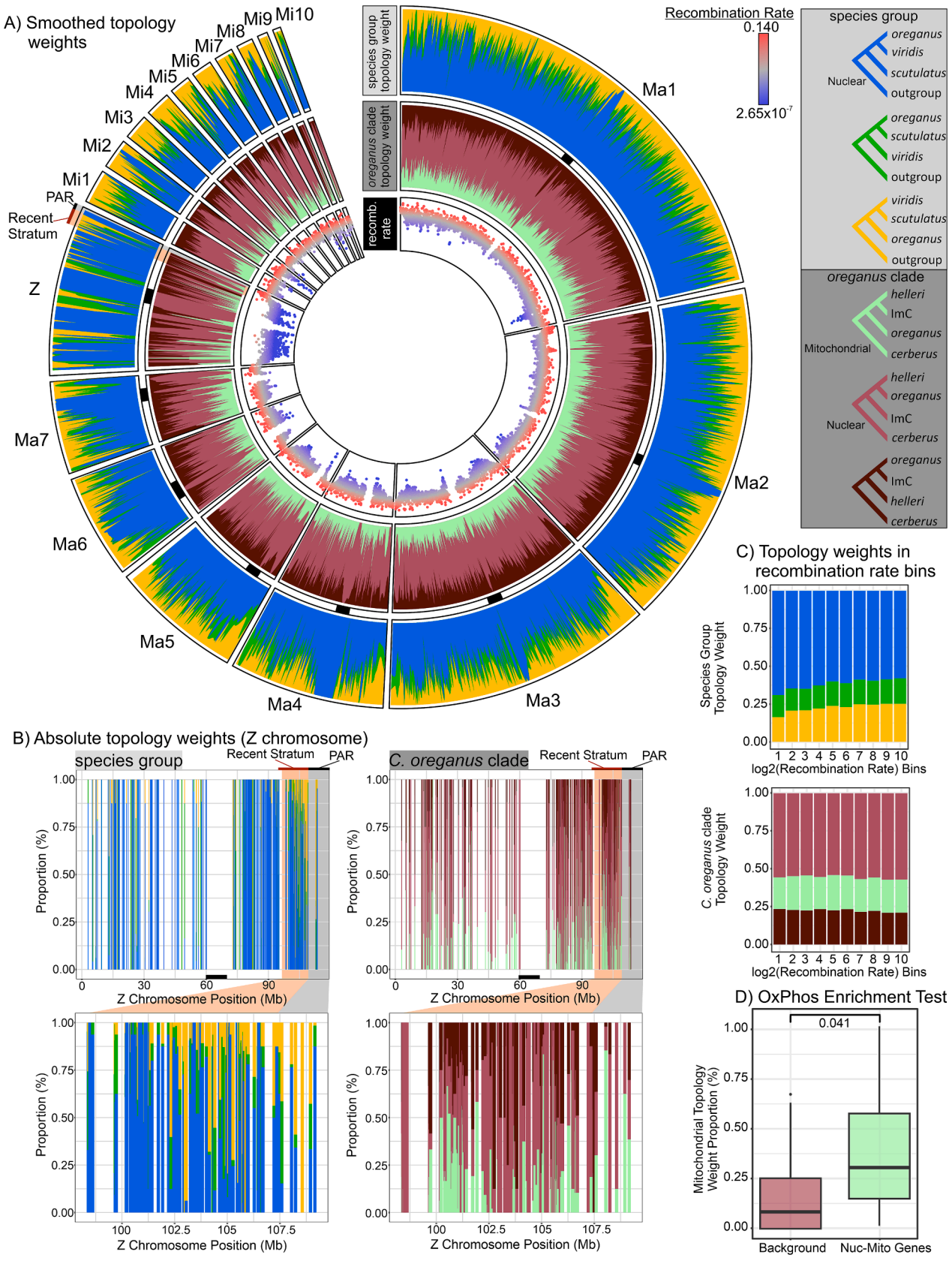


Fig. 3. K-means clustering of Hellinger's distances obtained from PSTD. Chromosome names are shown next to points, which are colored by chromosome type. The grey inset bubble (right) shows a weighted network of Hellinger's distances from PSTD. Nodes are colored based on chromosome type. Connections all represent positive correlation; darker connections represent greater connectivity (association) in the network. Phylogenetic trees are shown for the 'all-chromosome' coalescent tree (the nuclear species tree) and the coalescent tree for microchromosome 9.



(caption on next page)

Fig. 4. Topological weightings across the CVOS genome. A) Circos plot showing LOESS smoothed distribution of topological weights for various trees along with recombination rate variation (inner track; recombination rate scale in upper-right). Outer track: topology weights of a 3-taxon tree with the species groups *C. scutulatus*, *C. viridis*, and *C. oreganus*. Outgroups are *C. atrox* and *C. ruber*. Middle track: topology weights within the *C. oreganus* group of the Intermontane Clade (ImC), *C. oreganus*, and *C. helleri*. *Crotalus cerberus* is used as an outgroup. Black bars between tracks represent centromere regions. Colored topology key on right represents the respective color's topology in the topology weight analysis. Mitochondrial and Nuclear designations in the key represent the mitochondrial and nuclear tree topologies (both species and concatenated), respectively. B) Top: Zoomed in view of absolute (unsmoothed) topology weights (stacked bar plots) of the Z chromosome for the outer and middle tracks. Pink and grey colored columns highlight the recent stratum and pseudoautosomal region (PAR), which are shown in the bottom plots and highlighted in the circos plot in panel A. Black bars represent centromere regions. Colors are identical to the topologies in the key of Panel A. C) Frequency of topology weights of the species group (top) and *C. oreganus* clade (bottom) across 10 equal bins of log2 recombination rates. Colors of bars correspond to the tree key in the upper right. Log2 recombination bins from 1 to 10 are as follows: −21.1 to −19.3, −19.3 to −17.5, −17.5 to −15.7, −15.7 to −13.9, −13.9 to −12, −12 to −10.2, −10.2 to −8.42, −8.42 to −6.6, −6.6 to −4.78, −4.78 to −2.97. D) Enrichment test results for nuclear-encoded mitochondrial genes (Nuc-Mito Genes; e.g., oxidative phosphorylation genes [OxPhos]). Fisher's exact test (enrichment) shows OxPhos genes are enriched for the mitochondrial topology (p-value = 0.041). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

topology that matches the topology estimated from the mitochondrial genome, we tested whether this signal is driven by cyto-nuclear coevolution of nuclear-encoded mitochondrial genes. Among the genomic regions sampled for gene tree estimation, 36 contain a nuclear-encoded mitochondrial gene (Supplementary Fig. S8). Of these, 12 nuclear-encoded mitochondrial genes overlapped with at least one region that had a topology weight of greater than 25 % for the mitochondrial topology. These genes include 9 ATP synthase subunits (ATP) genes (ATP5PB, ATP6V0A4, ATP6V0B, ATP6AP1.1, ATP6V1A, ATP6V1D, ATP6V01D, ATP6V0B.1, ATP6V1H), 4 cytochrome *c* oxidase (COX) genes (COX10.1, COX10, COX11, COX15), 1 NADH dehydrogenase gene (NDUFS6), and 1 ubiquinol-cytochrome *c* reductase core protein 1 (UQCRC1) gene. Notably, the ATP6V1H, ATP6AP1.1, COX10, COX11, and UQCRC1 genes overlap with 10 kb alignments that have topology weights for the mitochondrial tree > 50 %, and the alignment overlapping with COX10 specifically has a mitochondrial tree weight of 100 %. A statistical enrichment test showed that trees reconstructed from 10 kb alignments from nuclear-encoded mitochondrial genes are indeed enriched for the mitochondrial topology (higher weight proportion for the mitochondrial tree) compared to the genomic background (Fisher's exact test, p-value = 0.041; Fig. 4D).

3.4. Evidence of introgression and reticulation

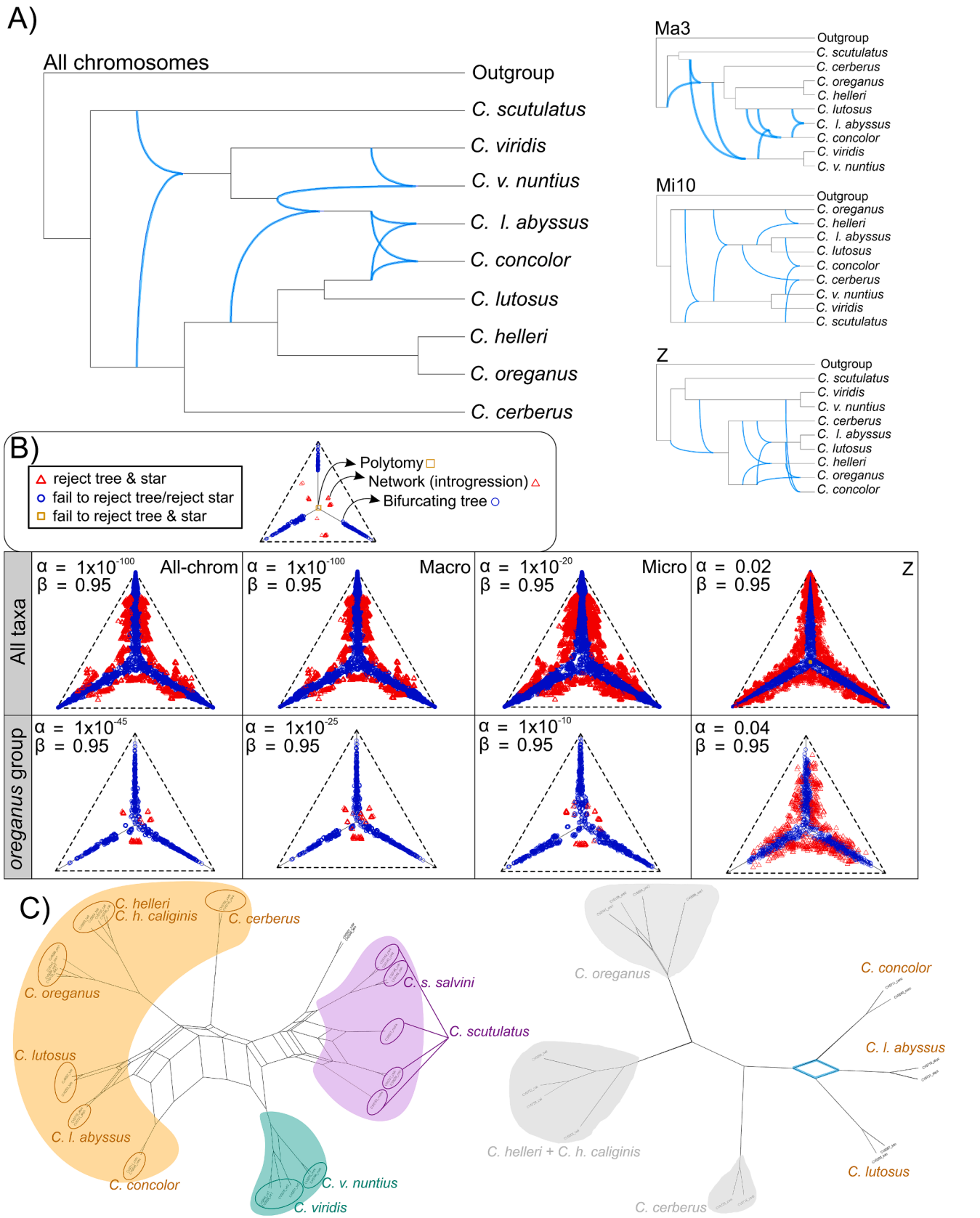
Our PhyloNet analyses provide broad support that reticulate evolution is a pervasive feature of the evolution of the Western Rattlesnake complex (Fig. 5A), with support for 3–5 reticulations (Supplementary Table S2). All chromosome-specific analyses recovered at least 3 reticulations between lineages, but the lineages inferred to share reticulation events varied between chromosomes (Fig. 5A; Supplementary Fig. S9). Many reticulations are observed within the *C. oreganus* group, especially within the Intermontane Clade. These results are corroborated by our NANUQ analyses, where we found pervasive evidence of quartets representing networks rather than trees in our tests using macrochromosomes, microchromosomes, and the Z chromosome (Fig. 5B). Using all taxa, every chromosome contains quartet trees that are better described as a network, with the Z chromosome showing the greatest number of reticulation quartet trees (Fig. 5B). This pattern was also observed when only testing for reticulations in the *C. oreganus* clade. While our NANUQ results for the *C. oreganus* group estimate fewer quartets with reticulations than the analysis with all taxa, the Z chromosome still has the most quartets that support a network compared to the other chromosomes (Fig. 5B). SplitsTree analysis of the quartet matrix with all taxa and all chromosomes revealed reticulations at the base of the tree and between the Intermontane Clade, suggesting both ancient and recent introgression in this group (Fig. 5C). This is also seen in the SplitsTree network for the *C. oreganus* group, in which the highest weighted reticulation is one in which *C. concolor* and *C. lutosus* are the parental lineages of a hybrid *C. l. abyssus* lineage (Fig. 5C).

3.5. Site- and locus-specific likelihood analyses of alternative topologies

To complement our Twisst analysis of variation in topology weights genome-wide, we calculated site-specific likelihoods (SSL) for pairs of competing trees and analyzed these for 10 kb genomic windows. We used this approach to examine the proportion of sites that supported the mitochondrial or Z topology (versus the nuclear concatenated tree) across 10 kb windows. We first examined support for the alternative mitochondrial topology, and found that specific regions of the genome are associated with different levels of support for the mitochondrial tree (Fig. 6). Specifically, we find a larger number of sites within 10 kb windows that support the mitochondrial tree on macrochromosome (ma) 3, microchromosomes (mi) 1, 2, 4, 6, 9, and 10, and towards the end of the Z chromosome nearest to the PAR (green; Fig. 6A). The average proportion of sites supporting the mitochondrial tree across these eight chromosomes ranged from 17.2–22.9 % (average across these eight chromosomes = 20.9 %). This contrasted to all other chromosomes which had an average proportion of sites supporting the mitochondrial tree ranging from 0.8–1.4 % (average across chromosomes = 1.1 %; Fig. 6A).

Given that topology weights of 10 kb alignments containing nuclear-encoded mitochondrial genes were enriched for support for the mitochondrial topology, we tested whether nuclear-encoded mitochondrial loci were also associated with locus-level likelihood support for the mitochondrial topology (based on the sum of SSL across a 10 kb window for the two alternative trees). We focused on the analysis of ma2 and ma3 because they are similar sizes and contain similar numbers of nuclear-encoded mitochondrial genes, yet only ma3 contained a high chromosome-wide fraction of sites that supported the mitochondrial topology (Fig. 6A). Despite the differences in the fraction of sites supporting the mitochondrial tree (Fig. 6A), these two chromosomes do not show substantially different distributions of locus-level likelihood support for the mitochondrial tree, and both show greater densities of support for the nuclear concatenated tree chromosome-wide (Fig. 6B, C). However, we do observe 10 kb windows associated with five nuclear encoded mitochondrial genes (ma3: ATP6V0B; ma2: UQCRC1, COX10, NDUFB11, ATP6V1A; Fig. 6B, C) that do support the mitochondrial tree over the nuclear tree.

Based on the analysis of site-specific likelihoods, we also find widely varying levels of support for the alternative Z chromosome topology across chromosomes (purple; Fig. 6D–F), with most chromosomes falling into one of two classes: chromosomes that show an average of around 29.2 % of sites per 10 kb window supporting the Z tree (e.g., ma1, ma3, ma4, mi1, mi2, mi3, mi5, mi6, mi7, mi9, mi10), and others with an average proportion of sites supporting the Z tree less than 2 % (e.g. ma2, ma5, ma6, ma7, mi1, mi4, mi8). The exception to these patterns is the Z chromosome itself, which shows very high proportions of sites (averaging 78 %, with a range = 67.2–95 %) for the Z topology (Fig. 6D). The Z chromosome is also the only chromosome that had a negative sum of all ΔSSLs for this alternative Z topology, and this support is distributed widely across the Z chromosome. These findings suggest that the evolutionary history of the Z chromosome is indeed distinct from that of



(caption on next page)

Fig. 5. Introgression in the CVOS group. A) Phylogenetic reticulations produced from PhyloNet, using 10 kb window trees from all chromosomes (left), only macrochromosome 3 (upper right) microchromosome 10 (middle right), and the Z chromosome (lower right). Blue edges represent reticulations. B) Simplex plots representing the results of two hypothesis tests in NANUQ, performed on all chromosomes, all macrochromosomes, all microchromosomes, and the Z chromosome datasets; secondary analysis for each dataset was also run on just the *C. oreganus* group (bottom row). Symbols indicate when a quartet rejects a tree & star topology (i.e., represents a network; red triangle), fails to reject a tree and rejects a star topology (i.e., represents bifurcating tree; blue circle), fails to reject a tree but supports a star topology (i.e., polytomy; yellow square), or rejects both a tree and a star topology (green X; not present in plots). Alpha (α) and beta (β) thresholds are given for each simplex plot. Top row = all taxa. Note: PhyloNet trees lump the following taxa: *C. scutulatus* with *C. s. salvini*; *C. helleri* with *C. h. caliginis,* and both Outgroups. C) SplitsTree results of the all-chromosome dataset with all taxa (left) and just the *C. oreganus* group (right). The blue diamond represents a reticulation formed among lineages of *C. concolor* and *C. l. lutosus* that appears to give rise to *C. l. abyssus.* Outgroups in the all-chromosome tree not highlighted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

autosomes, and inferences of a distinct Z chromosome topology are unlikely to be an artifact driven by model misspecifications and extreme site-specific likelihoods.

4. Discussion

Recent increases in available information derived from whole genomes have not necessarily yielded similar magnitudes of phylogenetic resolution, but instead have revealed new complexity and challenges for meaningful interpretation. From one genomic region to the next, numerous processes may influence molecular evolution and associated substitution patterns, leading to varying signal of phylogenetic ancestry. While this poses challenges for phylogenetic inference and systematics, diverse genome-wide signals can provide valuable insight into the distinct evolutionary processes that result in support for alternative topologies. In practice, alternative topologies are often discarded as artifactual or less relevant than inference of a single primary species tree. Our analyses highlight the value of exploring these conflicting signals, with the potential to gain perspectives on the extent to which recombination, gene flow, coevolution, and other evolutionary processes shape genome-wide variation. Our findings further underscore the power of dissecting phylogenetic signal to study these processes at different scales, from specific regions to chromosomes and the entire genome itself.

4.1. Distinct genomic regions harbor conflicting phylogenetic signals

We observed distinct conflicting signals for the relative branching order among lineages within the *C. oreganus* species group from different genomic regions and from different analytical approaches (Fig. 1; Supplementary Figs. S2, S3). Most notably, our phylogenetic reconstructions yielded several strongly supported, yet conflicting relationships within the Intermontane Clade of the *C. oreganus* species group that vary in support between the mitochondrial and nuclear genome, and across chromosomes.

We found that topological heterogeneity and support for alternative topologies are elevated in microchromosomes compared to macrochromosomes. Among classes of chromosomes, microchromosomes also have other distinct characteristics compared to macrochromosomes, including higher recombination rates, higher GC content, and higher densities of genes and euchromatin (Perry et al., 2021; Schield et al., 2019; Schield et al., 2022; Smith et al., 2000; Srikulnath et al., 2021; Waters et al., 2021). These distinct characteristics of microchromosomes, especially their relatively high recombination rate, likely explain higher gene tree variation observed on these chromosomes. Indeed, the hypothesis that higher recombination rates contribute to elevated gene tree heterogeneity in microchromosomes is also supported by our observation that as recombination rate increases genome-wide, the frequency of topology weights supporting an alternative gene tree increased. Generally, recombination acts to unlink genealogies from one another along a chromosome, leading to an expectation of greater gene tree heterogeneity in regions of higher recombination rate (Lanier and Knowles, 2012). Species tree methods typically assume a lack of recombination within a locus and free recombination among loci, and prior work has demonstrated that unrecognized recombination can influence phylogenetic inferences (Mirarab et al., 2024; Posada and Crandall, 2002; Sakoparnig et al., 2021), suggesting gene tree error may contribute some observed variation. However, higher-recombination regions are also more permeable to neutral introgression, suggesting that this elevated gene tree variation may reflect elevated density and variation of introgressed alleles on microchromosomes. Lastly, the high gene density and high recombination of microchromosomes makes them particularly prone to highly effective natural selection (e.g., high density of fitness-critical sites with low linkage). This suggests that variation in the mode, tempo, or direction of natural selection, which can strongly impact gene tree variation and estimation error (Adams et al., 2018; Castoe et al., 2009; Edwards, 2009; Thomas and Hahn, 2015), may also contribute to elevated gene tree variation on microchromosomes.

We also found that the Z chromosome and microchromosome 9 (Fig. 3) differ greatly in overall phylogenetic signal compared to the other chromosomes. However, only the Z chromosome showed support for a clearly differentiated Z topology. Different regions of the Z chromosome are associated with varying levels of recombination and dosage compensation effects (Schield et al., 2019a), with the PAR effectively acting as a highly recombining autosome, and the ancient strata not undergoing recombination with the W. Our results reject the hypothesis that a small set of extreme SSLs (as might be predicted with strong natural selection or convergence) explain support for the alternative Z tree on the Z chromosome. Instead, we found sites chromosome-wide with strong and consistent support for an alternative topology on the Z, consistent with the Z chromosome having a truly distinct evolutionary history from autosomes. There are a few hypotheses that may explain this observed signal. A plausible explanation for this is that ancient episodes of introgression may have resulted in the introgression and fixation of all or most of the Z chromosome. A similar scenario involving ancient introgression of the Z chromosome was inferred to explain broad evidence for a distinct Z chromosome topology in *Heliconius* butterflies (Zhang et al., 2016). An alternative hypothesis is that the combination of drift and selection on the Z chromosome, which has reduced effective population size, has led to greater fixation of alternative phylogenetic signal.

4.2. Mitochondrial versus nuclear trees and evidence for coevolution of nuclear-encoded mitochondrial genes

We find evidence that the mitochondrial tree is distinct from both the majority of nuclear genome signal and coalescent species tree estimates. While concatenated and species tree analyses of entire individual chromosomes do not recover the mitochondrial topology, topology weights indicate that some microchromosomes and regions of the Z chromosome and several macrochromosomes show high support for the mitochondrial topology (Fig. 4A; Supplementary Fig. S8). Both topology weights and locus-specific likelihood analyses indicate that nuclear genomic regions that overlap with nuclear-encoded mitochondrial genes also tend to exhibit elevated signal for the mitochondrial topology. Specifically, we identified several nuclear-encoded mitochondrial genes with strong support of the mitochondrial topology, including those that encode components of the ATP synthase, cytochrome *c* oxidase, and NADH dehydrogenase complex (topology weights for the mitochondrial topology = 50–100 %; Fig. 4D; Supplementary Fig. S8). Such strong

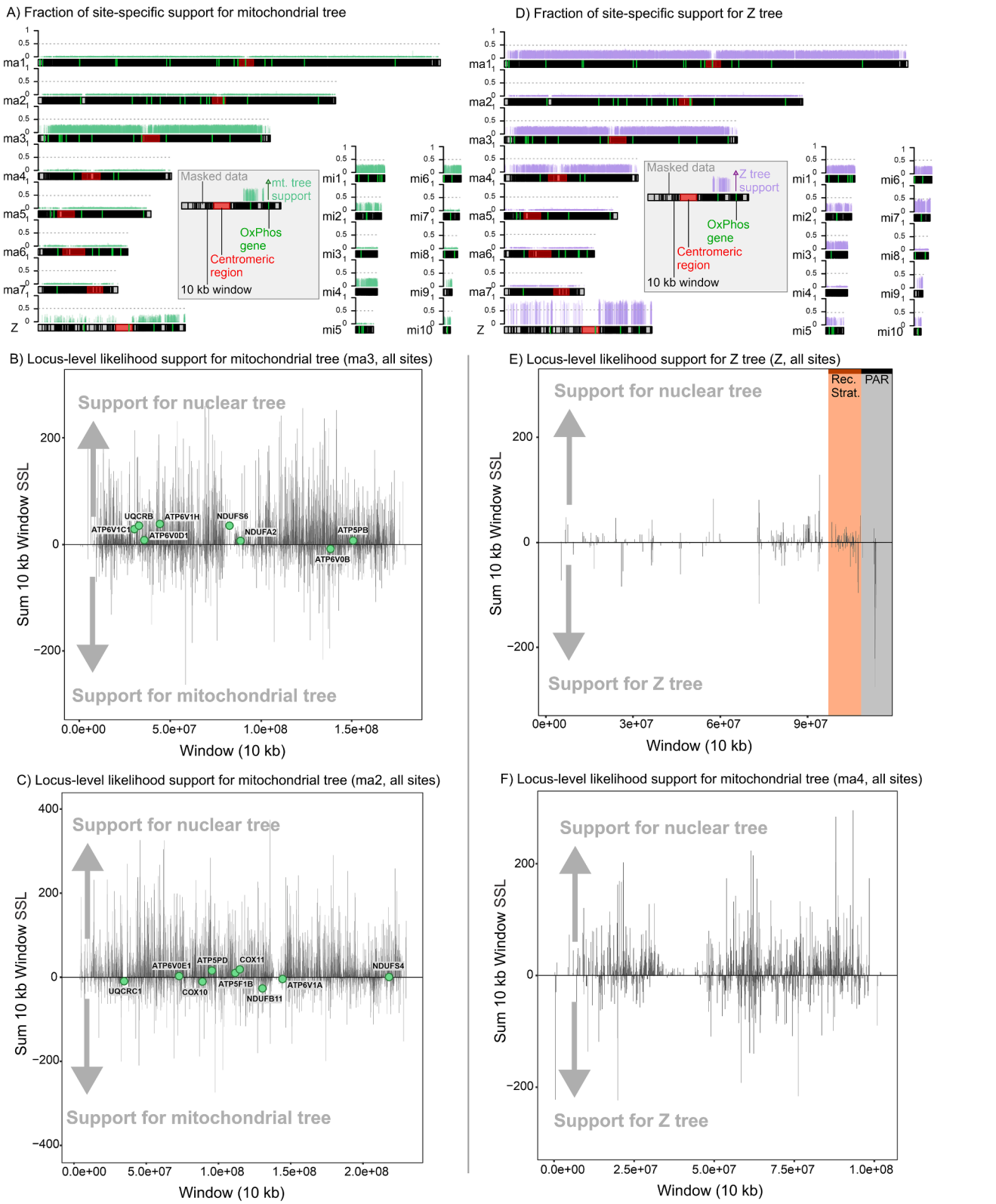


Fig. 6. Difference in site-specific likelihoods for the nuclear concatenated tree and mitochondrial (panels A–C) and Z (panels D–F) topologies. Each chromosome is represented as a horizontal bar; grey regions = regions with masked data, red = centromere, green = nuclear-encoded mitochondrial genes. The colored, vertical bars above each chromosome show the proportion of summed ΔSSL supporting the alternative topology (mitochondrial topology for panel A; Z topology for panel D); each bar represents one 10 kb window. Grey dashed horizontal line is shown for visualization at proportion = 0.5. Chromosome names are shown to the left of the chromosomes. Below each panel shows the ΔSSL summed over 10 kb windows (ma3 and ma2 for panel B and C; Z chromosome and ma4 for panels E and F). The Y-axis position of oxidative phosphorylation (OxPhos) genes in panels B and C represents the summed ΔSSL of all sites in the respective gene (positive = support for nuclear tree; negative = support for mitochondrial tree). Summed ΔSSL values over 10 kb windows above and below 0 in panels B–C and E–F show support for the nuclear species tree or the alternative topology, respectively. The recent stratum (Rec. Strat.) and pseudoautosomal region (PAR) of the Z chromosome are shown in panel E. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

patterns of concordance are consistent with a scenario of introgression of the mitochondria and subsequent coevolution of nuclear-encoded mitochondrial gene alleles to match the introgressed mitochondrial genome. Our findings likely reflect the strong coevolutionary trajectories of nuclear- and mitochondrial-encoded oxidative phosphorylation proteins (Barreto et al., 2018; Ben Slimen et al., 2018; Gershoni et al., 2010; Morales et al., 2015; Piccinini et al., 2021; Sunnucks et al., 2017; Zhou et al., 2020), which has also been suggested to have occurred in other systems based on phylogenetic evidence similar to what we observed in our study (Formaggioni et al., 2022).

Our analyses of introgression and reticulate evolution suggest that introgression is a widespread phenomenon within this group and likely contributes to genome-wide variation in gene trees. Phylogenetic networks from PhyloNet and NANUQ both indicate the existence of reticulations towards the tips and base of the tree, suggesting the roles of both recent (and likely ongoing) and ancient introgression events (Fig. 5; Supplementary Fig. S9). PhyloNet also identified reticulation events consistent with evidence of introgression among these lineages from prior studies, including evidence of contemporary introgression between *C. viridis* and *C. scutulatus* (Glenn and Straight, 1990; Pook et al., 2000; Schield et al., 2019; Zancolli et al., 2016), and among members of the *C. oreganus* group (Myers, 2021; Nikolakis et al., 2022; Schield et al., 2019b; Smith and Mackessy, 2016). Interestingly, tests for reticulations using quartet trees in NANUQ indicate that most reticulation signatures for the *C. oreganus* group are concentrated on the Z chromosome, suggesting high levels of introgression on this chromosome. This is consistent with evidence that the Z chromosome harbors a distinct, alternative, but real phylogenetic history compared to autosomes that may be driven in large part by introgression across the Z.

Relevant to our inferences of introgression and reticulation, our analyses also suggest that the taxon *C. l. abyssus* may represent a population derived from hybridization between *C. concolor* and *C. lutosus* (Fig. 5); alternatively, this signal could reflect historical introgression in the Intermontane Clade. These three taxa comprise the Intermontane Clade, within which we observed substantial genome-wide topological heterogeneity. These findings shed new light on the origins of this population and highlights the potential for such whole-genome data for identifying complex evolutionary scenarios involving introgression.

Genome-wide, recombination rate tends to be inversely correlated with topology weights for the coalescent species tree (Fig. 4C), indicating that high recombination areas tend to harbor greater gene tree diversity and signal of alternative trees. Generally, this is consistent with the fact that recombination acts to break up haplotypes, disrupting linkage and enabling a greater maintenance of ancestral polymorphism, which influences gene tree estimation. Considering the strong evidence for widespread introgression in this group, it is also notable that high recombination regions are expected to be more prone to introgression (Martin et al., 2019). Elevated gene tree variation in higher recombination regions may therefore also be the result of a greater density of introgressed alleles at detectable frequencies. We also find evidence for specific, narrow genomic regions which show dramatic shifts in topology weights (as well as site-specific likelihoods), indicating strong support for an alternative tree. These regions, some of which exhibit high recombination rates (Fig. 4A), may have undergone adaptive introgression (Martin and Van Belleghem, 2017). Together, our results highlight how introgression, recombination, and natural selection impact variation in the evolutionary history across regions of the genome, with downstream impacts on phylogenetic inference at genomic scales.

4.3. The phylogeny of the Western Rattlesnake species complex and future studies to address taxonomy

The taxonomy of the Western Rattlesnake species complex has been controversial and unstable for several decades (Campbell and Lamar, 2004; Davis et al., 2016; Douglas et al., 2002; Pook et al., 2000). High levels of phenotypic variation (Smith and Mackessy, 2016; Watson et al., 2019; Zancolli et al., 2016) combined with widespread hybridization (Meik et al., 2008; Nikolakis et al., 2022; Pook et al., 2000; Schield et al., 2019b) have played major roles in the phylogenetic instability of this group, and, not surprisingly, have made species-level relationships and taxonomy challenging to resolve. Recent studies have elevated many subspecies to species status, recognizing *C. scutulatus*, *C. viridis*, *C. oreganus*, *C. cerberus*, *C. helleri*, *C. concolor*, and *C. lutosus* as distinct lineages (Davis et al. 2016). Other studies have suggested the recognition of some populations as subspecies (e.g., *C. v. nuntius* as a subspecies of *C. viridis*, *C. h. caliginis* as an insular population of *C. helleri*, and *C. l. abyssus* as a stunted form of *C. lutosus*; (Davis et al., 2016; Pook et al., 2000).

Although we find evidence of conflicting phylogenetic signals to varying extents across the genome, we find that most of the nuclear genome supports a single majority species tree (i.e., the nuclear tree with all chromosomes; Fig. 1C, D; Supplementary Figs. S1, S3). We also find that the nuclear coalescent versus concatenated trees differ only slightly in their placement of *C. l. abyssus*, which based on reticulation analyses appears to be derived from hybridization between *C. lutosus* and *C. concolor* (Fig. 5; Supplementary Fig. S9). This finding is also not surprising, considering that the ranges of *C. lutosus* and *C. concolor* extend to the northern and eastern sides of the Grand Canyon, and "*abyssus*" is the name applied to all Western Rattlesnakes within the Grand Canyon mostly because of geographic province. Based on our findings, we suggest that *C. l. abyssus* be considered a hybrid population of *C. lutosus* and *C. concolor*, rather than a distinct subspecies or a synonym of another taxon. We do note, however, that our analyses are based on only two samples, and additional samples that confirm the hypothesis that "*abyssus*" is indeed a hybrid population would be useful to verify our conclusions. Aside from this case, it is notable that our inferences of relationships among major lineages do generally support the current taxonomy, and further indicate varying levels of recent and/or historical introgression among lineages has occurred. Future studies that incorporate additional (and more geographically dispersed) sampling per lineage would be important for understanding the range-wide extent of such introgression, and for informing decisions for further taxonomic revisions. Over the past decades, as we have collected greater resolution molecular datasets, we have found that hybridization is far more pervasive in nature than traditionally assumed (Payseur and Rieseberg, 2016), which will require reconsideration of how we incorporate degrees of hybridization into practical species delimitation (Abbott et al., 2013; Harrison and Larson, 2014). However, despite clear signals of hybridization we observe in Western Rattlesnakes, our findings also indicate that these lineages retain substantial portions of their divergent, unique genomes, which argues strongly for their recognition as distinct taxa.

As the size of molecular datasets become increasingly larger, the challenges of understanding how to interpret conflicting phylogenetic signal have increased by similar orders of magnitude. Our results demonstrate an extreme example of such conflict, with disparate topologies recovered from different genomes (mitochondrial vs. nuclear), different chromosomes, and from distinct regions of the same chromosomes. Our findings highlight the value in dissecting the sources of such conflicts to understand their distribution across large datasets and provide insight into why such conflict may exist. In our case, these conflicting phylogenetic signals appear to represent evidence of past introgression and may relate to signatures of selection, including those related to cytonuclear incompatibilities. Importantly, these approaches to dissect different signals also provided more insight into the primary underlying species tree (e.g., the tree supported by the majority of the genome), rather than be misled by other patterns and processes from disparate evolutionary histories present in some regions of the genome.

5. Conclusion

In this study, we highlight the power and the challenges of using whole genomes for inferring evolutionary histories. We show that whole genomes may not provide a single simple answer about the evolutionary history of a particular radiation, but instead can provide many well-supported and biologically meaningful conclusions about complex events associated with the evolutionary history of species. Our results support that introgression and natural selection have contributed to a mosaic of phylogenetic signal and evolutionary histories in this group of snakes. This complexity, together with substantial phenotypic and ecological variation among Western Rattlesnake lineages, has undoubtedly contributed to disagreement regarding relationships within this group as well as broad taxonomic instability. It is evident that future taxonomic decisions will be most effective if they embrace, rather than ignore, the variation and complexity within this species complex.

6. Data statement

Raw shotgun genomic sequencing read data are available on the Sequence Read Archive (SRA) under Project PRJNA1150930 and PRJNA1137891; all sample information can be found in Supplementary Table S1. Mitochondrial reference genome is available on GenBank (accession number: PV035282); GenBank accession numbers of individual mitochondrial genes for all samples are available in Supplementary Table S3. Code and input files for all analyses can be found on Github (https://github.com/jbernst) and Zenodo (DOI: 10.5281/zenodo.14774372).

CRediT authorship contribution statement

Justin M. Bernstein: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. **Yannick Z. Francioli:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Drew R. Schield:** Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Formal analysis. **Richard H. Adams:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis. **Blair W. Perry:** Writing – review & editing, Resources. **Keaka Farleigh:** Writing – review & editing, Resources. **Cara F. Smith:** Writing – review & editing. **Jesse M. Meik:** Writing – review & editing. **Stephen P. Mackessy:** Writing – review & editing, Resources. **Todd A. Castoe:** Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Funding

Support for this work was provided by National Science Foundation grants DEB-1655571 to T.A.C., J.M.M. and S.P.M., IOS-2307044 to T.A. C, S.P.M. and R.H.A. JMB was supported by an NSF Postdoctoral Research Fellowships in Biology Program (NSF-DEB 2208959). KF was supported by an NSF Postdoctoral Research Fellowship in Biology (DBI-2409958).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the following people and institutions for contributions of tissues: Jens Vindum and Rayna Bell at the California Academy of Sciences, Elda Sanchez at the National Natural Toxin Research Center, Robert Murphy at the Royal Ontario Museum, Carol Spencer and Jimmy McGuire at the Museum of Vertebrate Zoology, and Christopher Parkinson, Rulon Clark, Gordon Schuett, Doug Duerre, Brandon La Forest, Corey Roelke, Andrew Holycross, Hans-Werner Herrmann, Joshua Parker, and Stephen Spear.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2025.108309.

Data availability

Raw sequence data is available on the Sequence Read Archive (Project PRJNA1150930 and PRJNA1137891). Mitochondrial reference genome is available on GenBank (accession PV035282); GenBank accession numbers for mitochondrial genes of all samples are available in Supplementary Table S3. Analysis and associated code is available at https://github.com/jbernst.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.

Adams, R.H., Castoe, T.A., 2020. Probabilistic Species tree distances: implementing the multispecies coalescent to compare species trees within the same model-based framework used to estimate them. Sys. Biol. 69, 194–207. https://doi.org/10.1093/sysbio/syz031.

Adams, R.H., Schield, D.R., Card, D.C., Castoe, T.A., 2018. Assessing the impacts of positive selection on coalescent-based species tree estimation and species delimitation. Syst. Biol. 67, 1076–1090. https://doi.org/10.1093/sysbio/syy034.

Allman, E.S., Baños, H., Rhodes, J.A., 2019. NANUQ: a method for inferring species networks from gene trees under the coalescent model. Algorithms Mol. Biol. 14, 24. https://doi.org/10.1186/s13015-019-0159-2.

Arab, M.A., zu Siederdissen, C.H., Tout, K., Sahyoun, A.H., Stadler, P.F., Bernt, M., 2017. Accurate annotation of protein-coding genes in mitochondrial genomes. Mol. Phylogenet. Evol. 106, 209–216. https://doi.org/10.1016/j.ympev.2016.09.024.

Ashton, K.G., 2001. Body size variation among mainland populations of the western rattlesnake (*Crotalus viridis*). Evolution 55, 2523–2533. https://doi.org/10.1111/j.0014-3820.2001.tb00766.x.

Ashton, K.G., Queiroz, A.D., 2001. Molecular systematics of the Western Rattlesnake, *Crotalus viridis* (Viperidae), with comments on the utility of the D-Loop in Phylogenetic studies of snakes. Mol. Phylogenet. Evol. 21, 176–189. https://doi.org/10.1006/mpev.2001.1013.

Barreto, F.S., Watson, E.T., Lima, T.G., Willett, C.S., Edmands, S., Li, W., Burton, R.S., 2018. Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. Nat. Ecol. Evol. 2, 1250–1257. https://doi.org/10.1038/s41559-018-0588-1.

Ben Slimen, H., Awadi, A., Tolesa, Z.G., Knauer, F., Alves, P.C., Makni, M., Suchentrunk, F., 2018. Positive selection on the mitochondrial *ATP synthase 6* and the *NADH dehydrogenase 2* genes across 22 hare species (genus *Lepus*). J. Zool. Syst. Evol. Res. 56, 428–443. https://doi.org/10.1111/jzs.12204.

Blom, M.P.K., Bragg, J.G., Potter, S., Moritz, C., 2016. Accounting for uncertainty in gene tree estimation: summary-coalescent species tree inference in a challenging radiation of australian lizards. Syst. Biol. https://doi.org/10.1093/sysbio/syw089.

Campbell, J.A., Lamar, W.W., 2004. The Venomous Reptiles of the Western Hemisphere. Cornell University Press, Ithaca.

Castoe, T.A., Spencer, C.L., Parkinson, C.L., 2007. Phylogeographic structure and historical demography of the western diamondback rattlesnake (*Crotalus atrox*): a perspective on North American desert biogeography. Mol. Phylogenet. Evol. 42, 193–212. https://doi.org/10.1016/j.ympev.2006.07.002.

Castoe, T.A., de Koning, A.P.J., Kim, H.-M., Gu, W., Noonan, B.P., Naylor, G., Jiang, Z.J., Parkinson, C.L., Pollock, D.D., 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. Proc. Natl. Acad. Sci. u.s.a. 106, 8986–8991. https://doi.org/10.1073/pnas.0900233106.

Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information., in: German Conference on Bioinformatics. Citeseer, pp. 45–56.

Cloutier, A., Sackton, T.B., Grayson, P., Clamp, M., Baker, A.J., Edwards, S.V., 2019. Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the presence of an empirical anomaly zone. Syst. Biol. 68, 937–955. https://doi.org/10.1093/sysbio/syz019.

Davis, M.A., Douglas, M.R., Collyer, M.L., Douglas, M.E., 2016. Deconstructing a species-complex: geometric morphometric and molecular analyses define species in the Western Rattlesnake (*Crotalus viridis*). PLoS One 11, e0146166. https://doi.org/10.1371/journal.pone.0146166.

Donath, A., Jühling, F., Al-Arab, M., Bernhart, S.H., Reinhardt, F., Stadler, P.F., Middendorf, M., Bernt, M., 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res. 47, 10543–10552. Doi: 10.1093/nar/gkz833.

Douglas, M.E., Douglas, M.R., Schuett, G.W., Porras, L.W., Holycross, A.T., 2002. Phylogeography of the western rattlesnake (*Crotalus viridis*) complex, with emphasis on the Colorado Plateau. In: Schuett, G.W., Hoggren, M., Douglas, M.E., Greene, H. W. (Eds.), Biology of the Vipers. Eagle Mountain Publishing, Eagle Mountain (UT), pp. 11–50.

Edwards, S.V., 2009. Natural selection and phylogenetic analysis. Proc. Natl. Acad. Sci. u.s.a. 106, 8799–8800. https://doi.org/10.1073/pnas.0904103106.

Elworth, R., Ogilvie, H.A., Zhu, J., Nakhleh, L., 2019. Advances in computational methods for phylogenetic networks in the presence of hybridization. In: Warnow, T. (Ed.), Bioinformatics and Phylogenetics. Springer, Cham, pp. 317–360.

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. https://doi.org/10.1093/sysbio/sys004.

Formaggioni, A., Plazzi, F., Passamonti, M., 2022. Mito-nuclear coevolution and phylogenetic artifacts: the case of bivalve mollusks. Sci. Rep. 12, 11040. https://doi.org/10.1038/s41598-022-15076-y.

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

Gel, B., Serra, E., 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33, 3088–3090. https://doi.org/10.1093/bioinformatics/btx346.

Gershoni, M., Fuchs, A., Shani, N., Fridman, Y., Corral-Debrinski, M., Aharoni, A., Frishman, D., Mishmar, D., 2010. Coevolution predicts direct interactions between mtDNA-encoded and nDNA-encoded subunits of oxidative phosphorylation complex I. J. Mol. Biol. 404, 158–171. https://doi.org/10.1016/j.jmb.2010.09.029.

Glenn, J.L., Straight, R.C., 1990. Venom characteristics as an indicator of hybridization between *Crotalus viridis viridis* and *Crotalus scutulatus scutulatus* in New Mexico. Toxicon 28, 857–862. https://doi.org/10.1016/S0041-0101(09)80008-1.

Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res. 41, e129.

Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. J. Hered. 105, 795–809. https://doi.org/10.1093/jhered/esu033.

Hibbins, M.S., Hahn, M.W., 2022. Phylogenomic approaches to detecting and characterizing introgression. Genetics 220, iyab173. https://doi.org/10.1093/genetics/iyab173.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the Ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.

Holman, J.A., 2000. Fossil snakes of North America: origin, evolution, distribution, paleoecology. Indiana University Press, Bloomington.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. https://doi.org/10.1093/molbev/msj030.

Jiang, X., Edwards, S.V., Liu, L., 2020. The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. Syst. Biol. 69, 795–812. https://doi.org/10.1093/sysbio/syaa008.

Jiao, X., Flouri, T., Yang, Z., 2021. Multispecies coalescent and its applications to infer species phylogenies and cross-species gene flow. Natl. Sci. Rev. 8, nwab127. https://doi.org/10.1093/nsr/nwab127.

Jombart, T., Kendall, M., Almagro-Garcia, J., Colijn, C., 2017. treespace : Statistical exploration of landscapes of phylogenetic trees. Mol. Ecol. Resour. 17, 1385–1392. https://doi.org/10.1111/1755-0998.12676.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods. 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Kanehisa, M., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic. Acids. Res. 28, 27–30. https://doi.org/10.1093/nar/28.1.27.

Kanehisa, M., 2019. Toward understanding the origin and evolution of cellular organisms. Prot. Sci. 28, 1947–1951. https://doi.org/10.1002/pro.3715.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 51, D587–D592. https://doi.org/10.1093/nar/gkac963.

Klauber, L.M., 1972. Rattlesnakes: Their Habits, Life Histories, and Influence on Mankind, second ed. University of California Press, Berkeley, CA.

Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L., Tamura, K., 2012. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29, 457–472. https://doi.org/10.1093/molbev/msr202.

Lanfear, R., Hahn, M.W., 2024. The meaning and measure of concordance factors in phylogenomics. MBE. 41, msae214. https://doi.org/10.1093/molbev/msae214.

Lanier, H.C., Knowles, L.L., 2012. Is recombination a problem for species-tree analyses? Syst. Biol. 61, 691–701. https://doi.org/10.1093/sysbio/syr128.

Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61, 727–744. https://doi.org/10.1093/sysbio/sys049.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Lin, H.-Y., Hao, Y.-J., Li, J.-H., Fu, C.-X., Soltis, P.S., Soltis, D.E., Zhao, Y.-P., 2019. Phylogenomic conflict resulting from ancient introgression following species diversification in Stewartia s.l. (Theaceae). Mol. Phylogenet. Evol. 135, 1–11. https://doi.org/10.1016/j.ympev.2019.02.018.

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

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

Mackessy, S.P., 2010. Evolutionary trends in venom composition in the Western Rattlesnakes (*Crotalus viridis* sensu lato): toxicity versus tenderizers. Toxicon 55, 1463–1474. https://doi.org/10.1016/j.toxicon.2010.02.028.

Martin, S.H., Davey, J.W., Salazar, C., Jiggins, C.D., 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17, e2006288. https://doi.org/10.1371/journal.pbio.2006288.

Martin, S.H., Van Belleghem, S.M., 2017. Exploring evolutionary relationships across the genome using topology weighting. Genetics 206, 429–438. https://doi.org/10.1534/genetics.116.194720.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.

Meik, J.M., Fontenot, B.E., Franklin, C.J., King, C., 2008. Apparent natural hybridization between the rattlesnakes *Crotalus atrox* and *C. horridus*. The Southw. Naturalist 53, 196–200. https://doi.org/10.1894/0038-4909(2008)53[196:ANHBTR]2.0.CO;2.

Meleshko, O., Martin, M.D., Korneliussen, T.S., Schröck, C., Lamkowski, P., Schmutz, J., Healey, A., Piatkowski, B.T., Shaw, A.J., Weston, D.J., Flatberg, K.I., Szövényi, P., Hassel, K., Stenøien, H.K., 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus. Mol. Biol. Evol. 38, 2750–2766. https://doi.org/10.1093/molbev/msab063.

Mendes, F.K., Hahn, M.W., 2018. Why concatenation fails near the anomaly zone. Syst. Biol. 67, 158–169. https://doi.org/10.1093/sysbio/syx063.

Minh, B.Q., Hahn, M.W., Lanfear, R., 2020a. New methods to calculate concordance factors for phylogenomic datasets. Mol. Bio. Evol. 37, 2727–2733. https://doi.org/10.1093/molbev/msaa106.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020b. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

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. https://doi.org/10.1093/sysbio/syu063.

Mirarab, S., Rivas-González, I., Feng, S., Stiller, J., Fang, Q., Mai, U., Hickey, G., Chen, G., Brajuka, N., Fedrigo, O., Formenti, G., Wolf, J.B.W., Howe, K., Antunes, A., Schierup, M.H., Paten, B., Jarvis, E.D., Zhang, G., Braun, E.L., 2024. A region of suppressed recombination misleads neoavian phylogenomics. Proc. Natl. Acad. Sci. U.S.A. 121, e2319506121. https://doi.org/10.1073/pnas.2319506121.

Mo, Y.K., Lanfear, R., Hahn, M.W., Minh, B.Q., 2023. Updated site concordance factors minimize effects of homoplasy and taxon sampling. Bioinformatics 39, btac741. https://doi.org/10.1093/bioinformatics/btac741.

Morales, H.E., Pavlova, A., Joseph, L., Sunnucks, P., 2015. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. Mol. Ecol. 24, 2820–2837. https://doi.org/10.1111/mec.13203.

Myers, E.A., 2021. Genome-wide data reveal extensive gene flow during the diversification of the western rattlesnakes (Viperidae: Crotalinae: *Crotalus*). Mol. Phylogenet. Evol. 165, 107313. https://doi.org/10.1016/j.ympev.2021.107313.

Myers, E.A., Rautsaw, R.M., Borja, M., Jones, J., Grünwald, C.I., Holding, M.L., Grazziotin, F.G., Parkinson, C.L., 2024. Phylogenomic discordance is driven by widespread introgression and incomplete lineage sorting during rapid species diversification within rattlesnakes (Viperidae: *Crotalus* and *Sistrurus*). Syst. Biol. syae018. https://doi.org/10.1093/sysbio/syae018.

Nikolakis, Z.L., Schield, D.R., Westfall, A.K., Perry, B.W., Ivey, K.N., Orton, R.W., Hales, N.R., Adams, R.H., Meik, J.M., Parker, J.M., Smith, C.F., Gompert, Z., Mackessy, S.P., Castoe, T.A., 2022. Evidence that genomic incompatibilities and other multilocus processes impact hybrid fitness in a rattlesnake hybrid zone. Evolution 76, 2513–2530. https://doi.org/10.1111/evo.14612.

Oliver, J.C., 2013. Microevolutionary processes generate phylogenomic discordance at ancient divergences. Evolution 67, 1823–1830. https://doi.org/10.1111/evo.12047.

Parins-Fukuchi, C., Stull, G.W., Smith, S.A., 2021. Phylogenomic conflict coincides with rapid morphological innovation. Proc. Natl. Acad. Sci. U.S.A 118, e2023058118. https://doi.org/10.1073/pnas.2023058118.

Payseur, B.A., Rieseberg, L.H., 2016. A genomic perspective on hybridization and speciation. Mol. Ecol. 25, 2337–2360. https://doi.org/10.1111/mec.13557.

Perry, B.W., Schield, D.R., Adams, R.H., Castoe, T.A., 2021. Microchromosomes exhibit distinct features of vertebrate chromosome structure and function with underappreciated ramifications for genome evolution. Mol. Bio. Evol. 38, 904–910. https://doi.org/10.1093/molbev/msaa253.

Piccinini, G., Iannello, M., Puccio, G., Plazzi, F., Havird, J.C., Ghiselli, F., 2021. Mitonuclear coevolution, but not nuclear compensation, drives evolution of OXPHOS complexes in bivalves. Mol. Biol. Evol. 38, 2597–2614. https://doi.org/10.1093/molbev/msab054.

Pook, C.E., Wüster, W., Thorpe, R.S., 2000. historical biogeography of the western rattlesnake (Serpentes: Viperidae: *Crotalus viridis*), inferred from mitochondrial dna sequence information. Mol. Phylogenet. Evol. 15, 269–282. https://doi.org/10.1006/mpev.1999.0756.

Posada, D., Crandall, K.A., 2002. The effect of recombination on the accuracy of phylogeny estimation. J. Mol. Evol. 54, 396–402. https://doi.org/10.1007/s00239-001-0034-9.

Rambaut, A., 2012. Molecular Evolution, Phylogenetics and Epidemiology FigTree v1, 4.

Reilly, S.B., Stubbs, A.L., Karin, B.R., Bi, K., Arida, E., Iskandar, D.T., McGuire, J.A., 2019. Leap-frog dispersal and mitochondrial introgression: phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. J. Biogeogr. 46, 757–769. https://doi.org/10.1111/jbi.13526.

Rhodes, J.A., Baños, H., Mitchell, J.D., Allman, E.S., 2021. MSCquartets 1.0: quartet methods for species trees and networks under the multispecies coalescent model in R. Bioinformatics 37, 1766–1768. https://doi.org/10.1093/bioinformatics/btaa868.

Roch, S., Warnow, T., 2015. On the robustness to gene tree estimation error (or lack thereof) of coalescent-based species tree methods. Syst. Biol. 64, 663–676. https://doi.org/10.1093/sysbio/syv016.

Roycroft, E.J., Moussalli, A., Rowe, K.C., 2020. Phylogenomics uncovers confidence and conflict in the rapid radiation of Australo-Papuan rodents. Syst. Biol. 69, 431–444. https://doi.org/10.1093/sysbio/syz044.

Sakoparnig, T., Field, C., Van Nimwegen, E., 2021. Whole genome phylogenies reflect the distributions of recombination rates for many bacterial species. eLife 10, e65366. https://doi.org/10.7554/eLife.65366.

Sanderson, B.J., Gambhir, D., Feng, G., Hu, N., Cronk, Q.C., Percy, D.M., Freaner, F.M., Johnson, M.G., Smart, L.B., Keefover-Ring, K., Yin, T., Ma, T., DiFazio, S.P., Liu, J., Olson, M.S., 2023. Phylogenomics reveals patterns of ancient hybridization and differential diversification that contribute to phylogenetic conflict in willows, poplars, and close relatives. Syst. Biol. 72, 1220–1232. https://doi.org/10.1093/sysbio/syad042.

Schield, D.R., Adams, R.H., Card, D.C., Corbin, A.B., Jezkova, T., Hales, N.R., Meik, J.M., Perry, B.W., Spencer, C.L., Smith, L.L., 2018. Cryptic genetic diversity, population structure, and gene flow in the Mojave Rattlesnake (*Crotalus scutulatus*). Mol. Phylogenet. Evol. 127, 669–681. https://doi.org/10.1016/j.ympev.2018.06.013.

Schield, D.R., Card, D.C., Hales, N.R., Perry, B.W., Pasquesi, G.I.M., Blackmon, H., Adams, R.H., Corbin, A.B., Smith, C.F., Ramesh, B., Demuth, J.P., Betran, E., Tollis, M., Meik, J.M., Mackessy, S.P., Castoe, T.A., 2019a. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. Genome Res. 29, 590–601.

Schield, D.R., Perry, B.W., Adams, R.H., Card, D.C., Jezkova, T., Pasquesi, G.I.M., Nikolakis, Z.L., Row, K., Meik, J.M., Smith, C.F., 2019b. Allopatric divergence and secondary contact with gene flow: a recurring theme in rattlesnake speciation. Biol. J. Linn. Soc. Lond. 128, 149–169. https://doi.org/10.1093/biolinnean/blz077.

Schield, D.R., Pasquesi, G.I.M., Perry, B.W., Adams, R.H., Nikolakis, Z.L., Westfall, A.K., Orton, R.W., Meik, J.M., Mackessy, S.P., Castoe, T.A., 2020. Snake recombination landscapes are concentrated in functional regions despite PRDM9. Mol. Biol. Evol. 37, 1272–1294. https://doi.org/10.1093/molbev/msaa003.

Schield, D.R., Perry, B.W., Adams, R.H., Holding, M.L., Nikolakis, Z.L., Gopalan, S.S., Smith, C.F., Parker, J.M., Meik, J.M., DeGiorgio, M., Mackessy, S.P., Castoe, T.A., 2022. The roles of balancing selection and recombination in the evolution of rattlesnake venom. Nat. Ecol. Evol. 6, 1367–1380. https://doi.org/10.1038/s41559-022-01829-5.

Schrempf, D., Szöllősi, G., 2020. The Sources of Phylogenetic Conflict, in: Scornavacca, C., Delsuc, F., Galtier, N. (Eds.), Phylogenetics in the Genomic Era. p. 3.1:1–3.1:23.

Singhal, S., Colston, T.J., Grundler, M.R., Smith, S.A., Costa, G.C., Colli, G.R., Moritz, C., Pyron, R.A., Rabosky, D.L., 2021. Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective. Syst. Biol. 70, 542–557. https://doi.org/10.1093/sysbio/syaa054.

Smit, A.F.A., Hubley, R., Green, P., 2015. RepeatMasker Open-4.0. 2013–2015. Institute for Systems Biology. http://repeatmasker.org.

Smith, J., Bruley, C.K., Paton, I.R., Dunn, I., Jones, C.T., Windsor, D., Morrice, D.R., Law, A.S., Masabanda, J., Sazanov, A., Waddington, D., Fries, R., Burt, D.W., 2000. Differences in gene density on chicken macrochromosomes and microchromosomes. Anim. Genet. 31, 96–103. https://doi.org/10.1046/j.1365-2052.2000.00565.x.

Smith, C.F., Mackessy, S.P., 2016. The effects of hybridization on divergent venom phenotypes: characterization of venom from *Crotalus scutulatus scutulatus* × *Crotalus oreganus helleri* hybrids. Toxicon 120, 110–123. https://doi.org/10.1016/j.toxicon.2016.08.001.

Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15, 150. https://doi.org/10.1186/s12862-015-0423-0.

Smith, C.F., Nikolakis, Z.L., Ivey, K., Perry, B.W., Schield, D.R., Balchan, N.R., Parker, J., Hansen, K.C., Saviola, A.J., Castoe, T.A., Mackessy, S.P., 2023. Snakes on a plain: biotic and abiotic factors determine venom compositional variation in a wide-ranging generalist rattlesnake. BMC Biol. 21, 136. https://doi.org/10.1186/s12915-023-01626-x.

Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690. https://doi.org/10.1093/bioinformatics/bts492.

Srikulnath, K., Ahmad, S.F., Singchat, W., Panthum, T., 2021. Why do some vertebrates have microchromosomes? Cells. 10, 2182. https://doi.org/10.3390/cells10092182.

Steenwyk, J.L., Li, Y., Zhou, X., Shen, X.-X., Rokas, A., 2023. Incongruence in the phylogenomics era. Nat. Rev. Genet. 24, 834–850. https://doi.org/10.1038/s41576-023-00620-x.

Strickland, J.L., Smith, C.F., Mason, A.J., Schield, D.R., Borja, M., Castañeda-Gaytán, G., Spencer, C.L., Smith, L.L., Trápaga, A., Bouzid, N.M., Campillo-García, G., Flores-Villela, O.A., Antonio-Rangel, D., Mackessy, S.P., Castoe, T.A., Rokyta, D.R., Parkinson, C.L., 2018. Evidence for divergent patterns of local selection driving venom variation in Mojave Rattlesnakes (*Crotalus scutulatus*). Sci. Rep. 8, 17622. https://doi.org/10.1038/s41598-018-35810-9.

Sunnucks, P., Morales, H.E., Lamb, A.M., Pavlova, A., Greening, C., 2017. Integrative approaches for studying mitochondrial and nuclear genome co-evolution in oxidative phosphorylation. Front. Genet. 8, 25. https://doi.org/10.3389/fgene.2017.00025.

Thomas, G.W.C., Hahn, M.W., 2015. Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. Mol. Biol. Evol. 32, 1232–1236. https://doi.org/10.1093/molbev/msv013.

Uetz, P., Freed, P., Aguilar, R., Reyes, F., Kudera, J., Hošek, J., 2023. The Reptile Database. http://www.reptile-database.org.

Waters, P.D., Patel, H.R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N.C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., Georges, A., Graves, J.A.M., 2021. Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. Proc. Natl. Acad. Sci. u.s.a. 118, e2112494118. https://doi.org/10.1073/pnas.2112494118.

Watson, J.A., Spencer, C.L., Schield, D.R., Butler, B.O., Smith, L.L., Flores-Villela, O., Campbell, J.A., Mackessy, S.P., Castoe, T.A., Meik, J.M., 2019. Geographic variation in morphology in the Mohave Rattlesnake (*Crotalus scutulatus* Kennicott 1861) (Serpentes: Viperidae): implications for species boundaries. Zootaxa 4683, 129–143. https://doi.org/10.11646/zootaxa.4683.1.7.

Wen, D., Yu, Y., Zhu, J., Nakhleh, L., 2018. Inferring phylogenetic networks using PhyloNet. Syst. Biol. 67, 735–740. https://doi.org/10.1093/sysbio/syy015.

Zancolli, G., Baker, T., Barlow, A., Bradley, R., Calvete, J., Carter, K., De Jager, K., Owens, J., Price, J., Sanz, L., Scholes-Higham, A., Shier, L., Wood, L., Wüster, C., Wüster, W., 2016. Is hybridization a source of adaptive venom variation in rattlesnakes? A test, using a *Crotalus scutulatus* × *viridis* hybrid zone in southwestern New Mexico. Toxins 8, 188. https://doi.org/10.3390/toxins8060188.

Zhang, W., Dasmahapatra, K.K., Mallet, J., Moreira, G.R.P., Kronforst, M.R., 2016. Genome-wide introgression among distantly related *Heliconius* butterfly species. Genome Biol. 17, 25. https://doi.org/10.1186/s13059-016-0889-0.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinform. 19 (Suppl. 6), 153. https://doi.org/10.1186/s12859-018-2129-y.

Zhao, M., Kurtis, S.M., White, N.D., Moncrieff, A.E., Leite, R.N., Brumfield, R.T., Braun, E.L., Kimball, R.T., 2023. Exploring conflicts in whole genome phylogenetics: a case study within manakins (Aves: Pipridae). Syst. Biol. 72, 161–178. https://doi.org/10.1093/sysbio/syac062.

Zhou, Y., Huang, D., Xin, Z., Xiao, J., 2020. Evolution of oxidative phosphorylation (OXPHOS) genes reflecting the evolutionary and life histories of fig wasps (Hymenoptera, Chalcidoidea). Genes 11, 1353. https://doi.org/10.3390/genes11111353.

Zhu, S., Degnan, J.H., 2016. Displayed trees do not determine distinguishability under the network multispecies coalescent. Syst. Biol. syw097. https://doi.org/10.1093/sysbio/syw097.