# Species Delimitation of Eastern Pine snakes (Pituophis melanoleucus)

Basanta Khakurel [1*], Zachary L. Nikolakis [1],[2], Brian I. Crother [1], April M. Wright [1]

January 3, 2023

ABSTRACT The eastern Pine snakes *(Pituophis melanoleucus)* is found throughout eastern United States. Taxonomy in this group has been controversial with several conflicting sub-specific designations. Three subspecies of the Eastern Pine snakes have been proposed according to their geographical locations: the northern Pine snake *(P. m. melanoleucus)*, the Florida Pine snake *(P. m. mugitus)*, and the Black Pine snake *(P. m. lodingi)*. There are consistently unresolved relationships among these subspecific taxa in previous studies. The goal of this study was to examine the taxonomic validity of the eastern Pine snake complex using single nucleotide polymorphisms (SNPs) in phylogenetic, population genetic, and species delimitation approaches.

Species delimitation indicated that the plurality of datasets supported an in-group of one species rather than three different subspecies. These results confirm prior findings of little divergence between the three putative subspecies and suggesting one single species.

Keywords: Pine snakes, species delimitation, bppr

**Author Affiliations:**
[1]Department of Biological Sciences, Southeastern Louisiana University, Hammond, LA 70402, USA
[2]Department of Biology, University of Texas at Arlington, Arlington, TX 76019, USA
**Correspondence:***
Basanta Khakurel
basanta.khakurel@selu.edu

# 1  Introduction

The southeastern United States is an area with rich biodiversity consisting of almost half of the country's reptiles and amphibians (Graham et al., 2010). The drainage of major rivers that flow south to the Gulf of Mexico (e.g., Mississippi, Apalachicola, Suwanee) have created differentiating barriers for the biota in the east and the west (Soltis et al., 2006). These complex barriers have provided opportunities for diversification of various flora and fauna in the region. About 20% of the total population of herpetofauna in the region is considered endemic (Graham et al., 2010; Tuberville et al., 2005). Longleaf pine *(Pinus palustris)*, in particular, provides critical habitat for a number of endemic species (Guyer and Bailey, 1993). One of such species is the eastern Pine snake, *Pituophis melanoleucus.* It has been hypothesized to have as many as three subspecies (Stull, 1940; Bonett et al., 2017). Longleaf pine barrens are generally disturbed with fire which makes the habitat suitable for Pine snakes (Burger and Zappalorti, 1988; Zappalorti and Burger, 1985). But with increased human activity, pine barrens is constantly under pressure from development and has become increasingly fragmented over the 20th century. This leads to isolated patches of pine barrens thus creating habitat islands for Pine snakes across its range (Baxley et al., 2011; Baxley and Qualls, 2009).

*Pituophis melanoleucus* occurs across a large range of southeastern United States where many geological barriers that may inhibit gene flow exist (Burbrink et al., 2000; Myers et al., 2020; McKelvy and Burbrink, 2017; Burbrink and Guiher, 2015). Examples of barriers are the Apalachicola and Mississippi river drainage which are believed to have created population differences among many groups of organisms (Pyron and Burbrink, 2009; Soltis et al., 2006). Some prior studies have also supported population structure differences across these barriers in other taxa. For example, some populations of tiger salamanders (Church et al., 2003), rat snakes (Burbrink et al., 2000), musk turtles, and snapping turtles (Thomas et al., 2014) exhibit different population structure in the eastern and western side of the Apalachicola barrier, while some populations of catfish show no genetic differences across the region (Avise et al., 1987).

The eastern Pine snakes have a wide range of habitats across the eastern United States and are thought to contain several distinct populations with high degrees of gene flow (Nikolakis et al., 2021). This species complex is currently classifies with three geographic sub-specific taxa, the northern Pine snake *(P. m. melanoleucus)*, the Florida Pine snake *(P. m.*
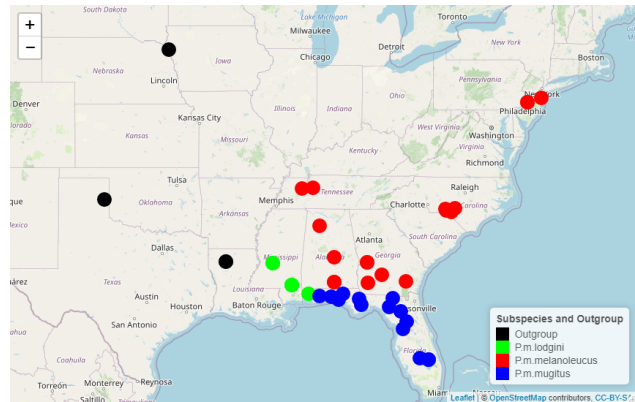


Figure 1: Sampling sites across southeastern United States. The colors indicate the respective subspecies and the outgroup (Figure made using R package *leaflet*)

*mugitus)*, and the Black Pine snake *(P. m. lodingi)* (Bonett et al., 2017). These snakes range widely in color from uniformly black to having red/bronze patches (Guyer et al., 2019). The Black Pine snake has brown or black dorsal coloration, the Florida Pine snake generally has gray anterior color and rusty brown in the posterior with some brown spots in the dorsal side, and the Northern Pine snakes are typically yellow in color with dark blotches in the entire body (Guyer et al., 2019). Besides the difference in coloration, molecular studies have shown that in this complex, there are some populations that are more closely related to the geographically closer populations of different subspecies rather than to more distantly occurring consubspecific population (Rodrıguez-Robles and De Jesús-Escobar, 2000). This species complex is an ideal candidate for testing whether or not the current sub-specific taxonomy represents evolutionary history through the use of species delimitation methods given the uncertainty of relationships from previous studies (Rodrıguez-Robles and De Jesús-Escobar, 2000; Nikolakis et al., 2021). Morphological characters, and particularly synapomorphies, have typically been considered an important component of determining valid species (Mayr, 1981; Mayden, 1999; Assis and Rieppel, 2011). However, Pine snake subspecies are not observed to have synapomorphies and those that are potential synapomorphies are fairly labile, such as scale coloration. Individuals of different subspecies appear to show signs of introgression in the wild (Dye, 2006; Messenger, 2015). In this case, morphology is not conducive to consistent and reproducible taxonomy.

In absence of traditional markers of species distinction according to the morphological or biologi-

cal species concepts, we can make use of molecular data to identify cryptic species. Molecular phylogenetic data have a long history of application to species delimitation problems (Donoghue, 1985). In the earliest forms, this took the form of phylogenetic species concept (De Queiroz, 2007), which posited that species were independent lineages on a phylogenetic tree. More recently, models that provide a more in-depth look at gene flow among taxa have been used for species delimitation from molecular data (Yang and Rannala, 2010). In many ways, this is a return to the biological species concept, placing emphasis on genetic introgression and lineage separation. The multi-species coalescent (MSC) method uses both molecular phylogenetics and population genetics to counter problems, such as long-branch attraction and the inherent subjectivity of interpreting the phylogeny, posed by delimiting species on the basis of phylogenetic tree (Yang and Rannala, 2010; Yang, 2015). Unlike the traditional phylogenetic methods which assume that the same tree underlies all gene loci, MSC accounts for coalescent processes in ancient and modern species and the resultant species-gene tree conflicts by allowing for multiple gene trees to underlie the data (Yang, 2015). There are various tree-based and non-tree-based species delimitation methods that can be used to determine the species' boundaries (Camargo et al., 2013). *bpp* is a software that generates a Bayesian posterior distribution of species delimitation models using multi-species coalescent framework (Yang and Rannala, 2014). It uses reversible-jump MCMC to move between models of the number of populations present in the sample, while calculating the posterior probabilities associated with the model of population differentiation. *bpp* allows us to calculate the model likelihoods for different number of species using the empirical data.

The populations of Pine snakes are geographically separated due to various causes. Some of them include natural geographic barriers along its range (rivers and drainages) and some are the results of human activities (development and logging). The geographic barrier systems along the range of pine snakes includes the Alabama river system, the Apalachicola river and the Suwanee river. In a study from Reichling Reichling (1995), the Tennessee river valley populations of *P. m. melanoleucus*, to the North of most other pine snake populations, could be seen as a geographically isolated OTU. The geographic barriers and previous works have allowed us to test several population grouping hypotheses to study the different populations and subspecies of pine snakes. We tested different population groups described in the sub-section **Species delimitation analyses** under

the section **Methods**.

Only one study related to population genomic study has been done on the *Pituophis melanoleucus* complex to this date (Nikolakis et al., 2021) and we hope this study will provide further crucial information regarding its taxonomy. The goal of this study was to examine the phylogenetic status and taxonomic validity of the three subspecies of Pine snakes. In this study, we make use of a dataset of SNPs obtained from UCE data, which provided us genome wide structural variation. Ultra-conserved elements (UCE) are the highly conserved regions within the genome that are shared among evolutionary distinct taxa (Bejerano et al., 2004) and are flanked by more variable regions. Single Nucleotide Polymorphisms (SNPs) are changes in one base pair along the genome, which allows us to capture more phylogenetically informative variation in the sequence (Morin et al., 2004).

Using the SNPs, we were able to perform Bayesian phylogenetic inference, which allowed us to infer the phylogenetic relationships among the population of each putative subspecies sampled across the eastern United States. We also performed several population genetic analyses to examine the genetic structure and the amount of diversity among the recovered lineages. In addition, we conducted species delimitation analysis using the full UCE dataset to confirm our conclusions on the diversity of various populations.

# 2   Methods

## 2.1   Sample collection, DNA extraction, and Bioinformatics

Tissue samples were collected from forty-two specimens of *Pituophis melanoleucus* from their geographical distribution (Fig. 1) by Nikolakis et al. for their 2021 publication. The samples provide coverage above most of the Pine snakes' ranges. The samples were de-multiplexed, filtered, and processed by removing adaptor sequences and ambiguous bases using the program *Illumiprocessor* which is incorporated in the software *Phyluce* v.1.5 (Faircloth, 2015) according to the protocol outlined in Nikolakis, Orton, and Crother (2021). The UCE loci were used to obtain SNPs and was used in further analyses.

## 2.2   Phylogenetic Analyses

We conducted phylogenetic analyses using a nucleotide substitution model from *RevBayes* software v.1.1.1 (Höhna et al., 2016). We used the general time reversible (GTR) model (Tavaré et al., 1986) of se-
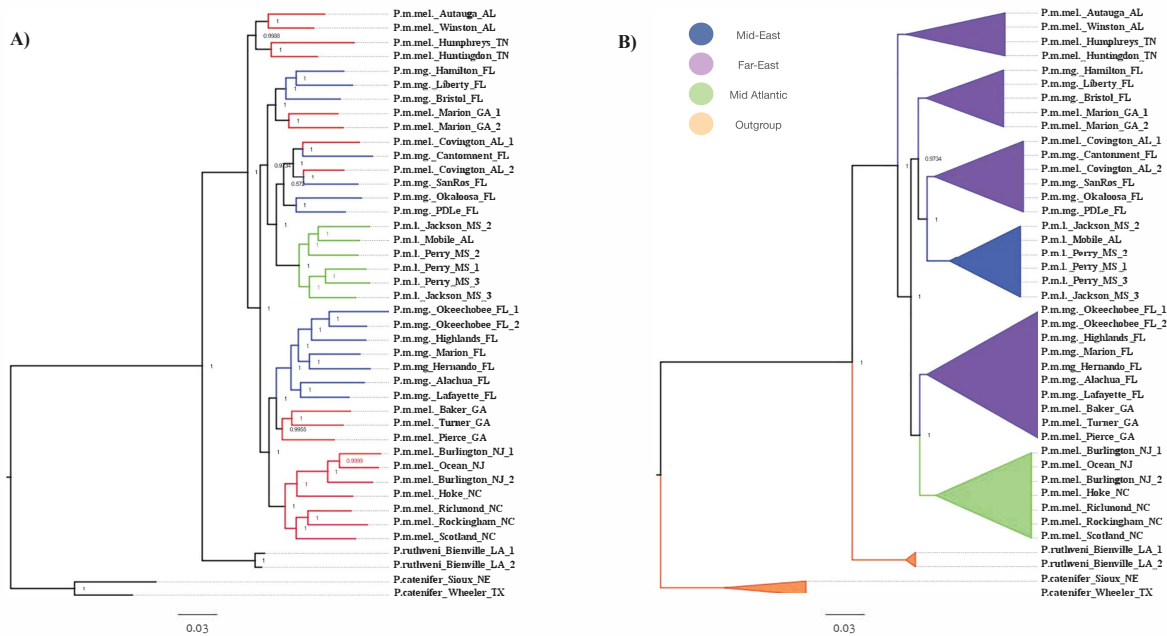
Figure 2: Phylogeny generated using GTR+I+Γ model in *RevBayes*. a) The colors of the tips represent current subspecific designations: Black represents the outgroup, red corresponds to *Pituophis melanoleucus melanoleucus*(Northern Pine snakes), blue corresponds to *Pituophis melanoleucus mugitus* (Florida Pine snakes), and green corresponds to *Pituophis melanoleucus lodingi* (Black Pine snakes). b) The colors in the clades represents the clusters from DAPC.

quence evolution, which allows six exchangeability rates between nucleotide states. The exchangeabilities are drawn from a Dirichlet distribution with an uninformative prior and allows the data to determine the value of the exchange abilities. We also used Gamma-distributed among site rate variation to allow different sites to evolve at different rates (Yang, 1994). The MCMC was run to replicate 150,000 generations and the resulting log files were viewed in *Tracer* v.1.7.1 (Rambaut et al., 2018) to check for convergence. The output files were summarized into maximum clade credibility trees (Helfrich et al., 2018) using *RevBayes*.

## 2.3   Population structure

In addition to phylogenetic work, we conducted Discriminant Analysis of Principal Components (DAPC) analysis using the R package adegenet (Jombart, 2008). Under this methodology, the researcher must predefine the study populations. Therefore, we tested the two, three, and four population hypotheses outlined in section **Introduction**. The hypotheses included using the traditional taxonomic grouping with

three subspecies, Tennessee northern Pine snakes as a separate group as suggested by Reichling (1995), and the population groupings as found by Nikolakis et al. (2021). Adegenet uses BIC to assess the optimal number of populations for the DAPC. In our case, the BIC was minimized by a four-population model. Additionally, we used the compoplot module in adegenet to determine the probability of assigning each individual to their population as assigned by DAPC.

## 2.4   Species delimitation analyses

In order to perform species delimitation analyses, we used the software *bpp*. All the *bpp* analyses did not use a guide tree for the clusters, and we used an inverse gamma prior on tau (3, 0.002) and theta (3, 0.004) as those corresponds to broad and uninformative priors (Flouri et al., 2020). The parameter theta is a measure of heterozygosity in the species and the parameter tau is the age of the root in the species tree, and both the priors depends on the species in the dataset used for the analyses (Flouri et al., 2020). For our analyses, we discarded the first 2000 samples
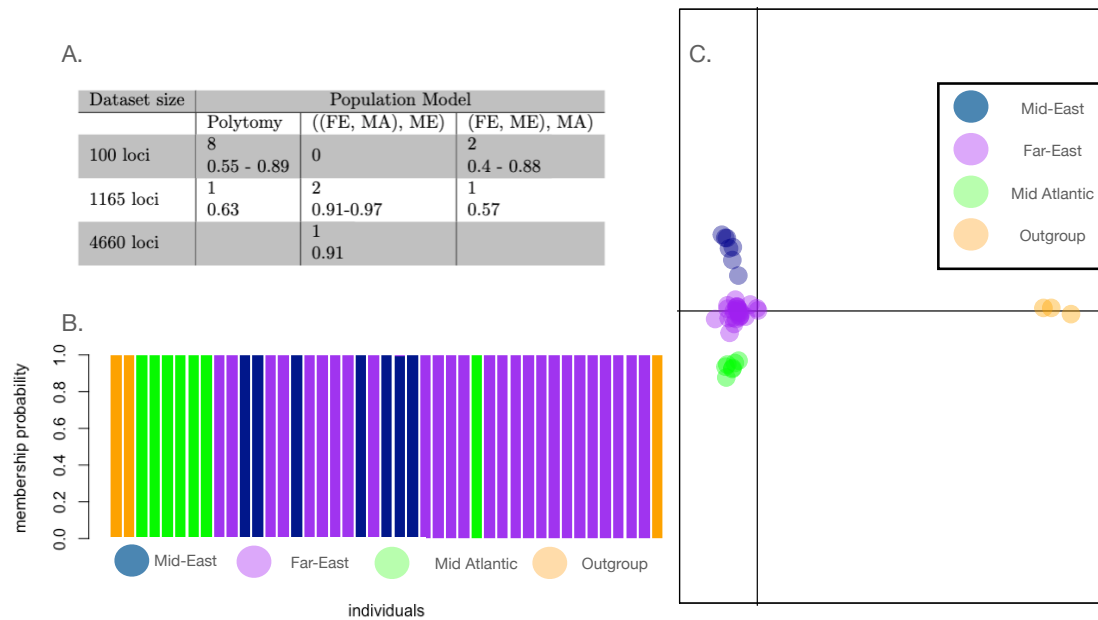
A.

| Dataset size | Population Model | | |
|---|---|---|---|
| | Polytomy | ((FE, MA), ME) | (FE, ME), MA) |
| 100 loci | 8<br>0.55 - 0.89 | 0 | 2<br>0.4 - 0.88 |
| 1165 loci | 1<br>0.63 | 2<br>0.91-0.97 | 1<br>0.57 |
| 4660 loci | | 1<br>0.91 | |

Figure 3: **Species delimitation results for *Pituophis melanoleucus.*** Panel A shows the support for different population models with their posterior probabilities among different dataset sizes. Panel B. shows the population assignments for each sample in the dataset, and Panel C shows a population PCA of the samples.

as burn-in and ran the analyses for 20,000 samples.

Due to the large size of the data, we tried several different schemas to process the data. First, we performed it using the concatenated SNP dataset. Because *bpp* does not include a variable sites correction (Lewis, 2001), this will tend to overestimate the branch lengths. Therefore, we also did several analyses involving whole UCE loci. We performed the analysis on several subsets of the complete 4600 UCE dataset with outgroup sequences for *Pituophis catenifer* and *Pituophis ruthveni*. *P. catenifer* and *P. ruthveni* are sister taxa to each other and as a whole, they constitute the sister clade to the *Pituophis* complex. We also performed two downsampling experiments (Smith and Carstens, 2020) for our species delimitation analyses. First, we did a random sampling of 100 loci into ten different replicates. Second, we randomly separated the 4600 loci to four replicates containing 1165 loci each. *bpp* uses multi-species coalescent method to estimate the divergence times and population sizes for both extant and ancestral sequences. This information is used to derive the probability distributions indicating whether the lineages can be differentiated from each other. We used the defined algorithm 'A10', which uses reversible jump

MCMC to test various species delimitation models. We used the clusters obtained from DAPC analyses to group the populations of the subspecies.

The groups in the *bpp* analyses consisted of Far-East (FE), Mid-East (ME), and Mid-Atlantic (MA) populations. The group FE consisted of samples from northern Alabama, Tennessee, and Florida, ME consisted of Alabama and Mississippi samples and MA consisted of North Carolina and New Jersey samples.

## 3 Results

### 3.1 Phylogenetic inference

The full dataset including the outgroups had 39077 SNPs. The dataset was used to infer a Bayesian phylogeny under the GTR+I+Γ model. A phylogenetic tree rooted with the outgroup is depicted in Figure 2. While what has been referred to as *Pituophis melanoleucus lodingi* is monophyletic, the other two proposed subspecies are not.

## 3.2   Population structure

Population genetic metrics were used to assess genetic variability among the individuals and clusters of the *P. melanoleucus* complex. The first method employed was a DAPC, which recovered three in-group clusters. These clusters are distinct and non-overlapping (Figure 3). However, the clusters do not align to the subspecies proposed in the literature. These results, instead, suggest that spatial structure is a better descriptor of the species complex. The compoplot indicates that individuals in these clusters typically have a 100% posterior probability of being assigned to their DAPC cluster. Using the results from the DAPC clusters, we separated the populations accordingly to perform species delimitation under the three-species DAPC model.

## 3.3   Species Delimitation

*bpp* was able to detect some structure in the clusters that we provided from our DAPC analyses. For the 100 UCE runs, there were mixed results among the 10 replicates. Amongst the 10 randomly sampled dataset, only 2 of them indicated that the three clusters can be considered separate species. Within the 8 remaining replicates that did not differentiate all three clusters as different species, 2 of them grouped FE and ME clusters together and the remaining 6 of the replicates grouped FE with MA indicating that ME has already diverged from the complex. All of the replicates have posterior probability ranging from 0.5 to 0.8.

In the 4 replicates with 1165 UCEs, two replicates show that FE and MA are sister to each other and ME is sister to (FE, MA) group with high posterior probability ( $< 0.92$ ). One replicate showed that FE is sister to ME and MA is sister to (FE, ME) with a posterior probability of 0.566. The remaining folder grouped FE and MA together as sister to ME cluster with posterior probability 0.626.

In the analysis with all 4660 UCEs, the results indicate presence of 2 different species of Pine snakes, it groups the FE and MA group together being sister to ME group. The posterior probability of presence of 3 different species (Far-East grouped with Mid-Atlantic, Mid-East and Outgroup) is 0.913.

## 4   Discussion

The Family Colubridae is the largest family of snakes found in every continent except Antarctica (Cogger et al., 1998). Among the many unresolved phylogenetic relationships in various genus, Pituophis is
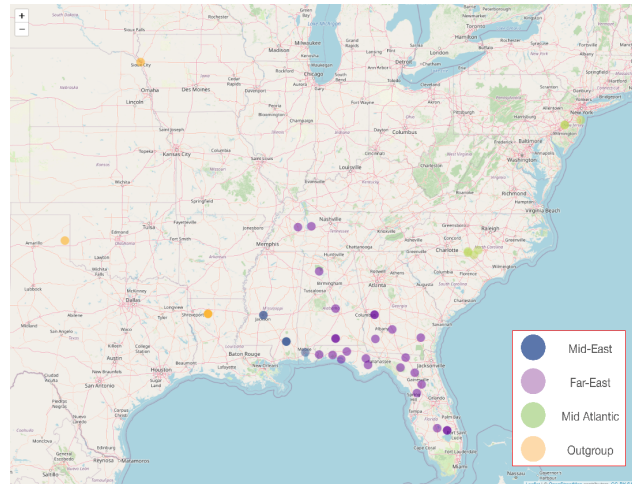


Figure 4: The samples map colored according to the clusters obtained from DAPC (Figure made using R package *leaflet*). Opacity of points corresponds to number of samples in that locality.

one of them having many subspecific designations. *Pituophis melanoleucus* is one of the species with three different subspecies. Morphologically the three subspecies appear different in superficial characters (scale color and pattern), but there has been only one other study (Nikolakis et al., 2021) regarding the taxonomic status of the three subspecies. The difference in coloration is not a reliable synapomorphy to diagnose the three subspecies as different species, so molecular methods have to be taken in account.

### 4.1   Phylogeny, DAPC and bpp

The phylogenetic analysis using *RevBayes* agrees with the population structure results obtained using DAPC with clear differentiation of clades in the tree (Figure 2 and 4). This partially agrees with the hypotheses that mentions the major barrier in the range of the Pine snakes could be the Apalachicola-Chattahoochee drainage system in the southern part and anthropogenic activities in the northern (Pyron and Burbrink, 2009; Burbrink et al., 2000; Weinell and Austin, 2017). The traditional taxonomic hypothesis groups three subspecies as an evolutionary entity which was contradicted by the phylogeny we obtained (Bonett et al., 2017). The SNP dataset that produced the phylogenetic tree pointed out that there were individuals of different subspecies that were geographically closer to each other formed a monophyletic group rather than individuals of same subspecies being grouped together. On our phylogeny, the current taxonomy of the Pine snakes is muddled in the

tree with different subspecies being sister to other subspecies rather than the ones of the same subspecies. Only *Pituophis melanoleucus lodingi* (Black Pine snake) is monophyletic on the tree. In addition to that there is a group of *P. m. melanoleucus* (Northern Pine snake) in the North Carolina and New Jersey region that appear monophyletic in the tree. All the other *P. m. melanoleucus* and *P. m. mugitus* (Florida Pine snake) have been placed in various places all across the phylogeny.

The DAPC analysis resulted in three clusters with a group in the Atlantic side comprising New Jersey and North Carolina populations, another group of populations from Alabama and Mississippi and the third group containing populations from Alabama, Tennessee, Florida, and Georgia. The phylogenetic tree also agrees with the clusters from DAPC, rendering population groups of Mid-Atlantic (NJ and NC), and Mid-East (MS and some AL) monophyletic. In the other hand, the Far-East group is rendered non-monophyletic in the tree. This suggests that the smaller groups (Mid-East and Mid-Atlantic) are diverging from a larger (Far-East) group. According to the results from DAPC and the phylogeny, the population structure maps with the Apalachicola-Chattahoochee drainage systems which splits the monophyletic group of *P. m. lodingi* (Mid-East) from the Far-East group and the small population of *P. m. melanoleucus* (Mid-Atlantic) being isolated due to anthropogenic activities such as logging and urbanization.

As more data are added, more population structure is uncovered using *bpp*. *bpp* also shows support for some structuring in the populations of Pine snakes. Most of the results in the species delimitation analyses recognize each cluster (Far-East, Mid-East, and Mid-Atlantic) as separate evolutionary entities. In the analyses with ten replicates of 100 loci each, there was indication of Far-East and Mid-Atlantic as a single cluster and Mid-East being a separate population group. This separation reflects the barrier caused by Apalachicola-Chattahoochee drainage system as a major separation in the range of *Pituophis*. In the analysis with four replicates of 1165 loci each, there was an indication of both Mid-East and Mid-Atlantic being a separate evolution entity in different runs. In one of the runs, Mid-Atlantic is grouped together with Far-East with Mid-East as the sister to the group of MA and FE. All the results from this analyses indicate that both Apalachicola-Chattahoochee drainage system and urbanization acts as a major barrier in the range of Pine snakes. One thing to keep in mind is the type of data we used for these studies. The UCE data are fairly conserved across a

long range of taxa which could lead to different results. Further comments in this is highlighted below in the sub-section **UCE in Species Delimitation**.

In the species delimitation analyses with all the 4660 loci, Far-East and Mid-Atlantic groups are indicated to be a same evolutionary entity and Mid-East as the sister to combined population group of FE and MA. This would suggest that Alabama river system has been acting as a major barrier in the range. The Mid-Atlantic group is still showing some introgression with the Far-East group. Although there exists a large area between the two groups, the high mobility of Pine snakes could possibly have contributed to the gene flow between the individuals of these populations. Although not fully resolved as separate species, the clusters obtained from DAPC analyses show some genetic differentiation according to the species delimitation studies.

Previous studies have also found that some populations of *P. melanoleucus* are more close to other geographically closer populations of other subspecies than distantly situated populations of same subspecies (Rodrıguez-Robles and De Jesús-Escobar, 2000). All of our studies show little phylogenetic support for the traditional taxonomic hypotheses suggesting more gene flow among the members of different subspecies that are geographically closer. Habitat fragmentation and loss of pine forests has become a major driving force in population structuring of the *P. melanoleucus* complex. Lack of prescribed burns have been impacting the health of remaining pine forests which has in turn been affecting the populations of *P. melanoleucus*. (Noss, 1988, 1989) Previous squamate phylogeographic studies across these regions (Burbrink et al., 2000; Soltis et al., 2006; Myers et al., 2020) indicated that there is significant genetic variation among clades that existed across the barrier. Despite the geographic barrier, the population structuring analyses and molecular species delimitation indicate that there is not enough variation for the groups to be distinct from each other. Regardless of that, our studies indicate that there are some populations of Pine snakes that are diverging. For example, the groups Mid-East and Mid-Atlantic containing populations of Black and Northern Pine snakes seem to be diverging from a larger group indicating there could be some geographic isolation within the groups. Due to the indication of little genetic differentiation, it can be concluded that gene flow has been maintained in the populations of eastern pine snakes across the region. Some movement studies have indicated that *P. melanoleucus* is a very mobile species with their home ranges spanning from approximately 35 hectares to over 105 hectares (Kapfer et al., 2010) which could

contribute to the gene flow among different populations.

## 4.2 UCE in Species Delimitation

The utility of UCEs has been well-demonstrated in deeper timescales (Faircloth, 2015; McCormack et al., 2013; Gilbert et al., 2015) but has not been less explored much in recent evolutionary events (Winker et al., 2018; Mason et al., 2018; Harvey et al., 2016). As UCEs are markers that are conserved across a large range of taxa, they are highly conserved regions within the genome that are shared among evolutionarily distant taxa (Bejerano et al., 2004). Having a large set of UCE loci led us to first attempt using SNPs for some analyses which proved valuable but for species delimitation using *bpp*, but due to the lack of invariant sites, this produced inflated branch lengths in our trees. We decided to try downsamples of different number of loci to test how it would influence the results we had. In addition to concatenating the dataset, we treated each UCE as a separate loci. As suggested by the scientific literature (Bryant and Hahn, 2020; Roch and Steel, 2015), concatenation approaches do not produce a reliable result. This is something we experienced with our study, concatenating all UCEs and even using concatenated variable sites only data produced results that were inconclusive whereas treating the UCE loci as separate showed some structure in the populations of Pine snakes across their range. *bpp* can accommodate many loci without concatenation, making it an ideal tool for exploring this problem.

Upon addition of more data in our species delimitation studies, the results somewhat varied. As depicted in Figure 3A, using fewer loci caused more uncertainty in our results. In a recently diverged group of Pine snakes, addition of more data is crucial to have enough variability among sites to support divergences on the tree. The posterior probabilities of proposed clusters were also higher when we added more loci to our analyses. One of the limitations for this study could be considered the use of UCE for our analyses. As mentioned before, UCEs are shared across a large range of evolutionarily distant taxa and this sharing of markers in evolutionarily distant taxa might have influenced our study. The lack of much distinction among the population groups might have been due to the use of UCEs and a more hypervariable marker would be a better option for study like this. This is because of hypothesized recent evolution of eastern Pine snakes from other *Pituophis* (Pyron and Burbrink, 2009). The oldest fossils of *P. melanoleucus* have been found in Florida dating 0.8 to 2.5 million years (myr) and more northern fossils from Pennsylvania have been dated from 0.1 myr (Holman, 2000). These dates indicate that the lineage diverged during the late Pleistocene when there were environmental fluctuations leading to periods of isolation and connection. This could also be an important factor that maintained gene flow through the contacts of different populations. Another possible factor that contributed in the gene flow of the Pine snakes are the home ranges as mentioned in subsection **Phylogeny, DAPC and bpp**. Our study indicates that the eastern Pine snakes have shown some population structuring due to the geographic barriers existing across its range. But the different groups of Pine snakes do not show enough support to establish them as their own species. As some populations of Pine snakes are more closely related to geographically closer populations of different subspecies than to more distant populations of same subspecies, we can conclude that the different Pine snakes subspecies prevent themselves from being elevated into a species level.

Combining multi-locus data with movement studies of the *P. melanoleucus* complex would possibly be helpful in understanding further about the phylogeography of these species. We hope that this study will be helpful for future research in the population dynamics and taxonomy of the Pine snakes complex.

## 5 Conclusions

Due to the recent period of isolation and connection during late Pleistocene and large home ranges of Pine snakes, they seem to have maintained gene flow which is slowly diminishing due to barriers and anthropogenic activities. Using a dataset of UCEs, we did not find enough evidence for multiple species in *(Pituophis melanoleucus)*. Using the multi-species coalescent and the phylogenetic species concept, we substantiate *Pituophis melanoleucus* as a single species across its whole range with three subspecies, a hypothesis suggested by previous authors (Nikolakis et al., 2021; Bonett et al., 2017).

## References

Assis, L. C. and O. Rieppel. 2011. Are monophyly and synapomorphy the same or different? revisiting the role of morphology in phylogenetics. Cladistics 27:94–102.

Avise, J. C., C. A. Reeb, and N. C. Saunders. 1987. Geographic population structure and species dif-

ferences in mitochondrial dna of mouthbrooding marine catfishes (ariidae) and demersal spawning toadfishes (batrachoididae). Evolution 41:991–1002.

Baxley, D., G. J. Lipps Jr, and C. P. Qualls. 2011. Multiscale habitat selection by black pine snakes (pituophis melanoleucus lodingi) in southern mississippi. Herpetologica 67:154–166.

Baxley, D. L. and C. P. Qualls. 2009. Black pine snake (pituophis melanoleucus lodingi): Spatial ecology and associations between habitat use and prey dynamics. Journal of Herpetology 43:284–293.

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

Bonett, R. M., J. Boundy, F. T. Burbrink, B. I. Crother, K. de Queiroz, D. R. Frost, R. Highton, J. B. Iverson, E. L. Jockusch, F. Kraus, et al. 2017. Scientific and standard English names of amphibians and reptiles of North America north of Mexico, with comments regarding confidence in our understanding. Society for the Study of Amphibians and Reptiles Lawrence, Kansas.

Bryant, D. and M. W. Hahn. 2020. The concatenation question.

Burbrink, F. T. and T. J. Guiher. 2015. Considering gene flow when using coalescent methods to delimit lineages of north american pitvipers of the genus agkistrodon. Zoological Journal of the Linnean Society 173:505–526.

Burbrink, F. T., R. Lawson, and J. B. Slowinski. 2000. Mitochondrial dna phylogeography of the polytypic north american rat snake (elaphe obsoleta): a critique of the subspecies concept. Evolution 54:2107–2118.

Burger, J. and R. T. Zappalorti. 1988. Habitat use in free-ranging pine snakes, pituophis melanoleucus, in new jersey pine barrens. Herpetologica Pages 48–55.

Camargo, A. et al. 2013. Species delimitation: a decade after the renaissance. *in* The species problem-ongoing issues. IntechOpen.

Church, S. A., J. M. Kraus, J. C. Mitchell, D. R. Church, and D. R. Taylor. 2003. Evidence for multiple pleistocene refugia in the postglacial expansion of the eastern tiger salamander, ambystoma tigrinum tigrinum. Evolution 57:372–383.

Cogger, H. G., R. G. Zweifel, D. Kirshner, et al. 1998. Encyclopedia of reptiles & amphibians. Weldon Owen.

De Queiroz, K. 2007. Species concepts and species delimitation. Systematic biology 56:879–886.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. Bryologist Pages 172–181.

Dye, D. D. 2006. Black Pine Snake.

Faircloth, B. C. 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

Flouri, T., B. Rannala, and Z. Yang. 2020. A Tutorial on the Use of BPP for Species Tree Estimation and Species Delimitation. Pages 5.6:1–5.6:16 *in* Phylogenetics in the Genomic Era (C. Scornavacca, F. Delsuc, and N. Galtier, eds.). No commercial publisher — Authors open access book.

Gilbert, P. S., J. Chang, C. Pan, E. M. Sobel, J. S. Sinsheimer, B. C. Faircloth, and M. E. Alfaro. 2015. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Molecular phylogenetics and evolution 92:140–146.

Graham, S. P., D. A. Steen, K. T. Nelson, A. M. Durso, and J. C. Maerz. 2010. An overlooked hotspot? rapid biodiversity assessment reveals a region of exceptional herpetofaunal richness in the southeastern united states. Southeastern Naturalist 9:19–34.

Guyer, C. and M. A. Bailey. 1993. Amphibians and reptiles of longleaf pine communities. Pages 139–158 *in* Proceedings of the tall timbers fire ecology conference vol. 18.

Guyer, C., M. A. Bailey, and R. H. Mount. 2019. Lizards and snakes of Alabama. University Alabama Press.

Harvey, M. G., B. T. Smith, T. C. Glenn, B. C. Faircloth, and R. T. Brumfield. 2016. Sequence capture versus restriction site associated dna sequencing for shallow systematics. Systematic biology 65:910–924.

Helfrich, P., E. Rieb, G. Abrami, A. Lücking, and A. Mehler. 2018. TreeAnnotator: Versatile visual annotation of hierarchical text relations. *in* Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC

2018) European Language Resources Association (ELRA), Miyazaki, Japan.

Höhna, S., M. J. Landis, T. A. Heath, B. Boussau, N. Lartillot, B. R. Moore, J. P. Huelsenbeck, and F. Ronquist. 2016. Revbayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. Systematic biology 65:726–736.

Holman, J. A. 2000. Fossil Snakes of North America: Origin, evolution, distribution, Paleoecology. Indiana University Press.

Jombart, T. 2008. adegenet: a r package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Kapfer, J., C. Pekar, D. Reineke, J. Coggins, and R. Hay. 2010. Modeling the relationship between habitat preferences and home-range size: a case study on a large mobile colubrid snake from north america. Journal of Zoology 282:13–20.

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

Mason, N. A., A. Olvera-Vital, I. J. Lovette, and A. G. Navarro-Sigüenza. 2018. Hidden endemism, deep polyphyly, and repeated dispersal across the isthmus of tehuantepec: Diversification of the white-collared seedeater complex (thraupidae: Sporophila torqueola). Ecology and Evolution 8:1867–1881.

Mayden, R. L. 1999. Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. Journal of Nematology 31:95.

Mayr, E. 1981. Biological classification: toward a synthesis of opposing methodologies. Science 214:510–516.

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

McKelvy, A. and F. Burbrink. 2017. Ecological divergence in the yellow-bellied kingsnake (lampropeltis calligaster) at two north american biodiversity hotspots. Molecular Phylogenetics and Evolution 106:61–72.

Messenger, K. 2015. Northern x Florida pine intergrade.

Morin, P. A., G. Luikart, R. K. Wayne, et al. 2004. Snps in ecology, evolution and conservation. Trends in ecology & evolution 19:208–216.

Myers, E. A., A. D. McKelvy, and F. T. Burbrink. 2020. Biogeographic barriers, pleistocene refugia, and climatic gradients in the southeastern nearctic drive diversification in cornsnakes (pantherophis guttatus complex). Molecular ecology 29:797–811.

Nikolakis, Z. L., R. Orton, and B. I. Crother. 2021. Fine scale population structure and extensive gene flow within an eastern nearctic snake complex (pituophis melanoleucus). bioRxiv .

Noss, R. F. 1988. The longleaf pine landscape of the southeast: almost gone and almost forgotten. Endangered Species Update 5:1–8.

Noss, R. F. 1989. Longleaf pine and wiregrass: keystone components of an endangered ecosystem. Natural Areas Journal 9:211–213.

Pyron, R. A. and F. T. Burbrink. 2009. Neogene diversification and taxonomic stability in the snake tribe lampropeltini (serpentes: Colubridae). Molecular Phylogenetics and Evolution 52:524–529.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology 67:901–904.

Reichling, S. B. 1995. The taxonomic status of the louisiana pine snake (pituophis melanoleucus ruthveni) and its relevance to the evolutionary species concept. Journal of Herpetology Pages 186–198.

Roch, S. and M. Steel. 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theoretical population biology 100:56–62.

Rodrıguez-Robles, J. A. and J. M. De Jesús-Escobar. 2000. Molecular systematics of new world gopher, bull, and pinesnakes (pituophis: Colubridae), a transcontinental species complex. Molecular Phylogenetics and Evolution 14:35–50.

Smith, M. L. and B. C. Carstens. 2020. Process-based species delimitation leads to identification of more biologically relevant species. Evolution 74:216–229.

Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern north america. Molecular ecology 15:4261–4293.

Stull, O. G. 1940. Variations and relationships in the snakes of the genus Pituophis. 175 US Government Printing Office.

Tavaré, S. et al. 1986. Some probabilistic and statistical problems in the analysis of dna sequences. Lectures on mathematics in the life sciences 17:57–86.

Thomas, T. M., M. C. Granatosky, J. R. Bourque, K. L. Krysko, P. E. Moler, T. Gamble, E. Suarez, E. Leone, K. M. Enge, and J. Roman. 2014. Taxonomic assessment of alligator snapping turtles (chelydridae: Macrochelys), with the description of two new species from the southeastern united states. Zootaxa 3786:141–165.

Tuberville, T. D., J. D. Willson, M. E. Dorcas, and J. W. Gibbons. 2005. Herpetofaunal species richness of southeastern national parks. Southeastern naturalist 4:537–569.

Weinell, J. L. and C. C. Austin. 2017. Refugia and speciation in north american scarlet snakes (cemophora). Journal of Herpetology 51:161–171.

Winker, K., T. C. Glenn, and B. C. Faircloth. 2018. Ultraconserved elements (uces) illuminate the population genomics of a recent, high-latitude avian speciation event. PeerJ 6:e5735.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from dna sequences with variable rates over sites: approximate methods. Journal of Molecular evolution 39:306–314.

Yang, Z. 2015. The bpp program for species tree estimation and species delimitation. Current Zoology 61:854–865.

Yang, Z. and B. Rannala. 2010. Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences 107:9264–9269.

Yang, Z. and B. Rannala. 2014. Unguided species delimitation using dna sequence data from multiple loci. Molecular Biology and Evolution 31:3125–3135.

Zappalorti, R. T. and J. Burger. 1985. On the importance of disturbed sites to habitat selection by pine snakes in the pine barrens of new jersey. Environmental Conservation 12:358–361.