**ABSTRACT**

The Eastern Pine Snake (*Pituophis melanoleucus*) is found throughout eastern North America. Taxonomy in this group has been controversial with several conflicting subspecific 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 resolved relationships among these subspecific taxa in previous studies. We analyzed ultra-conserved elements (UCEs) to perform species tree estimation and species delimitation approaches implementing Bayesian inference methods. Species delimitation indicated that the plurality of datasets supported an ingroup 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.

**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). About 20% of the total population of herpetofauna in the region is considered endemic (Graham et. al. 2010; Tuberville et. al. 2005). Longleaf pine, in particular, provides critical habitat for a number of endemic species (Guyer and Bailey 1993). One of such species is the eastern pine snakes, *Pituophis melanoleucus.* It has been hypothesized to have as many as three subspecies (Crother 2012). In this manuscript, we will use molecular species delimitation methods to examine the taxonomy in this group.

Molecular phylogenetic data have a long history of application to species delimitation problems (Donoghue 1985). In the earliest forms, this took the form of the phylogenetic species concept (de Queiroz 2007), which posited that species were independent lineages on a phylogenetic tree. 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. The multispecies coalescent (MSC) method utilizes both molecular phylogenetics and population genetics to counter problems, such long-branch attraction and the inherent subjectivity of interpreting the phylogeny, posed by traditional phylogenetic methods (Yang and Rannala 2010, Yang 2015). Unlike the traditional phylogenetic methods which assumes 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). Different evolutionary processes operate at different geographical locations which would lead to population genetic reconstructing over time (Soltis et al. 2006). In this manuscript, we use both the phylogenetic tree and the MSC to delimit species in the genus *Pituophis*.

Morphological characters, and particularly synapomorphies, have typically been considered an important component of determining valid species (Assis and Rieppel 2011). However, pine snakes are not observed to have synapomorphies and those that are potential synapomorphies are fairly labile, such as scale coloration. Individuals who appear to show signs of introgression have been observed (Scott 2008). In the absence of traditional markers of species distinction according to the morphological or biological species concepts, we can make use of molecular data to search for cryptic species variation. There are multiple types of molecular data that can be brought to bear this question. Each type has different evolutionary properties that lead to the marker capturing different types of variation. Some of such molecular data are DNA barcode (Herbert et. a. 2003), UCEs (Winker 2018), ddRADseq (Peterson et al. 2012, Reitzel et. al. 2013), Sequence capture (Anderman et. al. 2020). In this study, we make use of a UCE dataset collected for phylogenetics in the pine snakes group (Nikolakis 2018).

*Pituophis melanoleucus* occurs across a large range of southern and eastern United States. In this region, there are many geological barriers that may inhibit gene flow. (Burbrink et al. 2000; Burbrink and Guiher 2015; McKelvy and Burbrink 2017; Myers et al. 2020; Soltis et al. 2006). Examples of barriers are the Apalachicola and Mississippi river drainages which are believed to have created population differentiation among a lot of organisms (Pyron and Burbrink 2009; Soltis et. al. 2006). Prior studies have also supported population structure differences on different sides of these barriers. For example, the populations of tiger salamander, rat and corn snakes, musk turtle exhibit different population structure in the eastern and western side of the barrier (Church et. al. 2003; Burbrink et. al. 2000; Burbrink 2002), but the populations of snapping turtles and some 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 (Nikolakis 2018). The snakes are currently classified with three geographic sub-specific taxa the Northern Pine snakes (*P. m. melanoleucus*), (2) the Florida Pine snakes (*P. m. mugitus*), and (3) the Black Pine snakes (*P. m. lodingi*) (Crother 2012). These snakes range widely in color from uniformly black to having red/bronze patches (Guyer et al. 2019). This is an interesting group of organisms in which to perform species delimitation as the relationships between lineages, and indeed if there are multiple lineages, has been in doubt (Scott 2008, Rodriguez-Robles and De Jesus-Escobar 2000).

There are various tree-based and non-tree-based species delimitation methods that can be used to determine the species boundaries (Camargo and Sites 2013). bpp is a software that generates the Bayesian posterior distribution of species delimitation models using the multispecies coalescent framework (Yang and Rannala 2010). 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 numbers of species. We used bpp to test how many species are present in our ingroup.

We used UCEs to obtain the evolutionary relationships within *Pituophis melanoleucus*, to determine if UCEs are best suited for this type of analyses, and to determine if there is support for multiple subspecies in *Pituophis melanoleucus*. We used bpp(Flouri et. al. 2018) for delimiting species under the multispecies coalescent. We also used RevBayes (Hӧhna et. al. 2016) to determine if there is any existing phylogenetic structure within the group, as the phylogenetic species concept defines species as a distinct group of conspecific individuals (Woodruff 2001). This study builds on the recent work of Nikolakis (2018) and will help to provide a better understanding of diversity patterns of pine snakes in the eastern United States and may provide additional insight on the utility of UCEs to study phylogenetic relationships in recently diverged clades.

**MATERIALS AND METHODS**

**Sample collection and DNA extraction**

Tissue Samples were collected from forty-three specimens of Pituophis melanoleucus from their geographical distribution (Figure 1, Table 1). For the outgroup taxa, five samples each were collected from specimens of *Pituophis ruthveni*, *Pituophis catenifer*, and *Pantherophis obsoletus*. Samples were taken from the liver or muscle tissues and from the shed skins. Some samples were also obtained from ventral scale clips which are used for marking snakes. The collected samples were preserved in 95-100% ethanol until DNA extraction. Genomic DNA was extracted and isolated using Qiagen’s DNeasy kit using the manufacturers protocols and quantified using Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The DNA samples were diluted to concentrations ranging from 2 to 100 ng/µL and the samples were sent to the University of Georgia’s Department of Genetics for library preparation, sequence capture, and sequencing of Ultra conserved Elements (UCEs).

**Bioinformatics**

The samples were barcoded using Illumina TruSeq adapters with unique 8 base pair sequencetags for each individual. The UCEs were targeted by using a tetrapod 5060-locus probe set (fromultraconserved.org). The samples were de-multiplexed, filtered, and processed by removingadapter sequences and ambiguous bases using the program Illumiprocessor which is incorporatedin the software Phyluce v.1.5 (Faircloth 2015). Reads were assembled anew using standardsettings in Velvet v.1.1 (Zerbino and Birney 2008) and the assembled contigs were matched against the 5k UCE tetrapod probe kit to identify and extract UCEs using Phyluce. MAFFT   
v.7.397 (Katoh and Standley 2013) was used to generate sequence alignments for each individuallocus. To examine the effects of missing data on resulting topologies, concatenated dataalignment matrices of 50% and 75% were created. The data matrices represent the number ofalignments that include every individual (e.g., 75% data matrix would indicate that at least 75%of the alignments contain all individuals). The summary statistics of alignment matrices, totalbase pair reads, and UCE counts were generated using Python scripts from Phyluce. Threesamples that had more than 70% missing data in variable regions in UCE alignments wereexcluded.

**Phylogenetic Analyses**

We conducted phylogenetic analyses using a nucleotide substitution model from RevBayes software, v.1.1.1 (Hӧhna et. al. 2016). We used the general time reversible (GTR) model (Tavaré 1986) of sequence evolution, which allows six exchangeability rates between nucleotide states. The exchangeabilities are drawn from a Dirichlet distribution with the prior (1,1,1,1,1,1) which is 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 50,000 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 then summarized into maximum clade credibility trees (Helfrich et. al. 2018) and majority-rule consensus trees using RevBayes.

To delimit species, we used Bayesian Phylogenetics and Evolution, BPP (Yang & Rannala 2010), a genealogical method that estimates the time of origin, time of diversification, and the effective population multiplied by the mutation rate, for each species. We used the model A11 which estimates the species delimitation and the species tree (Fluori et. al. 2020). In this analysis the species delimitation model and the species phylogeny both change in the MCMC. The results showing posterior probability distributions indicate whether the lineages can be differentiated from each other. The subspecies of pine snakes were labelled according to their geographical distribution (FE – Far East; ME – Mid East; TN – Tennessee Valley and OG – Outgroup). The burnin was specified at 8000 and each dataset ran for 100000 generations. The output file contained probabilities of the best fit models and the arrangement of species labelled per their geographic location. The line containing all the probabilities for the best fit was extracted from the output files using a UNIX script and a histogram was created to visualize the number of species of pine snakes.

To create a concatenated tree for comparing consensus trees across UCEs, we then used the summarized trees from RevBayes (Höhna et. al. 2016) and built a consensus network in R (R core team) using the packages ape (Schliep & Paradis 2019), phangorn (Schliep 2011) and phytools (Revell 2012). The code and data for all the analyses are stored in GitHub (<https://github.com/basanta33/Pituophis_>).

**RESULTS**

**Sequence Data**

The average reads for each individual were 1,886,033 with a contig range of 7,136 to19,784. A range of 3,383 to 4,156 UCEs were recovered per individual with an average length of596 bp. The number of variable sites between individual UCEs ranged from 0-74 with anaverage of 7.18% of the total UCEs contained no variable sites and were omitted from the analysis. The majority of thevariation recovered was observed in the extreme regions of each UCE and there was littlecorrelation between the variable sites and locus lengths.

**Species Delimitation with BPP**

We carried out a study to examine whether there is any support for the recognition of subspecies within *Pituophis melanoleucus*. The model A11 accommodated for the gene tree uncertainty and variable population sizes over time to explore different species delimitation models and different species phylogenies. We processed the output from the datafiles with variable sites to obtain the posterior probabilities of different species groupings of pine snakes. When we combined the probabilities obtained from all the datasets to a single file and made a histogram, the plurality of the datasets indicated that there is an ingroup of one species of *Pituophis melanoleucus* and the outgroup of *P. ruthveni*, *P. catenifer*, and *Pantherophis obsoletus* (Figure 2). The posterior probabilities of delimitation were obtained based on a guide tree which accounts for the phylogenetic uncertainty.

**Consensus Network**

The output trees obtained from RevBayes were summarized to obtain maximum credibility trees and consensus trees. The consensus trees obtained were composed into a network of phylogenetic trees which reflected little to no variation amongst the subspecies of *P. melanoleucus*. The tree obtained from the phangorn package enabled us to evaluate the conflicting phylogenetic signals from the collected datasets. The tree (Figure 3) indicates that there is little variation among the topologies estimated from different pine snakes over the eastern side of the United States. A presence of variation would have the nodes of the trees connected to form a web-like structure.

**DISCUSSION**

**Corroboration of a single species of Eastern Pine Snakes**

Our results show that *Pituophis melanoleucus* is not composed of various distinct geographic lineages within the eastern United States. The Bayesian consensus tree (Figure 4) indicates that there is little to no variation among the subspecies of *P. melanoleucus*. Figure 3 also indicates that there is little to no phylogenetic structure among the samples collected from different parts of the eastern United States. This result agrees with Nikolakis’s study (2018) which used sequence-capture based approach to explore lineages of *P. melanoleucus*. The eastern pine snakes diverged within the *Pituophis* about 6 to 3 million years ago (Pyron and Burbrink 2009). The complex of *P. melanoleucus* appears to be of a single species with very little genetic differentiation, as indicated by the posterior probabilities of the best fit model for the species tree generated by BPP. In addition to the results from BPP, the majority rule consensus tree obtained from RevBayes also indicated that we get the same trees across the sites, a near-polytomy with little geographic structure (Fig. 3). The consensus network compiled using all the 4600 data files show that there is no connection between multiple tips and each node thus indicating little topology variation among the subspecies of pine snakes in the eastern region (Fig. 4).

**Gene flow across a geographic barrier**

The three different geographical lineages: far-eastern, mid-eastern and Tennessee clades would be separated by the Apalachicola/Chattahoochee River Basin acting as a geographical barrier (Nikolakis 2018). Previous phylogeographic studies across that region (Burbrink et al., 2008, 2000; Weinell and Austin, 2017) indicated that there is significant genetic variation among the clades existing across the barrier. Despite the geographical barrier, molecular species delimitation indicates that the three lineages of pine snakes are not distinct from each other. BPP should detect if there is gene flow across the barrier. In our analysis, BPP indicated that there was only one population 4650 times and two populations (one being the outgroup) 4654 times during the two runs of the datasets (Figure 2). The geographic barrier did not seem to have much effect on the isolation of the population of pine snakes across the mid-eastern and far-eastern sides. Due to the indication of little genetic differentiation, it can be concluded that gene flow has been maintained in the population of eastern pine snakes across that region. This disunity in the previous phylogeographic analyses (Burbrink et al. 2008, 2000; Weinell and Austin, 2017) of the snakes and our study could be due to the distribution of the species of the pine snakes across the eastern United States. Previous movement studies indicate that *P. melanoleucus* is a very mobile species with their home ranges spanning from approximately 35 hectares to over 105 hectares (Nikolakis 2018) which could contribute to the gene flow among the different populations.

**Use of UCE for recently diverged lineages**

Ultra-conserved Elements (UCEs) are highly conserved regions within the genome that are shared among evolutionarily distant taxa (Bejerano et al. 2004). The use of UCEs has been increasing in phylogeny inference across many vertebrate taxa (Gustafson et. al. 2019). Although UCE is an important molecular marker, for this complex the use of molecular markers other than UCE would be a better option. Some UCEs in this study did not even distinguish between the outgroup and the ingroup taxa, so we had very little information in those sites. This is due to recent evolution of the pine snakes from other *Pituophis*. The oldest fossils of *P. melanoleucus* have been found in Florida dating 0.8 to 2.5 mya and more northern fossils from Pennsylvania have been dated from 0.1 mya (Holman 2000). These data could indicate that the lineage diverged during the late Pleistocene when there were environmental fluctuations leading to periods of isolation and connection. These periods led to the maintenance of gene flow through the contacts of the different populations. As UCEs are conserved sequences, they evolve very slowly, thus decreasing the power to detect variation among the organisms that have recently diverged from its common ancestor (Winker et. al. 2018).

**CONCLUSION**

Using a dataset of ultra-conserved elements, we find no evidence for multiple subspecies in *Pituophis melanoleucus.* Using the multispecies coalescent and the phylogenetic species concept, we substantiate *Pituophis melanoleucus* as a single species across its whole range.

**LITERATURE CITED**

Andam, C.P., L. Challagundla, T. Azarian, W.P. Hanage, and D.A. Robinson. 2017. Population Structure of Pathogenic Bacteria. Genetics and Evolution of Infectious Diseases, Elsevier: 51–70.

Assis, L.C.S, O. Rieppel. 2011. Are monophyly and synapomorphy the same or different? Revisiting the role of morphology in phylogenetics. Cladistics 27(1):94-102.

Camargo, A., and J. Sites. 2013. Species Delimitation: A Decade After the Renaissance, p. 225- 247. *In:* The Species Problem - Ongoing Issues. Igor Pavlinov (eds.). InTech, 2013.

Andermann, T., M.F. Torres Jiménez, P. Matos-Maraví, R Batista, J.L. Blanco-Pastor, A.L.S. Gustafsson, L. Kistler, I.M. Liberal, B. Edelman, C.D. Bacon and A. Antonelli. 2020. A Guide to Carrying Out a Phylogenomic Target Sequence Capture Project. Frontiers in Genetics 10:1407.

Avise, J.C., A.C. Reeb, and N.C. Saunders. 1987. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). Evolution 41:991–1002.

Burbrink, F.T. 2002. Phylogeographic analysis of the cornsnake (*Elaphe guttata*) complex as inferred from maximum likelihood and Bayesian analyses. Molecular Phylogenetics and Evolution 25:465–476.

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.

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.

Crother, B. I. 2012. 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, 2012.

De Queiroz, K. 2007. Species Concepts and Species Delimitation. Systematic Biology 56:879–886.

Devitt, T. J., A.M. Wright, D.C. Cannatella, and D.M. Hillis. 2019. Species Delimitation in Endangered Groundwater Salamanders: Implications for Aquifer Management and Biodiversity Conservation. Proceedings of the National Academy of Sciences 116:2624–2633.

Donoghue, M.J. 1985. A Critique of the Biological Species Concept and Recommendations for a Phylogenetic Alternative. The Bryologist 88(3):172-181.

Felsenstein, J. 1981. Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. Journal of Molecular Evolution 17:368–376.

Flouri, T., J. Xiyun, B. Rannala, and Z. Yang. 2019. A Bayesian Implementation of the Multispecies Coalescent Model with Introgression for Phylogenomic Analysis. Molecular Biology and Evolution 37:1211–1223.

Flouri, T., B. Rannala, and Z. Yang. 2020. A Tutorial on the Use of BPP for Species Tree Estimation and Species Delimitation p.5.6:1-5.6:16. *In:* Phylogenetics in the Genomic Era. HAL-archives.

Flouri, T., J. Xiyun, B. Rannala, and Z. Yang. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Molecular Biology and Evolution 35:2585–2593.

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(1):19-34.

Gustafson, G.T., A. Alexander, J.S. Sproul, J.M. Pflug, D.R. Maddison, and A.E.Z. Short. 2019. Ultraconserved Element (UCE) Probe Set Design: Base Genome and Initial Design Parameters Critical for Optimization. Ecology and Evolution 9:6933–6948.

Guyer, C., and M.A. Bailey. 1993. Amphibians and reptiles of longleaf pine communities. p. 139–159. *In:* Proceedings of the Tall Timbers Fire Ecology Conference, No. 18. The Longleaf Pine Ecosystem: Ecology, Restoration and Management. Tall Timbers Research Station, Tallahassee, FL.

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

Hebert, P.D.N., A. Cywinska, S.L. Ball, and J.R. deWaard. 2003. Biological Identifications through DNA Barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270: 313–321.

Holland, B. R. 2004. Using Consensus Networks to Visualize Contradictory Evidence for Species Phylogeny. Molecular Biology and Evolution 21:1459–1461.

Höhna, S., M.J. Landis, T.A. Heath, B. Bissau, 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.

Korshunova, T., B. Picton, G. Furfaro, P. Mariottini, M. Pontes, J. Prkić, K. Fletcher, K. Malmberg, K. Lundin, and A. Martynov. 2019. Multilevel fine-scale diversity challenges the ‘Cryptic Species’ concept. Scientific Reports 9:1-23

Nikolakis, Zachary Lamar. 2018. Phylogenomics of the Eastern Pinesnake. Unpubl. M.S. diss., Southeastern Louisiana University, Hammond, Louisiana.

Patterson, C. 1982. Morphological characters and homology. p. 21-74. *In:* Problems of phylogenetic reconstruction. K. A. Joysey, and A. E. Friday, (eds.). Academic Press London.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E.. 2012. Double Digest RADseq: An Inexpensive Method for *De Novo* SNP Discovery and Genotyping in Model and Non-Model Species. PLOS ONE 7(5): e37135.

Pyron, A.R., and Frank B.T. 2009. Neogene Diversification and Taxonomic Stability in the Snake Tribe Lampropeltini (Serpentes: Colubridae). Molecular Phylogenetics and Evolution 52:524–529.

Rodrı́guez-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.

Reitzel, A.M., S. Herrera, M.J. Layden, M.Q. Martindale, and T.M. Shank. 2013. Going Where Traditional Markers Have Not Gone before: Utility of and promise for RAD Sequencing in Marine Invertebrate Phylogeography and Population Genomics. Molecular Ecology 22:2953–2970.

Ruane, S., R.W. Bryson, Jr., R.A. Pyron, and F.T. Burbrink. 2014. Coalescent Species Delimitation in Milksnakes (Genus Lampropeltis) and impacts on Phylogenetic Comparative Analyses. Systematic Biology 63:231–250.

Schliep, K., A.J. Potts, D.A. Morrison, and G.W. Grimm. 2017. Intertwining Phylogenetic Trees and Networks. Methods in Ecology and Evolution 8:1212–1220.

Schliep, K.P. 2011. Phangorn: Phylogenetic Analysis in R. Bioinformatics 27:592–593.

Scot, J. (2008). [Black Pine and Florida Pine Snake] [Photograph]. Flickr. https://www.flickr.com/photos/tamers1/3980065515

Solís-Lemus, C., L.L. Knowles, and C. Ané. 2014. Bayesian Species Delimitation Combining Multiple Genes and Traits in a Unified Framework. Evolution 69:492–507.

Tavaré S. 1986. Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. Some Mathematical Questions in Biology: DNA Sequence Analysis. 17:57–86.

Tuberville, T.D., J.D. Willson, M.E. Dorcas, and J.W. Gibbons. 2005. Herpetofaunal Species Richness of Southeastern National Parks. Southeastern Naturalist 4(3):537-569.

Sukumaran, J., and L.L. Knowles. 2017. Multispecies Coalescent Delimits Structure, Not Species. Proceedings of the National Academy of Sciences 114:1607–1612.

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.

Woodruff, D.S. 2001. Populations, Species, and Conservation Genetics. Encyclopedia of Biodiversity, Elsevier:811–829.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J Mol Evol 39:306–314.

Yang, Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc Natl Acad Sci USA 107(20):9264–9269.

Yang, Z. 2015. The BPP program for species tree estimation and species delimitation. Current Zoology 61(5):854-865.

Yang, Z., and B. Rannala. 2014. Unguided Species Delimitation Using DNA Sequence Data from Multiple Loci. Molecular Biology and Evolution 31:3125–3135.

Zhao, Y., Z. Yi, A. Warren, and W.B. Song. 2018. Species delimitation for the molecular taxonomy and ecology of the widely distributed microbial eukaryote genus *Euplotes* (Alveolata, Ciliophora). Proceedings Biological Sciences 285.