University of North Carolina Asheville
Journal of Undergraduate Research
Asheville, North Carolina
May 2019

Phylogenetics of the *Plethodon montanus* Species Complex with Mitochondrial DNA Analysis

Maaian Edery
Biology Major
The University of North Carolina Asheville
One University Heights
Asheville, North Carolina 28804 USA

Faculty Advisor: Dr. R. Graham Reynolds

Abstract

The Southern Appalachians are unique in the high diversity of salamanders that occur there. The genus *Plethodon* encompasses the terrestrial woodland salamanders that occur only in North America. This genus is the most diverse group of salamanders with 77 species currently described. Within this genus is the *Plethodon glutinosus* species complex, where several new species have been recently described, including the Gray-cheeked salamanders. Allozyme data have shown that there are four distinct Gray-cheeked salamander species comprising the *Plethodon montanus* species complex in Western North Carolina. These are the Northern Gray-cheeked (*P. monatnus*), Southern Gray-cheeked (*P. metcalfi*), Blue Ridge Gray-cheeked (*P. amplus*), and South Mountain Gray-cheeked (*P. meridianus*) salamanders. The little-understood geographic ranges of these species are the main distinction between them, though not much is currently known about the genetic relationships among these four species nor where each lineage actually occurs. Using samples collected by labmates, I isolated the mitochondrial gene ND2 from salamander tail tips of the four species via DNA extraction and PCR techniques and sent out for Sanger sequencing. I then constructed a series of phylogenetic trees to reveal the evolutionary history and relationships among these four species. This information helps us have a clearer idea of how closely related these species are, their unique evolutionary histories, and give us a better understanding of the current range of each species. These data also add to our understanding of the broader process of speciation in woodland salamanders. Importantly, we found that currently suggested species ranges do not correspond to mitochondrial lineages, and I offer a revision of the geographic range for each of these four species.

1. Introduction

The Southern Appalachians are unique in the high diversity of salamanders that occur there, with the highest salamander species diversity globally. Plethodontidae describes the terrestrial woodland salamanders that occur only in North America. This genus is the most diverse group of salamanders with 77 species currently described[1]. These species occur in mountainous forest with closed canopy, occupying burrows under covered objects for much of the year and venturing out onto the forest floor on warm rainy nights[2]. Within this genus is the sub group *Plethodon glutinosus*, where several new species have been recently described, including four recognized species of Gray-cheeked salamanders in Western North Carolina. These comprise what is currently described as the *Plethodon montanus* species complex, formerly included in the *Plethodon jordani* complex[3]. The four Gray-cheeked salamanders species are the Northern Gray-cheeked (*P. monatnus*), Southern Gray-cheeked (*P. metcalfi*), Blue Ridge Gray-cheeked (*P. amplus*), and Southern Mountain Gray-cheeked (*P. meridianus*) salamanders (Figs. 2,3). They are relatively small and identified by their uniform dark gray body and light gray mark behind their eyes. All four species are centered in Western North Carolina, with some populations extending into Tennessee, South Carolina, and Virginia. They inhabit high-elevation, moist, closed-canopy mixed deciduous forests, often under logs or rocks[1]. The four species exhibit no obvious morphological variation, making geographic range of these species the main distinction between them. These

ranges were defined in the original description of these species using allozyme data (Highton and Peabody 2000). The following are the current known ranges of the four Gray-cheeked species (see also Fig. 1):

Plethodon montanus occurs in southwest VA from Giles county southward. They can be found in the Flat Top, Buckhorn, Burkes Garden, Knob, Clinch, and Brumley Mountain Ridges within the Valley and Ridge province. They are in the Roan, Black, Bald, Max Patch, and Sandymush mountain ridges in the Blue Ridge Province. They also occur in Haywood County. This is the only species of the four that is restricted to higher mountains. *Plethodon amplus* occurs in the Blue Ridge Mountains in the counties of Buncombe, Henderson, and Rutherford. *Plethodon meridianus* occurs in the South Mountains in the counties of Burke, Cleveland, and Rutherford counties. *Plethodon metcalfi* occurs in the Blue Ridge Mountains in the counties of Haywood and Macon, NC, and Oconee, SC. This species intersects with both *P. amplus* in the southeast and *P. montanus* in the northwest. Though the exact locations of these contact zones are not known, several authors have posited the locations of these contact zones (Fig. 1) .

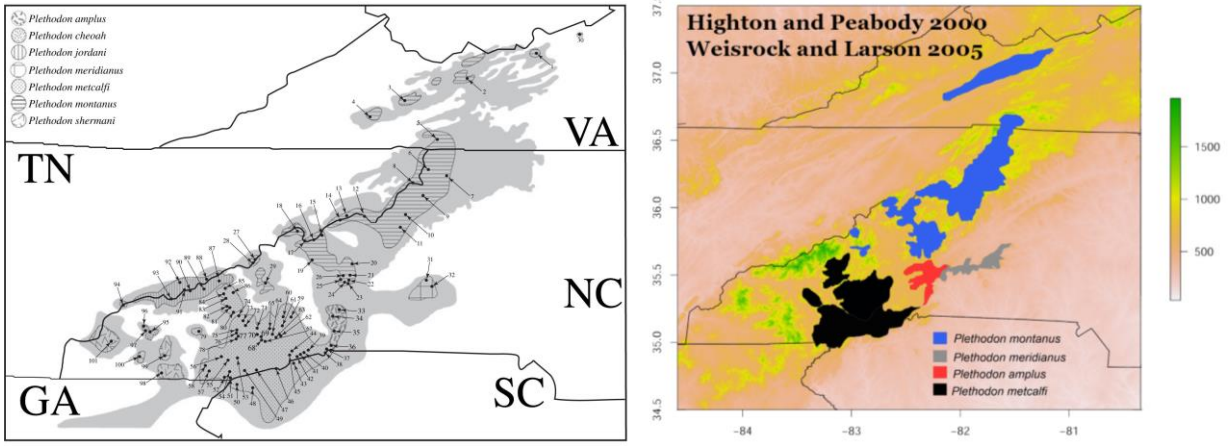

Figure 1. Distribution map of the *Plethodon jordani* (inclusive of the *P. montanus*) species complex (Weisrock and Larson, 2005); redrawn Highton & Peabody (2000; Fig. 1) (Left) and redrawn distribution map of *P. montanus* species complex ranges as described by Heighton and Peabody (2000) and Weisrock and Larson (2005) (Right).

Phylogenetic data suggest that a burst of speciation in the Miocene and Pliocene contributed to the large diversity of species within seen today in Plethodontidae[3]. Although many of the general relationships within the plethodontids are clearly defined, much of the terminal relationships, such as those of the *P. montanus* complex, remain largely unknown and disputed. The Gray-cheeked salamanders of Southern Appalachia were split into four distinct species based on allozyme data conducted by Highton and Peabody (2000)[4]. There are no distinguishing morphological traits that might be used to diagnose these species, instead the species are diagnosable almost entirely on their allozyme data, because this is how they were originally defined. Allozymes are variants of an enzyme that differ by single alleles within a single locus. Allelic variation arises from random mutations that occur over time at predictable rates, and can be used to identify how closely related species are. Allozyme data was used more frequently in early genetic studies, due to lack of access to genetic technologies used in isolating DNA. Importantly, allozyme studies can be less repeatable and offer less resolution than modern sequence-based methods.

Subsequently, Weisrock and Larson (2005) examined both additional allozyme data as well as mitochondrial DNA sequence data, using mitochondrial gene ND2, for the entire *P. jordani* species complex[5]. They sampled from 101 sites, which included numerous sites with the four Gray-cheek species. They found that allozymes continued to support recognition of these four Gray-cheek species, but the mtDNA did not. This is perhaps expected, given that the species were described based on allozyme data to begin with. While providing some overview of the evolutionary relationships of the Gray-cheeked Salamanders, they did not include detailed analysis of the implications of their results for this group. In addition, mtDNA show very clearly that the geography of allozyme variation does not correspond to molecular sequence data.

The mitochondrial data produced in the study by Weisrock and Larson were used to create a phylogenetic tree. In this tree, distinct clades were easily seen for *P. montanus*, *P. meridianus* and *P. amplus*. However, *P. metcalfi* individuals appeared throughout the tree among the three other groups, in addition to having its on clearly defined clade. I hypothesize that this is most likely due to misidentification of samples, or the discordance between the geography of allelic variation in the mitogenome versus the allozyme coding regions. Individuals were labeled as *P. metcalfi* due to the location at which they were collected and based on allozyme data but were not "true" *P. metcalfi*

based on phylogeny. Thus, they did not appear in the corresponding clade using sequence data. Given this, it seems that the currently described geographic areas of where the species occur and intersect is partly inaccurate. For example, there is likely an inaccurate division between the ranges of *P. mecalfi* and *P. amplus* shown in Figure 1. This causes doubt on the accuracy of the current identification based of allozyme data of the four Gray-cheeked species. Given these discrepancies, there is an opportunity to dive further into understanding the evolutionary relationships of these species. Through the acquisition and use of phylogenetic data derived from mtDNA we can create a more robust tree to explain the evolutionary history and relationships among these four species. These data will help us have a clearer idea of how closely related these species are and add to our understanding of the mechanisms contributing to the spectacular diversity of Plethodontid salamanders in this region.

Figure 2. *Plethodon metcalfi* (left; © Reynolds) and *Plethodon meridianus* (right; © Reynolds)

Figure 3. *Plethodon amplus* (left; ©R. Graham Reynolds) and *Plethodon montanus* (right; © Niemiller and Reynolds 2011).

2. Methods

2.1 Sample Collection

Tail tips of the Gray-cheeked salamander species complex were collected from various locations along their known range in North Carolina (Fig. 3) and were stored in 95% ethanol at -80 C. Samples were collected for *P. montanus* from Unaka Mountain (Yancey County) and Viking Mountain Road (Madison County). Samples of *P. metcalfi* were collected from Heintooga Road and Green River Gorge (Henderson County). *P. amplus* samples were collected from Bat Cave, CLC Florence Nature Reserve, and Bearwallow Mountain Trail (Henderson and Buncombe counties); and *P. meridianus* samples from South Mountain Game Lands (Burke County).

2.2 mtDNA Analysis

I extracted total genomic DNA from samples using the Wizard SV Genomic DNA Purification System. I performed a polymerase chain reaction amplification (PCR) for the mitochondrial DNA gene ND2 using the ND2-L4437-F2 forward primer (5’ -AAGCTTTCGGGCCCATACC- 3’) and the KND2-R2 reverse primer (5’ -

AAAGTTTGAGTTGCATTCA- 3'). This locus was chosen for analysis based on historical inertia, as data previously collected on the *Plethodon montnaus* complex used ND2. (Larson and Weisrock). PCR was carried out in a 25-µL reaction mixture containing 3 µL DNA template, 10.4 µL molecular biology-grade water, 2.5 µL 25mM $MgCl_2$, 5 µL 5X Colorless GoTaq Flexi Buffer, 1.5 µL DNTPs, 1.25 µL ND2-L4437-F2, 1.25 µL KND2-R2, and 0.125 µL GoTaq Flexi DNA Polymerase. PCR was performed according to the following program: 95.0°C for 2.5 minutes, 35 cycles of 95.0°C for 35 seconds, 59.0°C for 35 seconds, and 72.0°C for 2.5 minutes, 72.0°C for 2.5 minutes and incubated at 4.0°C until removal from PCR machine and storage in -20°C. Gel electrophoresis was run to check the quality of PCR products. Samples were then sent to the NC State Genomic Laboratory for sequencing.

2.3 Phylogenetic Analysis

Following sequencing, I manually cleaned up and aligned the sample sequences using Geneious® 10.2.6 (Biomatters, Auckland, NZ). I also imported ND2 sequence data for the four species from GenBank to supplement my own data and used these in my analysis. GenBank data were selected based on similar length of sequences once aligned. I used two methods to infer phylogenetic relationships among the *Plethodon montanus* complex. First, I used the RaxML plugin in Geneious to infer a maximum-likelihood phylogenetic consensus tree of the data. Next, I used the program Beauti v1.10.2 (Suchard et al. 2018)[6] to create an XML input file for the program BEAST v1.10.2 (Suchard et al. 2018). I used *jmodeltest* to determine which substitution model best fit the data and then BEAST to infer a time calibrated phylogenetic tree with a molecular clock (Bayesian Tree). I used a TIM+G model and added a standard vertebrate mtDNA rate of molecular evolution of 0.013% pairwise divergence per million years, a standard molecular clock for tetrapod mitochondrial genes (Macey et al. 1997)[7]. I further note that this molecular clock is intended to provide relative divergence times, not absolute divergence times, and because this is a molecular gene the times are actually coalescent times and not lineage divergence times. I loaded the xml file into BEAST and ran the analysis for 100 million generations.

2.4 Maps

I used RStudio v 1.1.447 running R v.3.5.3 to create all map figures. I downloaded a raster shapefile containing elevation data from the BioClim (https://www.worldclim.org/bioclim) database. I rendered the maps using the raster function from the R package raster (Hijmans and van Etten 2012) and the *readShapeSpatial* function in the R package maptools (Lewin-Koh et al. 2011)[8 9]

3. Results

3.1 Phylogenetics

From the 56 total samples I sequenced, I attained nine full-length sequences. This included two *P. amplus* from Bat Cave and one from CLC Florence, four *P. montanus* from Beauty Spot Road, two *P. metcalfi* from Heinatooga, and zero *P. meridianus* sequences. From nine novel sequences and 105 sequences downloaded from GenBank, I produced two trees to infer phylogenetic relationships among the *Plethodon montanus* complex species.

The Bayesian tree (Fig. 6) had eight distinct clades organized by species and geographic location. The earliest divergence occurred over 1.5 million years ago. *Plethodon amplus*, *P. meridianus* and *P. montanus* each had their own distinct clades, with a divergence of approximately 0.75 million years between the *P. amplus* and *P. merdinanus* clades. Several distinct clades were seen for *P. metcalfi*. These were separated into western, northern, and southern *P. metcalfi* populations. We also see a divergence of approximately 0.75 million years between the *P. montanus* and Northern *P. metcalfi* clade. An additional fourth *P. metcalfi* clade, which I called the central *P. metcalfi* population, occurs within the *P. montanus* group. Individuals were identified and named based on the current descriptions of geographic boundaries for the species, based of the allozyme data. By this classification system, the individuals in the central *P. metcalfi* clade were classified as such, however based on the mtDNA ND2 data, they appear within the *P. montanus* group. It is likely that these individuals, which were labeled as *P. metcalfi* due to where they were found, are instead actually *P. montanus*, as suggested by the mtDNA data. Our maximum-likelihood tree (Fig. 5) yielded a similar topology to the Bayesian tree, but we noted some topological aberrancies.

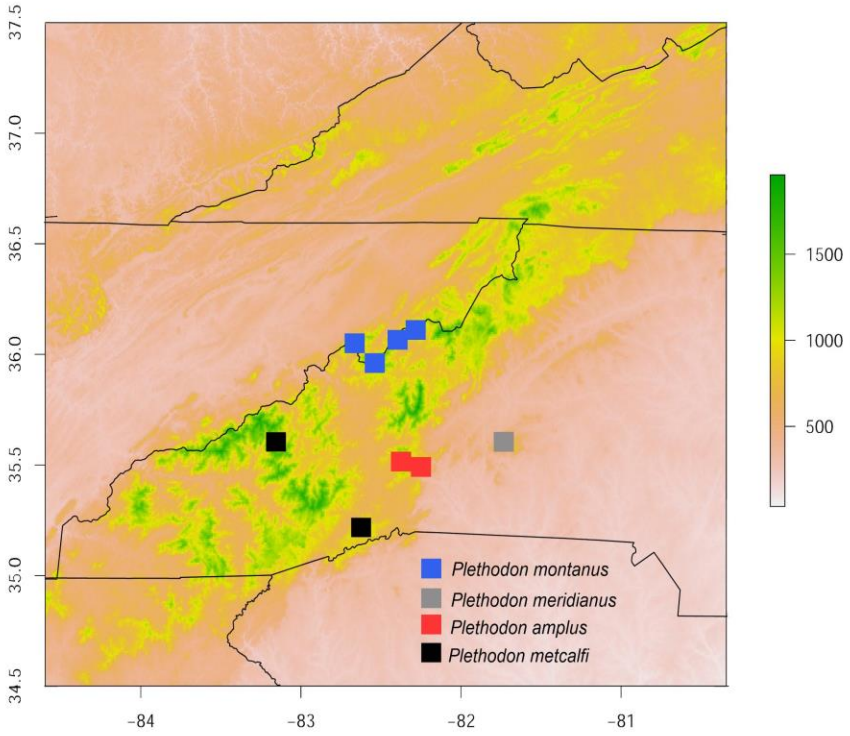

Figure 4. Sampling sites for the *Plethodon montanus* species complex samples collected by the Reynolds Lab to produce novel mtDNA ND2 sequences.

Figure 5. Maximum-likelihood phylogeny of 114 *P. montanus* species complex mtDNA sequences from the ND2 locus. Bootstrap values are given at respective nodes.

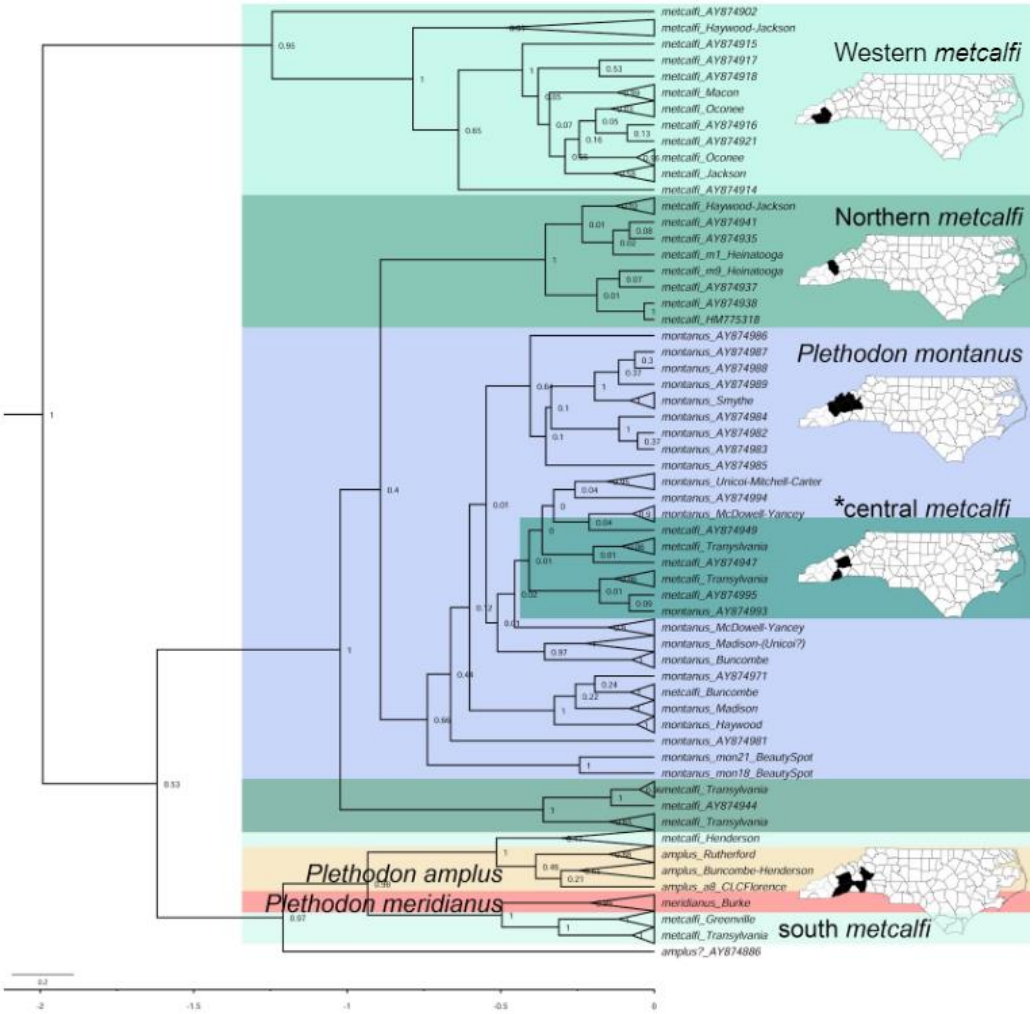

Figure 6. Molecular-clock calibrated Bayesian phylogeny of 114 *P. montanus species complex* mtDNA sequences from various geographic regions. Clades are collapsed and color coded by region. This includes 9 novel sequences and 105 sequences downloaded from GenBank.

4. Discussion

Our study's aim was to take a more focused look at the patterns of mtDNA variation in the Gray-cheeked Salamanders across Western North Carolina. We focused on increasing the sampling of known populations for the four species. From the forty PCR samples sent out only nine sequences came back usable for our data analysis. There are several possible explanations as to why such a large amount of my PCR sequences did not work. The DNA extractions for some samples were done in previous semester by other students, and it is possible they were not good extractions, i.e. they had low DNA amounts. Most likely, there was poor post-PCR cleanup at the sequencing lab to which the samples were sent. We then analyzed our newly generated data with that from Genbank to examine what mtDNA suggests about the geographic boundaries of the species.

The 2000 study by Highton and Peabody split the Gray-cheeked Salamanders into four separate species based on allozyme data[4]. They used allelic patterns among the allozymes of their sampled individuals to define the geographic boundaries of each species. In 2005, Larson and Weisrock used the same allozyme data and added mtDNA ND2 data to analyze the phylogenetic relationships of the four species. They found that the mtDNA data showed a rather different pattern of the geographic boundaries of the species than the allozyme data, especially with regard to the ranges of *P. montanus* and *P. metcalfi*[5]. Instead of revising the geographic boundaries of the recognized species, Larson and Weisrock reinforced the idea that allozymes define the species ranges and explained the discrepancies

between the allozyme and mtDNA data as hybridization ("hybrid introgression") between species[5]. This kind of reasoning is circular; the original definition of the four species and their ranges by allozymes was contradicted by the new mtDNA data, but rather than make adjustments, the existing lineages continued to be upheld and redefined based on those same allozyme allelic patterns.

We find that mtDNA data suggest that *P. montanus* occurs further south than previously suggested. Whether this is due to hybridization, as suggested by Larson and Weisrock (2005) or not is somewhat irrelevant given our basic question of where species alleles are found geographically. The genetic variation shows us a different geographic landscape of where these four species are occurring, which I displayed on a new map I made (Fig. 7). This map shows that two distinct populations of *P. montanus* occur farther south, between two populations of *P. metcalfi*. Based on the previous geographic map used to describe where species occurred (Fig 1.), these *P. montanus* populations were being inaccurately classified as *P. metcalfi*.

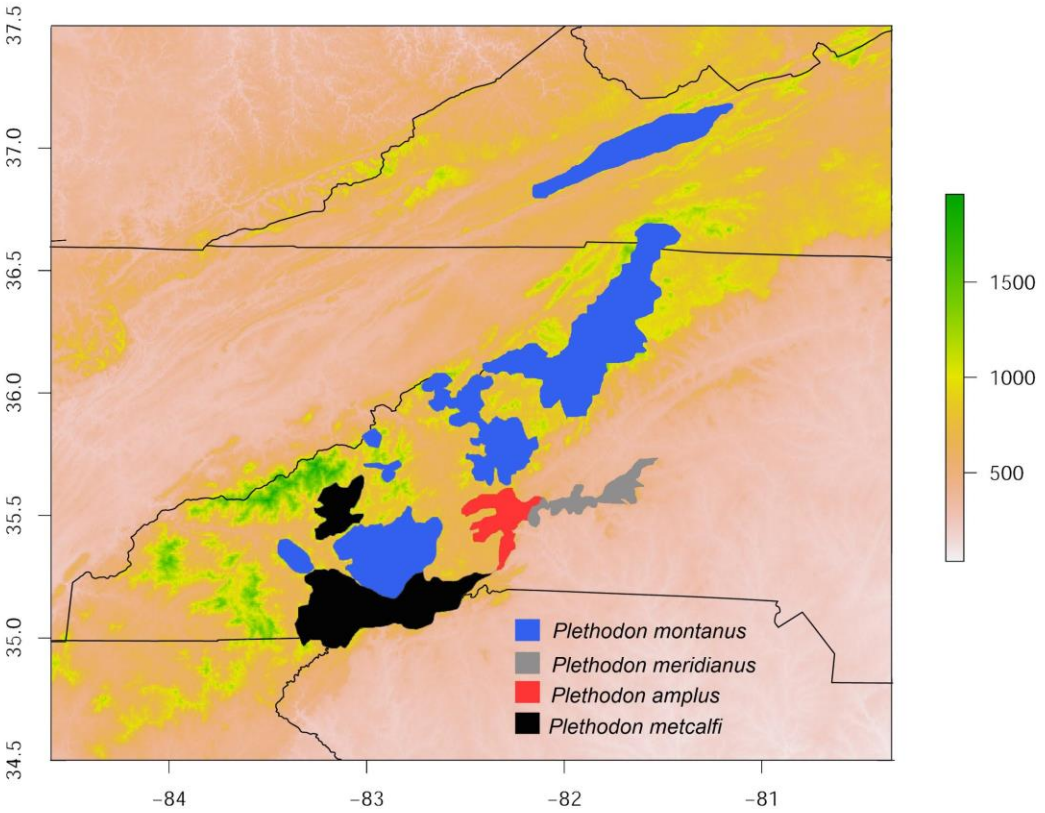

Figure 7. New proposed species ranges based on the distribution of mtDNA haplotypes.

Of course, even my data are only represented by a single mitochondrial locus, therefore in order to fully understand species boundaries in general, and evolutionary process such as hybridization, a more expansive approach incorporating data from the nuclear genome is needed. Nevertheless, I hope that by highlighting the reliance on one type of data to define species, we are constraining ourselves to a system of classification that is not necessarily reflective of the biology or evolutionary history of these lineages.

5. Acknowledgements

I thank Dr. Reynolds and the Reynolds Lab for collecting samples used in this study, in particular Ari Miller, Shannon Bodeau, and Robert Chambliss.

6. References

1. Mitchell J, Gibbons W. Salamanders of The South East. Athens, Georgia: The University of Georgia Press; 2010.

2. Kozak, Kenneth H., et al. 2006. Rapid Lineage Accumulation in a Non-Adaptive Radiation: Phylogenetic Analysis of Diversification Rates in Eastern North American Woodland Salamanders (Plethodontidae: Plethodon). Proceedings: Biological Sciences. 273(1586):539–546.

3. Highton R, Picard Hastings A., et al. 2011. Concurrent speciation in the eastern woodland salamanders (Genus Plethodon): DNA sequences of the complete albumin nuclear and partials mitochondrial 12s genes. Molecular Phylogenetics and Evolution. 63(2):278-290.

4. Highton R., Peabody R. B. 2000. Geographic Protein Variation and Speciation in Salamanders of the Plethodon jordani and Plethodon glutinosus Complexes in the Southern Appalachian Mountains with the Description of Four New Species. In: Bruce C. R, Jaeger G. R, Houck D. L., editors. The Biology of Plethodontid Salamanders. First Edition. New York: Kluwer Academic/Plenum Publishers. Pgs 58-72.

5. Weisrock D, Larson A. 2005. Testing hypotheses of speciation in the Plethodon jordani species complex with allozymes and mitochondrial DNA sequences. Biological Journal of the Linnean Society. 89:25–51.

6. Suchard M.A., Lemey P., Baele G., Ayres D. L., Drummond A. J., Andrew Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution. 4(1).

7. Macey, J.R., Larson, A., Ananjeva, N.B. and Papenfuss, T.J. 1997. Evolutionary shifts in three major structural features of the mitochondrial genome among iguanian lizards. Journal of Molecular Evolution, 44(6):660-674.

8. Hijmans R. J., van Etten J. 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-12. http://CRAN.R-project.org/package=raster

9. Lewin-Koh et al. 2011. maptools: Tools for reading and handling spatial objects. R package version 0.8-10, URL http://CRAN. R-project. org/package= maptools.

10. Wiens, John J., et al. 2006. Rapid Diversification, Incomplete Isolation, and the ‘Speciation Clock’ in North American Salamanders (Genus Plethodon): Testing the Hybrid Swarm Hypothesis of Rapid Radiation. Evolution. 60(12):2585–2603.