2022

# Hybridization between the Woodland Salamanders Plethodon cinereus and P. electromorphus Is Not Widespread

Shawn R. Kuchta

Maggie M. Hantak

Brian P. Waldron

Cari-Ann M. Hickerson

Richard M. Lehtinen

*See next page for additional authors*

## Authors

Shawn R. Kuchta, Maggie M. Hantak, Brian P. Waldron, Cari-Ann M. Hickerson, Richard M. Lehtinen, and Carl Anthony

*Ichthyology & Herpetology* 110, No. 3, 2022, 430–438

# Hybridization between the Woodland Salamanders *Plethodon cinereus* and *P. electromorphus* Is Not Widespread

**Shawn R. Kuchta[1], Maggie M. Hantak[2], Brian P. Waldron[1], Cari-Ann M. Hickerson[3], Richard M. Lehtinen[4], and Carl D. Anthony[3]**

**A recent study reported widespread hybridization between the Eastern Red-backed Salamander (*Plethodon cinereus*) and the Northern Ravine Salamander (*P. electromorphus*) in northern Ohio. In this study, DNA sequence data were obtained from three nuclear loci and 20 single nucleotide polymorphisms (SNPs) were identified from the sequences. They found that 48 out of 90 individuals from 13 populations were hybrids, and in some localities every individual possessed an admixed genotype. As these results contradict our observations, and because levels of hybridization impact our interpretation of past and ongoing studies, we revisited the data. First we reanalyzed the original SNPs using STRUCTURE, then we repeated the analysis using haplotypes instead of SNPs. We found that K = 2 was best supported by both analyses, and they agree in recovering lower levels of hybridization than originally reported. For example, five populations in the original study identified as highly admixed or composed entirely of admixed genotypes we found to be pure *P. cinereus* or to lack evidence of extensive admixture. Similar results were obtained using NEWHYBRIDS and analyses based on gene trees. We conclude that while hybridization between *P. cinereus* and *P. electromorphus* occurs, it is much more restricted than originally reported.**

A diversity of outcomes is possible when divergent groups meet, from sympatry without hybridization, to a narrow hybrid zone, to a hybrid swarm (Barton and Hewitt, 1989; Kuchta, 2007; Pereira and Wake, 2009). In extreme cases, hybridization can lead to the merger of formerly divergent lineages (Jockusch and Wake, 2002; Kleindorfer et al., 2014), resulting in the loss of biodiversity (Seehausen, 2006). While degrees of hybridization are variable across the tree of life, studies have consistently found that it is common, with at least 25% of plant species and 10% of animal species hybridizing with some other species (Mallet, 2005).

In this paper, we report on a reanalysis of hybridization between the Eastern Red-backed Salamander, *Plethodon cinereus*, and the Northern Ravine Salamander, *P. electromorphus*. Some hybridization between the species was apparent from the start, as in the original description of *P. electromorphus* Highton (1999) documented hybridization with *P. cinereus* using allozyme data. In that study, *P. cinereus* was found in 3 out of 16 localities with *P. electromorphus*. No hybridization was detected at two of the localities, but in Wayne County, Ohio, ten hybrid individuals were recovered in a sample that included $\geq 70$ of each parental species. One of the hybrids was a possible F1, while the others were backcrosses. Similarly, using microsatellite data, Waldron and Hantak (2020) reported on a single F1 in a sample of 23 individuals from Lorain County, Ohio. Hybridization between *P. cinereus* and *P. electromorphus* is interesting in part because the two species are not closely related. Both are members of the "*cinereus* group" (clade) within eastern *Plethodon* (Highton, 1995; Highton et al., 2012), which includes eight other species. However, *P. cinereus* and *P. electromorphus* are not sister taxa (Sites et al., 2004; Fisher-Reid and Wiens, 2011), and molecular clock studies suggest they diverged > 10 myr ago (Kozak et al., 2006; Wiens et al., 2006).

In an effort to better document the extent of hybridization, Lehtinen et al. (2016; hereafter "L16") analyzed 90 individuals from 13 populations in northeastern Ohio. In contrast with Highton (1999), they found evidence for extensive and geographically widespread hybridization and speculated that "these two lineages may be in the process of merging back into a single gene pool." In this paper, we reanalyze the data from L16 (hereafter the "linked data") and find that the amount of hybridization is substantially lower than was originally reported.

The linked data were obtained from short DNA sequences from three nuclear loci. From these loci, 20 diagnostic single nucleotide polymorphisms (SNPs) were selected and analyzed using the admixture model in the Bayesian clustering program STRUCTURE 2.3.4 (Pritchard et al., 2000). The population assignment probability (Q score) was used to assign individuals to groups: individuals were considered pure parental types if they had Q scores > 0.90 for either *P. cinereus* or *P. electromorphus*, and all remaining individuals were considered to be of mixed ancestry. Overall, 53% of the individuals analyzed had admixed genotypes, while 27% were pure *P. electromorphus* and 20% were pure *P. cinereus*. In addition, and in sharp contrast with Highton (1999), all syntopic sites in L16 were found to include hybrid individuals, and four localities were found to be composed entirely of hybrids.

***The problem with linkage disequilibrium.***—The linked data included 20 SNPs from three nuclear loci, but in many cases

[1] Ohio Center for Ecological and Evolutionary Studies, Department of Biological Sciences, Ohio University, Irvine 107, Athens, Ohio 45701; Email: (SRK) kuchta@ohio.edu; and (BPW) bw120818@ohio.edu. Send correspondence to SRK.
[2] Department of Natural History, Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611; Email: mhantak@ufl.edu.
[3] Department of Biology, John Carroll University, University Heights, Ohio 44118; Email: (CMH) chickerson@jcu.edu; and (CDA) canthony@jcu.edu.
[4] Department of Biology, The College of Wooster, 931 College Mall, Wooster, Ohio 44691; Email: rlehtinen@wooster.edu.
Submitted: 25 June 2021. Accepted: 11 January 2022. Associate Editor: B. L. Stuart.
 DOI: 10.1643/h2021081 Published online: 5 August 2022

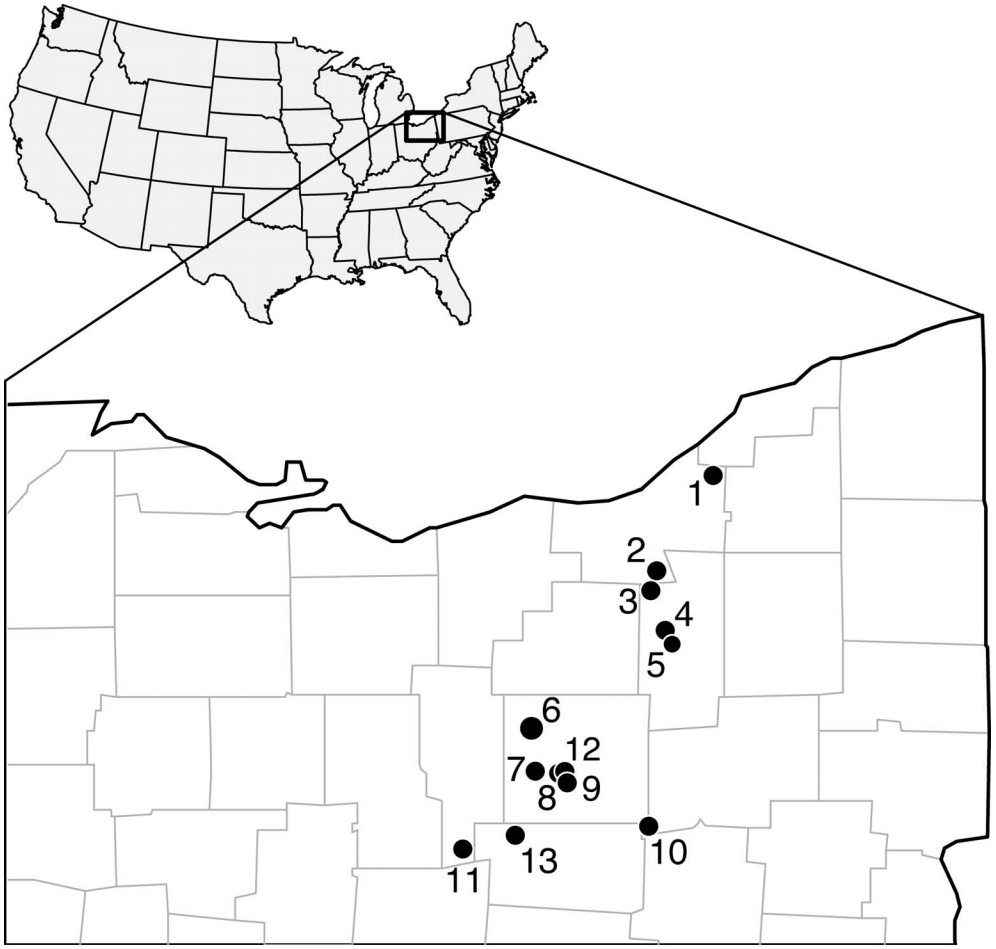


**Fig. 1.** Map of the sampled localities in central and northern Ohio. Localities are numbered, match Table 1, and are used throughout the text.

SNPs were only a few base pairs apart or even adjacent to one another. This is problematic because SNPs located even thousands of base pairs apart commonly exhibit high levels of linkage disequilibrium (Hohenlohe et al., 2012; Willis et al., 2017; O'Leary et al., 2018), yet most analyses assume molecular markers are statistically independent. The inclusion of tightly linked SNPs can result in false signals consistent with admixture and population structure (Willis et al., 2017; O'Leary et al., 2018). L16 inferred admixture using the program STRUCTURE, which estimates cluster membership by maximizing Hardy Weinberg equilibrium within all loci, while minimizing linkage disequilibrium between loci. That is, a key assumption of the program is that all markers provide independent information regarding ancestry. To quote the manual, "…if the data are dominated by one or a few non- or low-recombining regions, then STRUCTURE could be seriously misled" (Pritchard et al., 2010).

***The problem with K.***—Given genotypic data, STRUCTURE assigns individuals to K discrete clusters, where each cluster corresponds to a Mendelian population characterized by a set of allele frequencies at each locus. However, the choice of K is often not clear cut. One approach is to use Equation 12 in Pritchard et al. (2000) to calculate the estimated log probability of the data $\Pr(X|K)$ for various K, and select the K with the highest probability. Alternatively, the *ad hoc* statistic ΔK, developed by Evanno et al. (2005), can be used to infer the uppermost hierarchical level of population structure (thus requiring nested analyses to detect substructure; Converse et al., 2015; Janes et al., 2017). Using the linked data, L16 found that $\Pr(X|K)$ and ΔK were both maximized at $K = 5$. This is a surprising result given that two divergent species were being compared, suggesting ΔK, at least, should equal two.

In this paper we report on reanalyses of the data from L16. We include analyses that do not violate the assumption of linkage disequilibrium, but at the cost of fewer loci, and we evaluate the results of $K = 2$. We also conduct complementary analyses using NEWHYBRIDS and investigations of gene trees. We conclude that while hybridization between *P. cinereus* and *P. electromorphus* occurs, it is far less extensive than reported by L16. The species are not as risk of "merging back into a single gene pool."

## MATERIALS AND METHODS

***Sampling.***—The distributions of *P. cinereus* and *P. electromorphus* in Ohio are complex (figures 19-6 and 19-7 in Pfingsten, 2013; Deitloff, 2011). For their study, L16 obtained genetic data from 13 sites (90 individuals) from central and northern Ohio (Fig. 1). At 11 of these sites, samples were collected without knowledge of whether hybridization was present or not, but during field sampling an effort was made to find specimens that had traits consistent with *P. cinereus* and *P. electromorphus*, as well as morphological intermediates. Sample sizes ranged from 2–8 individuals per site, except site 7, from which 36 individuals were sampled (Table 1). L16

**Table 1.** Locality information for samples of *Plethodon cinereus* and *P. electromorphus*. Population numbers correspond to Figure 1 and are used throughout the text.

| Population | # Individuals | County | Locality | Latitude | Longitude |
|---|---|---|---|---|---|
| 1 | 2 | Lake | North Chagrin Reservation | 41.563160 | –81.430960 |
| 2 | 5 | Cuyahoga | Brecksville Reservation | 41.321310 | –81.618270 |
| 3 | 2 | Summit | Furnace Run Metropark | 41.271000 | –81.638460 |
| 4 | 4 | Summit | O'Neill Woods Metropark | 41.169900 | –81.590460 |
| 5 | 3 | Summit | Sand Run Metropark | 41.135000 | –81.567860 |
| 6 | 4 | Wayne | Pee Wee Hollow Boy Scout Camp | 40.921770 | –82.035920 |
| 7 | 36 | Wayne | Wooster Memorial Park | 40.812350 | –82.023230 |
| 8 | 3 | Wayne | Christmas Run Park | 40.806610 | –81.945230 |
| 9 | 5 | Wayne | Secrest Arboretum | 40.783270 | –81.916750 |
| 10 | 6 | Stark | The Wilderness Center | 40.672240 | –81.645090 |
| 11 | 6 | Ashland | Mohican State Park | 40.615220 | –82.264520 |
| 12 | 8 | Wayne | College of Wooster Golf Course | 40.811480 | –81.924310 |
| 13 | 5 | Holmes | Fern Valley Field Station | 40.649470 | –82.089660 |

sequenced up to three nuclear loci from each individual, and from these loci 20 putatively diagnostic SNPs were selected: recombination activation gene-1 (RAG1; 385 base pairs; 5 SNPs), pro-opiomelanocortin (POMC; 451 bp; 6 SNPs), and glyceraldehyde-3-phosphate dehydrogenase (GAPD; 590 bp; 9 SNPs).

In our reanalyses, all individuals used to construct the linked data were included except for RML 128 (population 11), from which the data has been lost. Before conducting analyses, we tested for intragenic recombination in all three loci using the difference in sum-of-squares (DSS) test implemented in TOPALi (Milne et al., 2009), including a ten-base-pair increment, a window size of 100, and 500 parametric bootstraps. We did not detect recombination at any locus, indicating that the SNPs from each locus are in complete linkage disequilibrium. We tested for hybridization by examining gene trees, as well as using STRUCTURE and NEWHYBRIDS.

***Gene trees.***—Prior to inferring gene trees from each locus, we phased the sequence data using PHASE v2.1.1 (Stephens et al., 2001). PHASE was run for 1000 iterations, with a thinning interval of two steps and a burn-in of 100 iterations. Homozygous loci were doubled in our files so that each individual included two haplotypes.

After each locus was phased, gene trees were inferred. In preliminary analyses, we included sequences from *P. cinereus* from GenBank and from Radomski et al. (2020) to distinguish between the *P. cinereus* and *P. electromorphus* clades, but for our final analyses these extra sequences were omitted. Sequence data were aligned using default conditions in MUSCLE (Edgar, 2004). For the introns POMC and GAPD, models of evolution were inferred using jModelTest 2.1.10 (Darriba et al., 2012), with the best model selected using AICc. For POMC, the best model was K80 + I, and for GAPD the best model was TPM2. For the exon RAG1, we selected a best-fit partitioning scheme using PartitionFinder v2.1.1 (Lanfear et al., 2012), with codon positions input as separate data blocks. The best scheme according to AICc combined all codons, and the JC model was selected. For the GAPD gene tree, sequences from RML 135 were omitted in the final analyses, and for RAG1 sequences from RML165 were omitted. This is because in both cases the haplotypes were recovered as distantly related to *P. cinereus* and *P. electromorphus* and thus were not informative for the present analysis. The source of these divergent sequences is unclear.

We inferred gene trees separately for each locus using BEAST v.2.6.3 (Drummond and Bouckaert, 2015). Multiple preliminary analyses were conducted to explore a diversity of priors. For all gene trees, a constant coalescent tree prior and a strict molecular clock were used. We ran the Markov chain Monte Carlo (MCMC) for 50 million generations with parameters sampled every 1000 generations, trees sampled every 5000 generations, and a burn-in of 25%. Effective samples sizes for all parameters in all runs were $> 200$. The maximum clade credibility tree with common ancestor node heights was obtained using TreeAnnotator 2.4.6 (Drummond and Bouckaert, 2015).

Given the >10 myr divergence times between the two species, we expected each gene tree to form two divergent clades. For all three nuclear genes, clades were recovered that included either mostly haplotypes sampled from *P. cinereus* or mostly haplotypes sampled from *P. electromorphus* (Supplemental Figs. 1–3; see Data Accessibility). Individuals with genotypes entirely composed of alleles that corresponded with a particular species were scored as parental types (*P. cinereus* or *P. electromorphus*), while individuals with genotypes composed of any kind of mixed ancestry (e.g., five *P. cinereus* haplotypes and one *P. electromorphus* haplotype) were considered hybrids. This approach using gene trees has the advantage of accounting for phylogeny, which methods based on allele frequencies typically ignore.

***Structure.***—Two datasets were analyzed. First, we reanalyzed the linked data from L16. This dataset has 20 SNPs, but suffers from high levels of linkage disequilibrium. Second, we inferred haplotypes for each locus using the *haplotype* function in the 'haplotypes' package in R (Aktas, 2020) and treated these as alleles (Willis et al., 2017; Kuchta et al., 2018). This new dataset (hereafter, the "unlinked data") is free of linkage disequilibrium, but only includes three loci. For all analyses, STRUCTURE was run from K = 1–10 populations, with each value of K replicated ten times with randomly generated starting seeds. Each MCMC run consisted of 500,000 iterations, with the first 100,000 discarded as burn-in. The admixture model, an inferred $\alpha$, and fixed $\lambda = 1$ were used, and sampling localities were used as priors. We also used a population-specific ancestry prior and an initial alpha value of 0.5 (as recommended for K = 2; Wang, 2017).

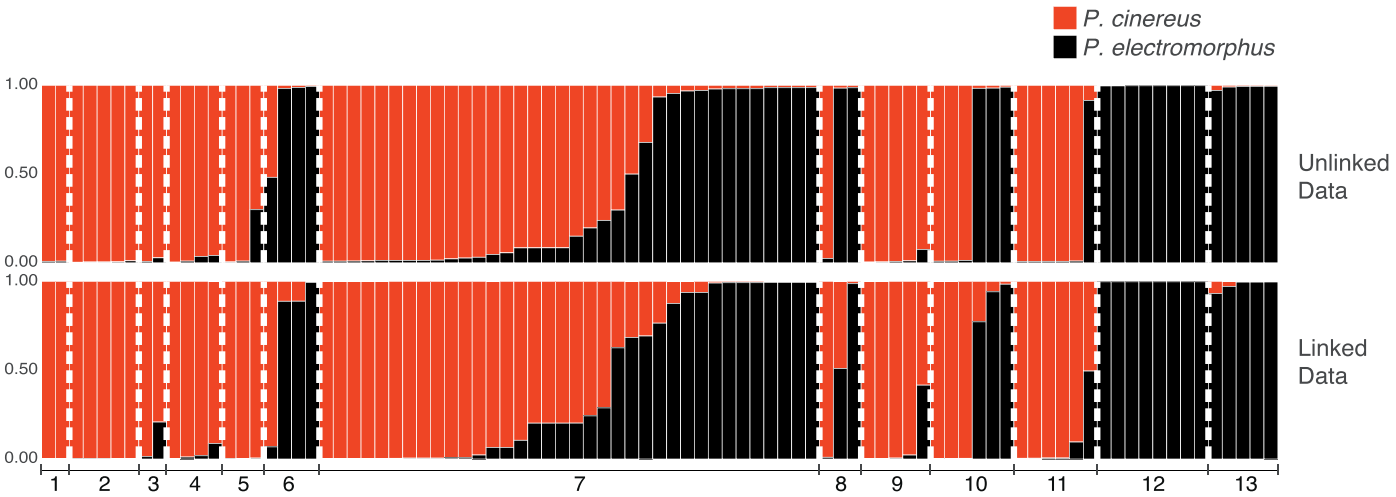


**Fig. 2.** STRUCTURE results using nuclear sequence data coded as haplotypes (unlinked) and as SNPs (linked). Results are for K = 2. The red group corresponds with *P. cinereus* and the black group corresponds with *P. electromorphus*. Individuals with Q scores < 0.9 for either parental type were scored as hybrids.

We did not use the correlated allele frequency model, as *P. cinereus* and *P. electromorphus* represent deeply divergent populations. We used the log probability of the data Pr($X|K$) and the Evanno method (Evanno et al., 2005) to select K. Our STRUCTURE results were collated, analyzed, and visualized using the R package 'Pophelper' (Francis, 2017). For consistency with L16, all individuals with membership coefficients (Q) > 0.90 were classified as parental types, while all others were scored as hybrids.

***NEWHYBRIDS.***—As an additional approach to identifying hybrid individuals, we used the program NEWHBRIDS (Anderson and Thompson, 2002) to analyze the unlinked data. This program computes the posterior probabilities of individuals belonging to each parental type, as well as distinct hybrid classes such as parental, F1, F2, and backcrosses. After multiple preliminary analyses to evaluate the impact of alternative priors, we carried out a final analysis of 5,000,000 MCMC iterations after a burn-in of 1,000,000 iterations. Analyses used Jeffreys-like priors on allele frequencies and admixture proportions, and the default genotype frequency classes were used. Individuals with posterior probabilities ≥0.90 of belonging to either species were considered parental types. If the sum of the posterior probabilities of all the possible hybrid classes was ≥0.90, individuals were scored as a hybrid. All other individuals were scored as unresolved. Among our analyses, only NEWHYBRIDS accounts for uncertainty in classification.

## RESULTS

***Linked data in STRUCTURE.***—Using the linked data from L16 ($n$ = 20 SNPs), we found ΔK = 2 using the Evanno method and K = 3 using Pr($X|K$); however, for the latter method the log probabilities are similar for K = 2–10 (Supplemental Fig. 4; see Data Accessibility). Given that we are examining hybridization between two distinct species, we focus on K = 2 (see Discussion).

Our analysis recovered lower levels of admixture than reported by L16 (Fig. 2). All individuals at sites 1, 2, 4, and 5 were identified as pure *P. cinereus*, and all individuals at sites 12–13 were identified as pure *P. electromorphus*. Some admixture was found at all syntopic sites. Single hybrid individuals were identified at sites 3, 9, and 11, but otherwise these sites only included *P. cinereus*. Populations 6–8 and 10–11 included both species as well as hybrid individuals. Site 7 exhibited especially high levels of admixture, with 12 out of 36 (33%) individuals categorized as hybrids.

We also explored K = 5 for the linked data, as this was the K used by L16 (Supplemental Fig. 5; see Data Accessibility). Here too we found lower levels of admixture than L16, but substantially more than for K = 2. For example, we recovered populations 1 and 2 as pure *P. cinereus*, while L16 reported both populations as 100% admixed. Most other populations in our analyses included one or more hybrids, but generally fewer than reported in L16. One exception is all individuals in populations 12–13 in L16 were scored as pure *P. electromorphus* (by definition, as these populations were used to assign SNPs to *P. electromorphus*), but we recovered two individuals as admixed in population 13.

***Unlinked data in STRUCTURE.***—For comparison with L16, we analyzed DNA haplotypes in STRUCTURE, which lowers the number of loci from 20 to 3, but eliminates the problem of linkage disequilibrium. We found ΔK = 2 using the Evanno method and K = 3 using Pr($X|K$); however, for the latter method the log probabilities are similar for K = 2–4 (Supplemental Fig. 4; see Data Accessibility). Again, we recovered lower levels of admixture than L16 (Fig. 2). All individuals in populations 1–4 and 9 were pure *P. cinereus*, and all individuals in populations 12–13 were pure *P. electromorphus*. Single hybrid individuals were found in populations 5 and 6. Syntopic sites without any evidence of admixture included populations 8 and 10–11. Only population 7 exhibited high levels of admixture, with 6 out of 36 (17%) individuals categorized as hybrids.

When we analyzed the unlinked data for K = 5, we found less admixture than L16, but substantially more than for K = 2. Overall, the results were similar to K = 5 for the linked data, though the hybrid classes differ (Supplemental Fig. 5; see Data Accessibility).

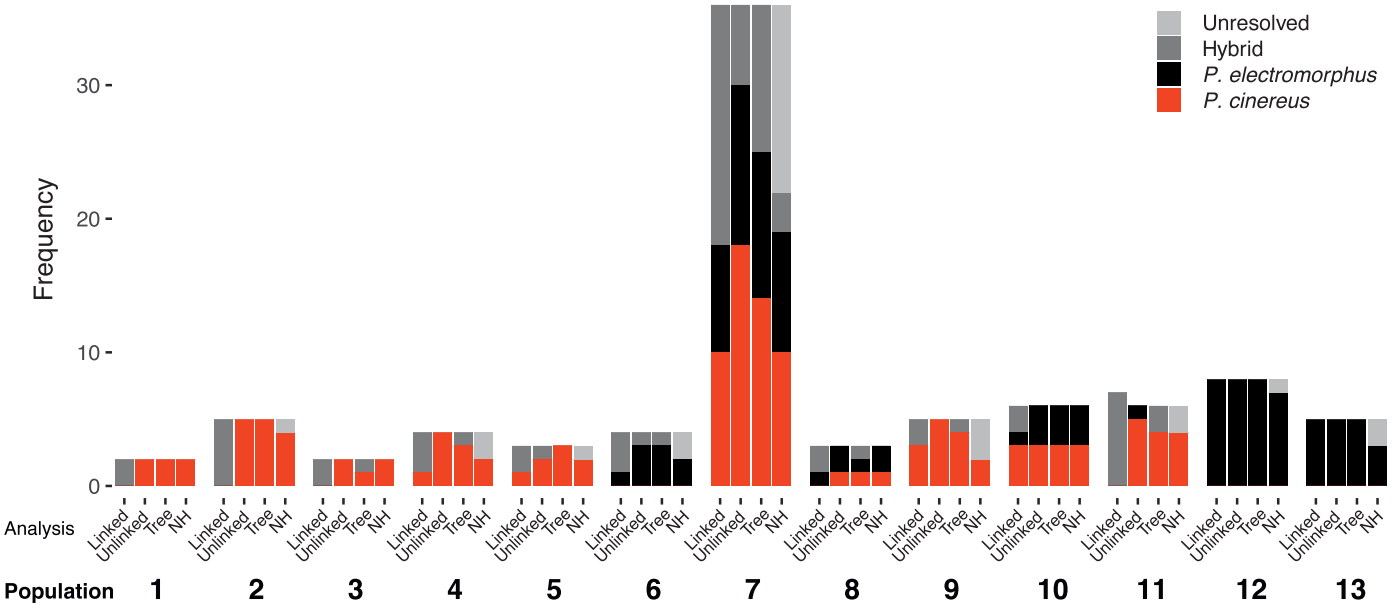


**Fig. 3.** Bar chart summarizing the assignments of individuals to parental and hybrid classes. Analyses include the linked data in STRUCTURE (K = 5; L16), unlinked data (haplotypes) in STRUCTURE (K = 2; this paper), analysis of gene trees (Tree), and unlinked data in NEWHYBRIDS (NH). For the last analysis, some individuals were not assigned to any class (parental or hybrid) with confidence ('Unresolved'). Localities correspond to Table 1 and Figure 1. See also Supplemental Table 1 (see Data Accessibility).

***Gene trees.***—For POMC and GAPD, the basal split in the gene tree separated haplotypes belonging to *P. cinereus* from haplotypes belonging to *P. electromorphus* (pp > 0.95, except for the *P. electromorphus* clade in POMC, which was pp = 0.82; Supplemental Figs. 1–2; see Data Accessibility). For POMC, we had sequence data for 80 out of 89 individuals in the dataset. Of these, 37 possessed two haplotypes from the *P. cinereus* clade, 33 possessed two haplotypes from the *P. electromorphus* clade, and ten possessed one haplotype from each clade (putative hybrids). For GAPD, we had sequence data from 53 individuals. Of these, 37 possessed two haplotypes from the *P. cinereus* clade and 16 possessed two haplotypes from the *P. electromorphus* clade. No individual possessed a hybrid genotype, which suggests that our dataset did not include any F1 hybrids.

For RAG1, we had sequence data from 88 individuals. While a *P. electromorphus* clade was recovered (pp = 0.96), the *P. cinereus* clade lacked statistical support (pp = 0.79; Supplemental Fig. 3; see Data Accessibility). Within the *P. cinereus* clade, one subclade that included 13 haplotypes from six sample localities may actually belong to *P. electromorphus*. By assuming this clade belongs to *P. electromorphus*, which is possible given the low posterior probabilities, ten individuals change status from putative hybrid to parental *P. electromorphus* (all haplotypes from all loci belonging to *P. electromorphus*). Assuming this is correct, we recover 49 individuals of *P. cinereus*, 34 of *P. electromorphus*, and 5 admixed genotypes. If this is not correct, ten more admixed genotypes are recorded (far more than POMC or GAPD), but our qualitative message in this paper does not change.

We assumed that all individuals with any evidence of admixture were hybrids. Analyzed by population, our results were congruent with the STRUCTURE analysis of the unlinked data for K = 2 (Figs. 2, 3). Populations 1–2 were recovered as pure *P. cinereus*, and 12–13 were recovered as pure *P. electromorphus*. Single hybrid individuals were found in populations 3–4, 6, and 8–9, while population 10 lacked evidence of hybridization. In population 7, 11 out of 36 individuals (31%) were scored as hybrids.

***NEWHYBRIDS.***—These results differed from the other analyses in that the status of some individuals was unresolved (the posterior probability of being either parental type or some type of hybrid was < 0.90). The NEWHYBRIDS results paralleled the analyses of the unlinked data using STRUCTURE (Figs. 2, 3). Populations 1 and 3 were recovered as pure *P. cinereus*, while populations 2, 4, and 5 were pure *P. cinereus* except for single individuals that were unresolved. For these unresolved individuals, it was unclear whether they were *P. cinereus* or some sort of hybrid, but the posterior probability of any being *P. electromorphus* was < 0.01. Populations 9 and 11 were also composed of *P. cinereus* and unresolved genotypes. Populations 6 and 12–13 were either pure *P. electromorphus* or included unresolved genotypes. Populations 8 and 10 were syntopic sites without evidence of admixture. Finally, 3 hybrid individuals (8%) were found in population 7, while 14 individuals (39%) were unresolved.

## DISCUSSION

An earlier study (L16) quantified hybridization between *P. cinereus* and *P. electromorphus* and found that hybridization was pervasive and widespread across northern Ohio. Given such high levels of interbreeding, the authors questioned how the two species could maintain their separation. Could they be merging into a single entity? If so, it would be a remarkable example of reticulate evolution given that the two species diverged on the order of 10 myr ago (Wiens et al., 2006).

In this paper we reanalyzed the data in L16, as the results contradict our field observations and also impact our interpretation of past and ongoing research. We had two alternative hypotheses to explain why levels of hybridization

might have been overestimated: that high levels of linkage disequilibrium between sets of SNPs created a false signal of admixture (Willis et al., 2017; O'Leary et al., 2018), or that the selection of five populations instead of 2 (K = 5 or 2) in STRUCTURE misled the analyses.

For both the linked and unlinked data, we obtained support for $\Delta K = 2$ (Supplemental Fig. 4; see Data Accessibility). This result is intuitive given that two species are being compared and ΔK is known to identify the basal level of hierarchical population structure (Janes et al., 2017). For both datasets, the estimated log probability of the data, $\Pr(X|K)$, was similar for K = 2–4. Tests for substructure within either species did not provide any results that would inform this paper, but for detailed analyses of genetic variation in this region see Highton (1999), Cameron et al. (2019), Hantak et al. (2019), Waldron et al. (2019), and Radomski et al. (2020). In contrast with our results, L16 reported finding support for K = 5 using ΔK and $\Pr(X|K)$, and this is supported by a saved screenshot of STRUCTURE HARVESTER output (Earl and vonHoldt, 2012). We ran a number of analyses in STRUCTURE with diverse priors and were never able to replicate the K = 5 result.

Our analyses in STRUCTURE using the linked (L16) and unlinked (this paper) datasets for K = 2 gave very similar results (Fig. 2), indicating that linkage disequilibrium did not mislead STRUCTURE. Both analyses identified several populations that were pure *P. cinereus* or *P. electromorphus*. In addition, syntopic populations without hybrids were identified, or only a single hybrid individual was recovered. Only population 7 was highly admixed, yet even in this population both parental types were also present. When the unlinked data were analyzed using DNA haplotypes at each locus in NEWHYBRIDS, or when genotypes were reconstructed from gene trees, similar lower levels of hybridization were inferred (Fig. 3). No clear F1 hybrids were recovered in any population by these analyses.

In contrast with K = 2, when we used K = 5 (as in L16) on both the linked and unlinked data, high levels of admixture were inferred (Supplemental Fig. 5; see Data Accessibility). L16 postulated that K = 5 corresponded with pure *P. cinereus*, pure *P. electromorphus*, F1 hybrids, backcrosses to *P. cinereus*, and backcrosses to *P. electromorphus*. This represents a dubitable interpretation of the data as STRUCTURE assumes each individual has ancestry from one or more of K genetically distinct sources. For example, the expectation is that an F1 hybrid should be identified as including roughly equal membership to two groups, rather than being recovered as a third group.

In summary, our analyses indicate that the high frequency of hybridization reported in L16 is a consequence of an inappropriate selection of K and not linkage disequilibrium. When K is correctly inferred, substantially lower levels of admixture are recovered.

Indeed, the lower levels of admixture we report in this paper likely overestimate the amount of hybridization in natural populations. This is because the sampling of individuals by L16 was not random. Rather, an effort was made to find individuals of each species as well as individuals of intermediate phenotype. This maximizes the chance of detecting hybridization, which is a reasonable starting point for such a study; however, one cannot use this data to infer the frequency of hybridization within natural populations. The exception is population 7, where most individuals were randomly sampled. In that population, L16 reported 50% hybrids. In our analyses, population 7 included between 8% and 31% hybrids, depending on the analysis (Fig. 2, 3; Supplemental Table 1; see Data Accessibility). Highton's (1999) analyses of this same population, which included "over 70 specimens" of both parental types and ten hybrid individuals, recovered a maximum of 10/150 = 7% hybrids.

In conclusion, though sample sizes are limited at many sites, it appears from our analyses as well as Highton (1999) that *P. electromorphus* and *P. cinereus* are largely reproductively isolated, with occasional hybridization. The exception is population 7, where levels of hybridization are relatively high. The cause of this shift in reproductive isolation is unknown and presents a fruitful avenue for future research.

***Study limitations.***—While analyzing 20 SNPs as three loci avoids the problems associated with linkage disequilibrium (Willis et al., 2017; O'Leary et al., 2018), studying hybridization with three nuclear loci is statistically dubious. For example, NEWHYBRIDS only detected three hybrid genotypes with posterior probability > 0.90 (all in population 7); no other individual in any other population was classified as a hybrid with confidence. On the other hand, the program could not assign 28 individuals to any category (parental or hybrid), including some individuals in populations that STRUCTURE and the gene tree approach inferred to be composed entirely of *P. cinereus* or *P. electromorphus*. These results highlight a fundamental deficiency of the study: there are not enough loci to address levels of admixture with statistical confidence. Even 20 independent SNPs is insufficient for many scenarios. More data are needed, and the analyses in this paper should be interpreted with caution.

A confounding factor in both studies is separating hybridization from introgression. Hybridization is the crossing of genetically distinguishable groups that produces individuals with admixed genotypes, such as F1s and backcrosses. Introgression occurs when one entity incorporates alleles from another entity into its genome. Introgression is thus a relative term: alleles at one locus introgress with respect to alleles at other loci. The process of introgression is initiated by hybridization, but such hybridization may have occurred in the distant past. In animals, the introgression of mtDNA has been best studied (Weisrock et al., 2005; Kuchta and Tan, 2006; McGuire et al., 2007), but alleles at nuclear loci can also introgress (Edelman et al., 2019). Our method of examining gene trees to infer genotypes may overestimate hybridization because it cannot distinguish between hybridization and introgression. For instance, if RAG1 alleles introgressed from *P. cinereus* into the genome of *P. electromorphus* thousands of years ago, our method would score this as evidence of hybridization. Indeed, within our RAG1 gene tree, one subclade within the *P. cinereus* clade appears to be composed entirely of individuals of *P. electromorphus* (Supplemental Fig. 3; see Data Accessibility). This could reflect historical introgression instead of hybridization; however, there are other viable explanations (e.g., phylogenetic error) and more data are needed to distinguish among them.

***The importance of hybridization.***—Why levels of hybridization between *P. electromorphus* and *P. cinereus* at Wooster Memorial Park (population 7) are more extensive than elsewhere is unclear. More importantly, is hybridization rampant across

northern Ohio? This is crucial to understand because the amount of hybridization between *P. cinereus* and *P. electromorphus* impacts the interpretation of many past and ongoing studies. Multiple color morphs of *P. cinereus* exist, one of which—the unstriped morph, which like *P. electromorphus* has a black dorsum—is tricky to distinguish from *P. electromorphus*. In northern Ohio, *P. cinereus* has long been studied as an exemplar of color polymorphism, including studies of diet (Stuczka et al., 2016; Hantak et al., 2020), mating (Anthony et al., 2008; Acord et al., 2013; Jaworski et al., 2018), territoriality (Reiter et al., 2014; Anthony et al., 2017), predation (Venesky and Anthony, 2007; Hantak and Kuchta, 2018), and many other axes of differentiation. If unstriped salamanders were often genetically admixed individuals, alternative explanations for the results of this work on polymorphism would need to be invoked.

Fortunately, our reanalysis of the data collected by L16, as well as Highton (1999), indicate that hybridization between *P. cinereus* and *P. electromorphus* is not pervasive or geographically widespread. That said, our understanding of the amount of hybridization and introgression remains incomplete. Future work should take advantage of next generation sequencing technologies to obtain many loci (hundreds or thousands, not three) to quantify the reproductive interactions between *P. cinereus* and *P. electromorphus*. Ideally, such a study should include populations where the two species are syntopic, as well as regions where the two species present a distributional patchwork with contacts along range edges, as both patterns are found in Ohio (Pfingsten, 2013).

## DATA ACCESSIBILITY

Supplemental material is available at https://www.ichthyologyandherpetology.org/h2021081. Unless an alternative copyright or statement noting that a figure is reprinted from a previous source is noted in a figure caption, the published images and illustrations in this article are licensed by the American Society of Ichthyologists and Herpetologists for use if the use includes a citation to the original source (American Society of Ichthyologists and Herpetologists, the DOI of the *Ichthyology & Herpetology* article, and any individual image credits listed in the figure caption) in accordance with the Creative Commons Attribution CC BY License.

## ACKNOWLEDGMENTS

We would like to think E. Watts-Whitehead for comments on the manuscript.

## LITERATURE CITED

**Acord, M. A., C. D. Anthony, and C.-A. M. Hickerson.** 2013. Assortative mating in a polymorphic salamander. Copeia 2013:676–683.

**Aktas, C.** 2020. haplotypes: manipulating DNA sequences and estimating unambiguous haplotype networks with statistical parsimony. R package version 1.1.2. https://CRAN.R-project.org/package=haplotypes

**Anderson, E. C., and E. A. Thompson.** 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229.

**Anthony, C., C.-A. M. Hickerson, K. Jaworski, and M. Messner.** 2017. Differences in prey availability within the territories of striped and unstriped Eastern Red-backed Salamanders (*Plethodon cinereus*). Herpetological Review 48: 509–514.

**Anthony, C. D., M. D. Venesky, and C.-A. M. Hickerson.** 2008. Ecological separation in a polymorphic terrestrial salamander. Journal of Animal Ecology 77:646–653.

**Barton, N. H., and G. M. Hewitt.** 1989. Adaptation, speciation and hybrid zones. Nature 341:497–503.

**Cameron, A. C., R. B. Page, J. I. Watling, C.-A. M. Hickerson, and C. D. Anthony.** 2019. Using a comparative approach to investigate the relationship between landscape and genetic connectivity among woodland salamander populations. Conservation Genetics 20:1265–1280.

**Converse, P. E., S. R. Kuchta, W. M. Roosenburg, P. F. P. Henry, G. M. Haramis, and T. L. King.** 2015. Spatiotemporal analysis of gene flow in Chesapeake Bay Diamondback Terrapins (*Malaclemys terrapin*). Molecular Ecology 24: 5864–5876.

**Darriba, D., G. L. Taboada, R. Doallo, and D. Posada.** 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772.

**Deitloff, J.** 2011. Sympatry between two wide-ranging salamander species. IRCF Reptiles & Amphibians 18:7–11.

**Drummond, A. J., and R. R. Bouckaert.** 2015. Bayesian Evolutionary Analysis with BEAST. Cambridge University Press, Cambridge.

**Earl, D. A., and B. M. vonHoldt.** 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361.

**Edelman, N. B., P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-Accinelli, S. M. Van Belleghem, N. Patterson, D. E. Neafsey, R. Challis, S. Kumar, G. R. P. Moreira, C. Salazar . . . J. Mallet.** 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366:594–599.

**Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.

**Evanno, G., S. Regnaut, and J. Goudet.** 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

**Fisher-Reid, M. C., and J. J. Wiens.** 2011. What are the consequences of combining nuclear and mitochondrial data for phylogenetic analysis? Lessons from *Plethodon* salamanders and 13 other vertebrate clades. BMC Evolutionary Biology 11:300.

**Francis, R. M.** 2017. POPHELPER: an R package and web app to analyse and visualize population structure. Molecular Ecology Resources 17:27–32.

**Hantak, M. M., K. M. Brooks, C.-A. M. Hickerson, C. D. Anthony, and S. R. Kuchta.** 2020. A spatiotemporal assessment of dietary partitioning between color morphs of a terrestrial salamander. Copeia 108:727–736.

**Hantak, M. M., and S. R. Kuchta.** 2018. Predator perception across space and time: relative camouflage in a colour polymorphic salamander. Biological Journal of the Linnean Society 123:21–33.

**Hantak, M. M., R. B. Page, P. E. Converse, C. D. Anthony, C.-A. M. Hickerson, and S. R. Kuchta.** 2019. Do genetic structure and landscape heterogeneity impact color morph

frequency in a polymorphic salamander? Ecography 42: 1383–1394.

**Highton, R.** 1995. Speciation in eastern North American salamanders of the genus *Plethodon*. Annual Review of Ecology, Evolution, and Systematics 26:579–600.

**Highton, R.** 1999. Geographic protein variation and speciation in the salamanders of the *Plethodon cinereus* group with the description of two new species. Herpetologica 55: 43–90.

**Highton, R., A. P. Hastings, C. Palmer, R. Watts, C. A. Hass, M. Culver, and S. J. Arnold.** 2012. Concurrent speciation in the eastern woodland salamanders (genus *Plethodon*): DNA sequences of the complete albumin nuclear and partial mitochondrial 12s genes. Molecular Phylogenetics and Evolution 63:278–290.

**Hohenlohe, P. A., S. Bassham, M. Currey, and W. A. Cresko.** 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. Philosophical Transactions of the Royal Society B: Biological Sciences 367:395–408.

**Janes, J. K., J. M. Miller, J. R. Dupuis, R. M. Malenfant, J. C. Gorrell, C. I. Cullingham, and R. L. Andrew.** 2017. The K = 2 conundrum. Molecular Ecology 26:3594–3602.

**Jaworski, K. E., M. S. Lattanzio, C.-A. M. Hickerson, and C. D. Anthony.** 2018. Male mate preference as an agent of fecundity selection in a polymorphic salamander. Ecology and Evolution 8:8750–8760.

**Jockusch, E., and D. B. Wake.** 2002. Falling apart and merging: diversification of slender salamanders (Plethodontidae: *Batrachoseps*) in the American West. Biological Journal of the Linnean Society 76:361–391.

**Kleindorfer, S., J. A. O'Connor, R. Y. Dudaniec, S. A. Myers, J. Robertson, and F. J. Sulloway.** 2014. Species collapse via hybridization in Darwin's tree finches. American Naturalist 183:325–341.

**Kozak, K. H., D. W. Weisrock, and A. Larson.** 2006. Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). Proceedings of the Royal Society B: Biological Sciences 273:539–546.

**Kuchta, S. R.** 2007. Contact zones and species limits: hybridization between lineages of the California Newt, *Taricha torosa*, in the southern Sierra Nevada. Herpetologica 63:332–350.

**Kuchta, S. R., A. D. Brown, and R. Highton.** 2018. Disintegrating over space and time: paraphyly and species delimitation in the Wehrle's Salamander complex. Zoologica Scripta 47:285–299.

**Kuchta, S. R., and A. Tan.** 2006. Lineage diversification on an evolving landscape: phylogeography of the California newt, *Taricha torosa* (Caudata: Salamandridae). Biological Journal of the Linnean Society 89:213–239.

**Lanfear, R., B. Calcott, S. Y. W. Ho, and S. Guindon.** 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29:1695–1701.

**Lehtinen, R. M., A. F. Steratore, M. M. Eyre, E. S. Cassagnol, M. L. Stern, and H. A. Edgington.** 2016. Identification of widespread hybridization between two terrestrial salamanders using morphology, coloration, and molecular markers. Copeia 104:132–139.

**Mallet, J.** 2005. Hybridization as an invasion of the genome. Trends in Ecology & Evolution 20:229–237.

**McGuire, J. A., C. W. Linkem, M. S. Koo, D. W. Hutchison, A. K. Lappin, D. I. Orange, J. Lemos-Espinal, B. R. Riddle, and J. R. Jaeger.** 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution 61: 2879–2897.

**Milne, I., D. Lindner, M. Bayer, D. Husmeier, G. McGuire, D. F. Marshall, and F. Wright.** 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25:126–127.

**O'Leary, S. J., J. B. Puritz, S. C. Willis, C. M. Hollenbeck, and D. S. Portnoy.** 2018. These aren't the loci you're looking for: principles of effective SNP filtering for molecular ecologists. Molecular Ecology 27:3193–3206.

**Pereira, R. J., and D. B. Wake.** 2009. Genetic leakage after adaptive and nonadaptive divergence in the *Ensatina eschscholtzii* ring species. Evolution 63:2288–2301.

**Pfingsten, R. A.** 2013. Northern Ravine Salamander, *Plethodon electromorphus*, p. 361–368. *In*: Amphibians of Ohio. R. A. Pfingsten, J. G. Davis, T. O. Matson, G. J. Lipps, D. Wynn, and B. J. Armitage (eds.). Ohio Biological Survey, Columbus, Ohio.

**Pritchard, J. K., M. Stephens, and P. Donnelly.** 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

**Pritchard, J. K., X. Wen, and D. Falush.** 2010. Documentation for STRUCTURE software: version 2.3.

**Radomski, T., M. M. Hantak, A. D. Brown, and S. R. Kuchta.** 2020. Multilocus phylogeography of Eastern Red-backed Salamanders (*Plethodon cinereus*): cryptic Appalachian diversity and postglacial range expansion. Herpetologica 76:61–73.

**Reiter, M. K., C. D. Anthony, and C.-A. M. Hickerson.** 2014. Territorial behavior and ecological divergence in a polymorphic salamander. Copeia 2014:481–488.

**Seehausen, O.** 2006. Conservation: losing biodiversity by reverse speciation. Current Biology 16:R334–R337.

**Sites, J. W., Jr., M. Morando, R. Highton, F. Huber, and R. E. Jung.** 2004. Phylogenetic relationships of the endangered Shenandoah Salamander (*Plethodon shenandoah*) and other salamanders of the *Plethodon cinereus* group (Caudata: Plethodontidae). Journal of Herpetology 38:96–105.

**Stephens, M., N. J. Smith, and P. Donnelly.** 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68:978–989.

**Stuczka, A., C.-A. Hickerson, and C. Anthony.** 2016. Niche partitioning along the diet axis in a colour polymorphic population of Eastern Red-backed Salamanders, *Plethodon cinereus*. Amphibia-Reptilia 37:283–290.

**Venesky, M. D., and C. D. Anthony.** 2007. Antipredator adaptations and predator avoidance by two color morphs of the Eastern Red-backed Salamander, *Plethodon cinereus*. Herpetologica 63:450–458.

**Waldron, B. P., and M. H. Hantak.** 2020. *Plethodon cinereus* (Eastern Red-backed Salamander) and *Plethodon electromorphus* (Northern Ravine Salamander). Hybridization. Herpetological Review 51:809–810.

**Waldron, B. P., S. R. Kuchta, M. M. Hantak, C.-A. M. Hickerson, and C. D. Anthony.** 2019. Genetic analysis of

a cryptic contact zone between mitochondrial clades of the Eastern Red-backed Salamander, *Plethodon cinereus*. Journal of Herpetology 53:144–153.

**Wang, J.** 2017. The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. Molecular Ecology Resources 38:42–49.

**Weisrock, D. W., K. H. Kozak, and A. Larson.** 2005. Phylogeographic analysis of mitochondrial gene flow and introgression in the salamander, *Plethodon shermani*. Molecular Ecology 14:1457–1472.

**Wiens, J. J., T. N. Engstrom, and P. T. Chippindale.** 2006. Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (genus *Plethodon*): testing the hybrid swarm hypothesis of rapid radiation. Evolution 60:2585–2603.

**Willis, S. C., C. M. Hollenbeck, J. B. Puritz, J. R. Gold, and D. S. Portnoy.** 2017. Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Molecular Ecology Resources 17:955–965.