# Genomic Structure Among Populations of a Regionally Rare Perennial Plant Indicates the Need for Reevaluation of Management Units

Authors: Moore, Cheyenne L., McDonnell, Angela J., Schuette, Scott, and Martine, Christopher T.

Source: Natural Areas Journal, 45(2) : 124-136

Published By: Natural Areas Association

URL: https://doi.org/10.3375/2162-4399-45.2.8

*Research Article*

# Genomic Structure Among Populations of a Regionally Rare Perennial Plant Indicates the Need for Reevaluation of Management Units

**Cheyenne L. Moore,[1,2,5] Angela J. McDonnell,[1,3] Scott Schuette,[4] and Christopher T. Martine[1]**

[1]Bucknell University, Lewisburg, PA
[2]Pennsylvania Department of Conservation and Natural Resources, Harrisburg, PA (current)
[3]St. Cloud State University, St. Cloud, MN (current)
[4]Western Pennsylvania Conservancy, Pittsburgh, PA

[5]Corresponding author: cheyennelmoore17@gmail.com

Associate Editor: Betsy Miller Vixie

## ABSTRACT

Conservation of rare species often relies on the delineation of management units and genomic tools can now be applied for this purpose. However, this is not as common on a local scale, where populations are often small and fragmented, despite the utility of informing conservation and management practice at the regional level. We use a genotyping by sequencing (GBS) approach to assess the perennial riparian plant species *Baptisia australis* (Fabaceae) in Pennsylvania, where the taxon is at the edge of its natural distribution and considered threatened. In this system, we investigate whether sampled subpopulations exhibit genetic structure. We find that genetic structure largely aligns with the waterways that were sampled, with five genetic clusters, one for each of three waterways and two additional clusters along a fourth waterway. This work directly addresses conservation priorities in the state and informs delineation of management units that will be used in conservation practices.

*Index terms:* conservation; management units; population genomics; riparian

## INTRODUCTION

On the local scale, small populations of species of concern are often identified as conservation priorities (e.g., Ellstrand and Elam 1993), yet genetic data are infrequently available to aid in management decisions and, ultimately, the prevention of local extirpation through interventions such as conservation horticulture. At the same time, the increasing adoption of next generation sequencing (NGS) data for population studies has translated into an increased interest in applying these tools to questions related to conservation biology (Delgado et al. 2008; Allendorf et al. 2010; Frankham 2010; Noss et al. 2021). The use of genomic data to help delineate conservation and management units is well established (Allendorf and Luikart 2007; Palsbøll et al. 2007; Funk et al. 2012; Coates et al. 2018). Genetic tools have been used to aid in conservation efforts including establishment of management units (Alpers et al. 2004) and examination of population structure and gene flow in species of conservation concern (Bowen and Karl 2007). Expanded access to and greater affordability of high-throughput sequencing now offers researchers the ability to approach these same questions with increased amounts of sampling and genetic data as demonstrated in McLennan et al. (2019), Wright et al. (2020), McDonnell et al. (2021), and Hayes (2021). Reduced representation sequencing methods (e.g., restriction enzyme based approaches such as RADseq [Rowe et al. 2011] or genotyping by sequencing [GBS; Elshire et al. 2011]) and targeted enrichment (Hale et al. 2020) are valuable for population genomic study of non-model organisms because the data recovered are more representative of the entire genome than traditional approaches such as microsatellites, RFLPs, and AFLPs (Rowe et al. 2011; Seeb et al. 2011; Peterson et al. 2012; Andrews et al. 2016).

Two primary concerns in conservation genetics are population delineation for management targets (or assigning management units) and assessment of the genetic health of those populations. Funk et al. (2012) define a "management unit" (MU) as "a local population that is managed as a distinct unit because of its demographic independence" and conclude that "maintaining multiple MUs is important for ensuring long-term persistence of species." As per Frankham (2010), "Lack of genetic management in [wild] populations is not due to a lack of scientific guidelines, but due to failure to consider genetic issues in wild management." An especially important issue in genetic management of wild populations is addressing the relationship between genetic and spatial population structure. Understanding spatial structure and metapopulation composition can be integral to establishment of effective conservation plans and identification of management units (De Campos Telles et al. 2003; Lee et al. 2006; Delgado et al. 2008; Cipollini et al. 2017). In this paper we implement a GBS approach to address these concerns as they relate to a globally rare plant taxon.

We focus on the regional conservation of *Baptisia australis* (L.) R. Br. or blue wild indigo, a disturbance-dependent perennial legume that is native to riverbank scour areas and gravel banks occurring historically as a native species in parts of the eastern United States, including Virginia, Maryland, West Virginia, Pennsylvania, Ohio, Indiana, Kentucky, and Tennessee (Broyles 2006; Weakley 2024). It is considered rare in all of the states throughout its distribution and is at moderate risk of decline due to a fairly restricted range and habitat requirements. Following the NatureServe global conservation status definitions, this species is ranked as vulnerable (G3; NatureServe 2021). In Pennsylvania, the Department of Conservation and Natural Resources (PA DCNR) recently ranked the taxon as threatened (2017), and the Natural Heritage Program (PNHP) considers it imperiled (S2) (NatureServe 2021). Pennsylvanian populations of this taxon are found along four waterways in the western half of the state, within the Allegheny and Monongahela River watersheds. This study was designed with the purpose of aiding conservation agencies such as PA DCNR and PNHP in future management plans of this state-imperiled species, and applies molecular data in a way that has so far rarely been accomplished in the region for plants (Cipollini et al 2017; McDonnell et al. 2021) but has been successfully used elsewhere (Schilling et al. 2014; Resh et al. 2018; Silliman 2019). Such work often finds that when plants are at the edge of their range they exhibit lower genetic diversity and high rates of inbreeding compared to plants at the center of the range.

Pennsylvania populations of *B. australis* are relatively isolated from populations outside of the state, making them a relatively "closed system" for studying gene flow in riparian systems and understanding such gene flow in this system may also aid in delineating other riparian plant metapopulations in the state. In addition to being isolated, Pennsylvania populations of *B. australis* are at the edge of the taxon's distribution. Edge-of-range populations are of particular interest because of their potential contribution to the genetic diversity and integrity to more widely distributed taxa, for example private alleles found in edge populations (Diekmann and Serrão 2012; Assis et al. 2013). A better understanding of the genetic diversity, inbreeding, and connectivity of edge-of-range populations can be used to infer levels of these genetic traits in central populations, which are anticipated to have greater genetic diversity than edge populations susceptible to bottlenecks or genetic drift (Sexton et al. 2009, 2011).

We adopt a GBS approach to obtain single nucleotide polymorphisms (SNPs) from multiple individuals from all populations of *B. australis* in Pennsylvania. We employ the resultant SNP dataset to estimate population genetic statistics including $F_{ST}$, $F_{IS}$, and heterozygosity and also conduct clustering analyses to infer group membership of individuals. We test two hypotheses based on knowledge of *B. australis* populations in Pennsylvania: (1) the sampled subpopulations exhibit genetic structure with the most geographically disjunct population (Youghiogheny River) being the most genetically distinct population due to its isolation within Pennsylvania, and (2) genomic data will align with that of range edge populations—high differentiation of populations, low diversity, and high inbreeding.

## MATERIALS AND METHODS

### Study Species

*Baptisia* Vent. is a genus endemic to North America composed of 21 species, 2 infraspecific taxa, and 7 named hybrids (NatureServe 2021; Weakley 2024) that range from eastern Texas, north to Minnesota into Canada, east to southern Maine through the southeastern United States, and throughout the Midwest (Kartesz 2003; Woods and Diamond 2014). *Baptisia australis* has a scattered range from northeastern Illinois through Indiana and Ohio to western Pennsylvania, then south to northwestern Georgia. This species primarily occurs in riverbank scours, cobble bars, and other disturbed areas generally associated with midwestern prairie grasses (e.g., *Andropogon gerardii* Vitman., *Sorghastrum nutans* (L.) Nash., *Spartina pectinata* Link.). In Pennsylvania, the habitat is described as riverscour wet meadow (Zimmerman et al. 2012).

### Sampling and Sequencing

We sampled 22 of 25 known populations of *Baptisia australis* in Pennsylvania in an effort to determine whether reservoirs of diversity exist in the state (Roberts et al. 2007; Frankham et al. 2017). Sampling was conducted at populations monitored by the PNHP along 4 waterways in western Pennsylvania in the summers of 2018 and 2019 and included 15 sites along the Allegheny River, two sites along the Clarion River, three sites along the Red Bank Creek, and the only known population along the Youghiogheny River (Figure 1). These sites are well documented, naturally occurring populations at which ~15 plants were sampled. Individuals were $\geq$1 m apart and representative phenotypes were targeted. Additionally, cultivated plants grown from seed harvested along West Virginia's Greenbrier River as well as those grown from seed harvested from a previously unsampled Allegheny River population were also sampled and included in this study (provided by Ernst Conservation Seeds, Meadville, Pennsylvania, USA). In total, 326 samples representing 24 populations were included (Table 1). All populations sampled occur along waterways that are in the Ohio River drainage; the Clarion River and Red Bank Creek drain into the Allegheny River and the Youghiogheny River drains into the Monongahela River. Both the Allegheny and Monongahela Rivers confluence to form the Ohio River at Pittsburgh in western Pennsylvania (Figure 1).

Genomic DNA was extracted from silica-dried tissue using the FastDNA kit following manufacturer instructions (MP Biomedicals, Santa Ana, California, USA) and was quantified using a Qubit dsDNA BR assay kit on a Qubit v2.0 fluorometer (ThermoFisher Scientific, Waltham, Massachusetts, USA). To ensure the quality of DNA, 2–3 µl of DNA was visualized on 1% agarose gels run at 100V for 1.5 hr. DNA and

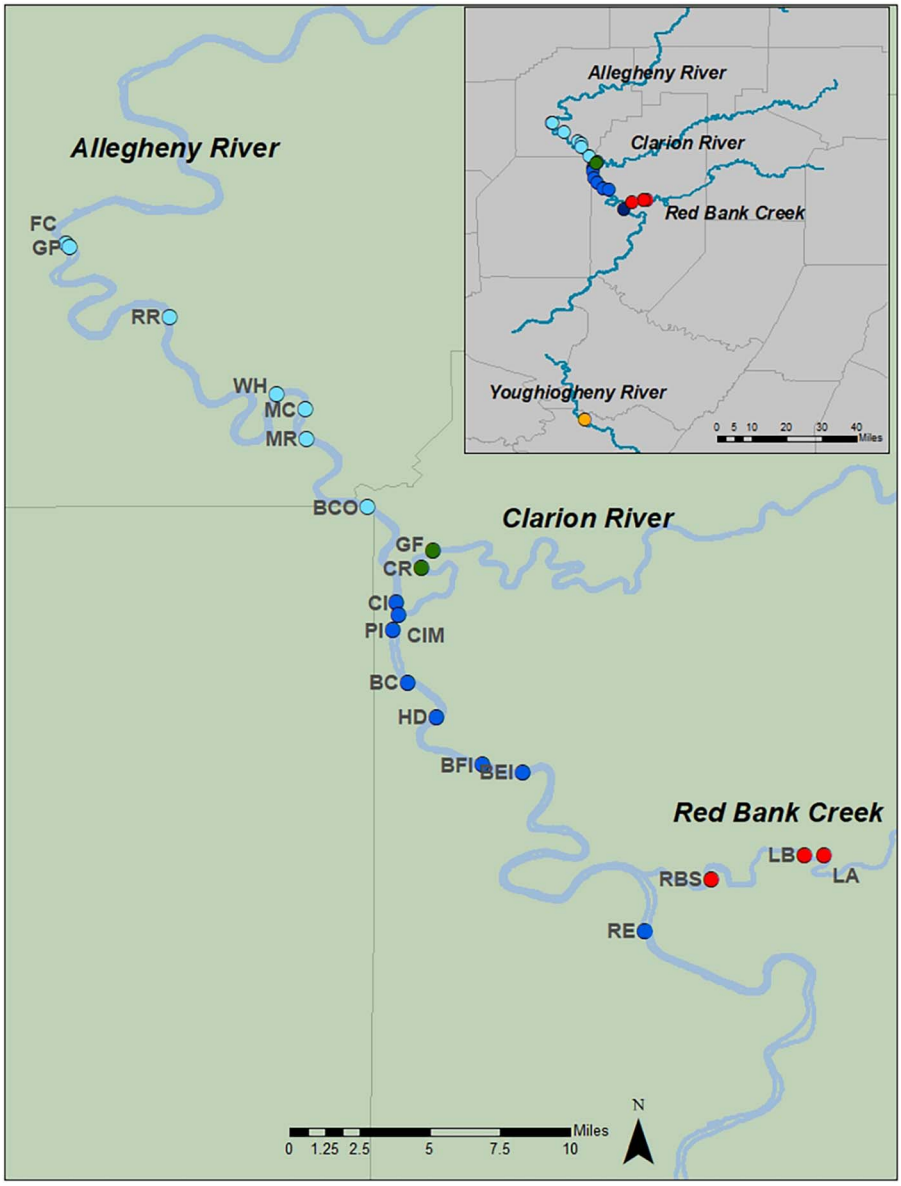


**Figure 1.**—Map of all Pennsylvania sites where tissue was collected. Excludes Ernst Seeds sites where *Baptisia* is in cultivation. Map made with ArcMap (ESRI 2011).

corresponding gel images were sent to the University of Wisconsin Biotechnology Center lab (https://www.biotech.wisc.edu/services/dnaseq) where additional enzyme testing as well as library preparation and sequencing were conducted. Fragment analyses revealed that the restriction enzyme *Ape*K1 showed the greatest activity in our samples. Therefore, a one-enzyme, genotyping by sequencing approach was used (Elshire et al. 2011). DNA was digested using *Ape*K1, libraries were prepared, quantified, pooled, and 150bp paired-end sequenced on a NovaSeq 6000 instrument (Illumina, San Diego, California, USA). All raw sequencing data are available online via the National Center for Biotechnology Information (NCBI) website, BioProject number PRJNA574847. The Sequence Read Archive (SRA) is available at https://www.ncbi.nlm.nih.gov/sra/PRJNA574847.

**Sequence Processing**

The Python program iPyrad 07.30 was used to process raw reads (Eaton 2014). Default parameters were used to demultiplex reads by sample, filter adapters, merge reads from separate sequencing runs, and cluster into de novo loci. Low-quality bases were removed and reads were subsequently clustered within samples into de novo loci using an 88% similarity threshold. Samples with fewer than six or more than 10,000 reads were removed. Loci were then clustered across samples at an 88% similarity threshold and processed into an assembly where all loci were shared among four or more samples. A total of 96,489 loci were recovered. The resulting Variant Call Format (.vcf) file was further filtered using VCFtools 0.1.16 (Danecek et al. 2011) with the following flags: –maf 0.01, –max-alleles 2, –min-alleles 2, –minDP 3,

**Table 1.**—Table of sampled sites and information about their location and sampling. Latitude and longitude is redacted due to PNHP imperiled status of taxon.

| Waterway | Site name | Abbrev. | Collector | Collection numbers | Tissue collection date | Survey date | Plants sampled | County |
|---|---|---|---|---|---|---|---|---|
| Allegheny River | Fisherman's Cove | FC | C.L. Moore | 69 | 7/8/2019 | 7/8/2019 | 5 | Venango |
| Allegheny River | Gas Pipeline | GP | C.L. Moore | 70 | 7/8/2019 | 7/8/2019 | 10 | Venango |
| Allegheny River | Robert's Run | RR | C.L. Moore | 85 | 7/11/2019 | 7/11/2019 | 15 | Venango |
| Allegheny River | Wood Hill | WH | C.L. Moore | 1 | 7/15/2018 | 6/11/18 7/15/18 | 15 | Venango |
| Allegheny River | Mill Creek | MC | C.L. Moore | 9 | 7/13/2018 | 6/12/18 7/13/18 | 15 | Venango |
| Allegheny River | Meadowsweet Run | MR | C.L. Moore | 14 | 7/13/2018 | 6/12/18 7/13/18 | 15 | Venango |
| Allegheny River | Butler County | BCO | C.L. Moore | 21 | 7/14/2018 | 6/13/18 7/14/18 | 15 | Butler |
| Allegheny River | Clarion Island | CI | C.L. Moore | 26 | 7/14/2018 | 6/13/18 7/14/18 | 15 | Clarion |
| Allegheny River | Clarion Island Mix | CIM | C.L. Moore | 31 | 7/14/2018 | 6/13/18 7/14/18 | 15 | Clarion |
| Allegheny River | Parker Island | PI | C.L. Moore | 71 | 7/9/2019 | 7/9/2019 | 15 | Clarion |
| Allegheny River | Bear Creek | BC | C.L. Moore | 67 | 9/6/2018, 7/9/2019 | 9/6/2018, 7/9/2019 | 25 | Clarion |
| Allegheny River | Heck Drive | HD | C.L. Moore | 77 | 7/9/2019 | 7/9/2019 | 10 | Clarion |
| Allegheny River | Black Fox Island | BFI | C.L. Moore | 79 | 7/10/2019 | 7/10/2019 | 15 | Clarion |
| Allegheny River | Bald Eagle Island | BEI | C.L. Moore | 83 | 7/10/2019 | 7/10/2019 | 15 | Clarion |
| Allegheny River | Ernst | EPA | C.L. Moore | 88 | 7/11/2019 | 7/11/2019 | 10 | Crawford |
| Allegheny River | River's Edge | RE | C.L. Moore | 89 | 7/12/2019 | 7/12/2019 | 15 | Armstrong |
| Clarion River | Grassy Flats | GF | S. Schuette | 2204 | 8/9/2018 | 8/9/2018 | 15 | Clarion |
| Clarion River | Clarion River | CR | C.L. Moore | 60–66,68 | 9/5/2018 | 9/5/2018 | 15 | Clarion |
| Red Bank Creek | Lawsonham A | LA | S. Schuette | 2245 | 8/28/2018 | 8/28/2018 | 15 | Clarion |
| Red Bank Creek | Lawsonham B | LB | S. Schuette | 2246 | 8/28/2018 | 8/28/2018 | 15 | Clarion |
| Red Bank Creek | Red Bank Station | RBS | C.L. Moore | 58 | 9/4/2018 | 9/4/2018 | 20 | Clarion |
| Youghiogheny River | Layton | YR | C.L. Moore | 92 | 7/26/2019 | 7/4/2018 | 11 | Fayette |
| Greenbrier River | Ernst | EWV | C.L. Moore | 87 | 7/11/2019 | 7/11/2019 | 10 | Crawford |
| Greenbrier River | Ernst Cultivated @ BU | EWVC | C.L. Moore | 93 | 7/25/2019 | n/a | 5 | Union |

–min-meanDP 3, –max-missing 0.4, –out all_B_filtered, and –remove-indels, returning biallelic SNPs without indels and having a minor allele frequency greater than 0.01, sites with mean depth greater than 3. Our final filtered dataset used for analyses contains 317 individuals from 24 populations and 11,323 SNPs.

**Genetic Diversity**

All analyses were conducted in the R Studio environment (R 3.6.2; R Core Team 2019). The R package *vcfR* (Knaus and Grünwald 2017) was used to convert the filtered .vcf file to genlight and genind objects, which were used as inputs for the majority of downstream analyses. The package *hierfstat* (Goudet 2005) was used to calculate population genetic parameters, fixation index ($F_{IS}$), observed heterozygosity and within-population gene diversity ($H_o$, $H_s$), as well as differentiation index ($F_{ST}$) using the Weir and Cockerham (1984) method with 100 bootstraps.

**Population Structureh Ahmet**

The R package *adegenet* (Jombart 2008; Jombart and Ahmed 2011) was used to complete a principal component analysis (PCA) of all samples using glPC. "adegenet" was also used to conduct a discriminant analysis of principal components (DAPC), which functions similarly to a PCA but also utilized discriminant analysis (DA) to provide membership probabilities (Jombart et al. 2010). SNP data were analyzed to estimate the population structure of *Baptisia australis* in Pennsylvania through a series of principal PCAs. These PCAs are retained to complete a discriminant analysis on all PCAs with the optimum *K*, or number of clusters, identified using the Bayesian information criterion and *K*-means clustering. The *Landscape and Ecological Association* or *LEA* R package (Frichot et al. 2014; Frichot and François 2015) was used to assess population structure and admixture. *K* was estimated with the cross-entropy criterion; 1000 iterations were run 10 times for *K* values of 1–24. *LEA* was then used to create a sNMF plot.

The R packages *dartR* and *adegenet* were used to perform a Mantel test to determine whether a pattern of isolation by distance (IBD) is present. All individuals except for those grown in cultivation (Table 1: EPA, EWV, EWVC) were included. A second Mantel test included only Allegheny River populations. For each test, 999 permutations were summarized. Finally, we conducted an analysis of molecular variance (AMOVA) using the *poppr* R package (Kamvar et al. 2014, 2015) to determine whether there is more genetic variation within or between sampled populations.

## RESULTS

**GBS Data**

Sequencing yielded a total of 794.3 million high-quality paired sequences. After demultiplexing, assembly, and filtering, 11,323 SNPs from 317 individuals collected from 24 populations were used for analyses.

**Genetic Diversity**

Between most populations, moderate differentiation was detected (global $F_{ST} = 0.185$, 95% confidence interval; Figure 2)

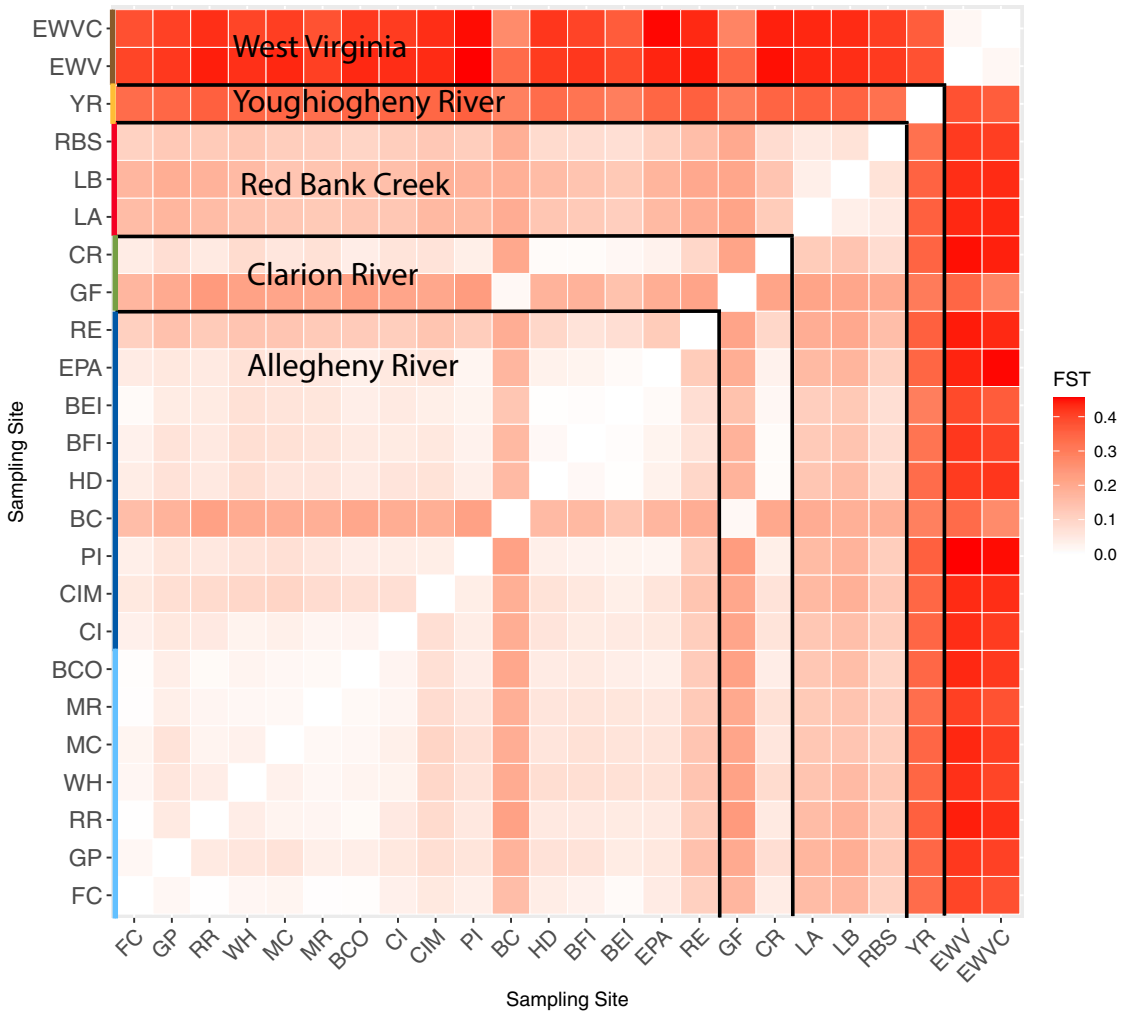


**Figure 2.**—Heatmap of pairwise $F_{ST}$ values (Weir and Cockerham 1984). Site abbreviations correspond to Table 1.

as well as high rates of inbreeding (global $F_{IS} = 0.173$ with a range of 0.0292 to 0.2201; Aguirre-Dugua et al. 2023; Wu et al. 2023), with particularly high inbreeding along the Clarion and Youghiogheny Rivers and the Red Bank Creek (Table 2). Within population gene diversity heterozygosity was higher than observed (global $H_S = 0.087$, global $H_O = 0.072$). The variance within $H_S$ and $H_O$ was found to be significant between all individual populations except for Grassy Flats (Clarion River), Lawsonham B (Red Bank Creek), and River's Edge (Allegheny River) (Supplemental Table S2). Overall global genetic diversity ($H_t$) was calculated to be 0.46.

**Population Structure**

We retained 30 principal components, explaining 35% of variance in our PCA analysis, which reveals separation between all waterways and shows an upstream to downstream separation along the Allegheny River (Figure 3, Supplemental Figure S2). The DAPC supports $K = 6$ with each river clustering separately and the Allegheny River containing three metapopulations (Supplemental Figures S1 and S2; Figure 3; Supplemental Table S3). The DAPC supports greater distinction between the Youghiogheny River and the Greenbrier River from the rest of the Pennsylvania individuals, with the Clarion River, Red Bank Creek, and Allegheny River clustering geographically while also being recovered as separate clusters by the *K*-means clustering function.

The sNMF analysis also indicates $K = 6$ to be the optimal number of genetic clusters. The clusters roughly correspond to waterways, with the Clarion River, Youghiogheny River, and (cultivated) Greenbrier River populations clustering as distinct populations. The Allegheny River was represented by two clusters, one composed of the majority of upstream individuals and the second composed of individuals from downstream and the Red Bank Creek, a tributary that joins the Allegheny in the downstream section (Figure 4). $F_{ST}$ calculations support the genetic distinction of populations from the Clarion River, Greenbrier River, and Youghiogheny River based on rough recommendations for interpreting $F_{ST}$ values (Hartl and Clark 1997; Frankham et al. 2011) (Supplemental Table S1). A Mantel test revealed a significant positive relationship between pairwise geographic and genetic distance (samples from cultivation were excluded) (Mantel statistic $r = 0.642$, $p = 0.002$), indicating a significant pattern of IBD. This trend was also present when

**Table 2.**—Inbreeding coefficient ($F_{IS}$) and within-population gene diversity and observed heterozygosity ($H_S$ and $H_O$) as calculated by "hierfstat". Nonsignificant difference between observed heterozygosity and within population gene diversity denoted by *.

| | | $H_O$ | $H_S$ | $F_{IS}$ |
|---|---|---|---|---|
| Allegheny River | FC* | 0.073 | 0.084 | 0.088 |
| Allegheny River | GP | 0.073 | 0.08 | 0.062 |
| Allegheny River | RR | 0.074 | 0.079 | 0.072 |
| Allegheny River | WH | 0.086 | 0.095 | 0.084 |
| Allegheny River | MC | 0.085 | 0.087 | 0.003 |
| Allegheny River | MR | 0.096 | 0.103 | 0.059 |
| Allegheny River | BCO | 0.075 | 0.083 | 0.085 |
| Allegheny River | CI | 0.073 | 0.085 | 0.129 |
| Allegheny River | CIM | 0.052 | 0.068 | 0.215 |
| Allegheny River | PI | 0.052 | 0.065 | 0.193 |
| Allegheny River | BC | 0.054 | 0.068 | 0.213 |
| Allegheny River | HD | 0.052 | 0.065 | 0.194 |
| Allegheny River | BFI | 0.062 | 0.079 | 0.21 |
| Allegheny River | BEI | 0.066 | 0.089 | 0.266 |
| Allegheny River | EPA | 0.049 | 0.058 | 0.153 |
| Allegheny River | RE | 0.064 | 0.074 | 0.14 |
| Clarion River | GF* | 0.124 | 0.15 | 0.163 |
| Clarion River | CR | 0.12 | 0.155 | 0.206 |
| Red Bank Creek | LA | 0.056 | 0.073 | 0.207 |
| Red Bank Creek | LB | 0.066 | 0.076 | 0.104 |
| Red Bank Creek | RBS | 0.059 | 0.076 | 0.208 |
| Youghiogheny River | YR | 0.074 | 0.089 | 0.151 |
| West Virginia | EWV | 0.093 | 0.113 | 0.154 |
| West Virginia (in cultivation) | EWVC | 0.061 | 0.109 | 0.32 |

examining only the Allegheny River, where the majority of populations occur (Supplemental Figure S2; Mantel statistic $r = 0.4427$, $p = 0.001$). This is supported by the PCA, with clustering from upstream to downstream.

## DISCUSSION

Our findings provide evidence that *Baptisia australis* (sensu Weakley 2024) occurs in genetically structured populations that roughly correspond to the Pennsylvania waterways it inhabits. Our results give us insight into gene flow (via the proxy of fixation index) along and between waterways. and inferences from this study can now be used to optimize management unit delineation and other conservation practices integral to protecting the small number of populations occurring at the northeastern edge of the species' range. Our study illuminates gene flow and population processes in a set of riparian plant populations that will inform modern conservation assessments of rare species.

### Population Genetic Structure

We found evidence for five genetically distinct groups of *Baptisia australis* in Pennsylvania. Based on recommendations for interpreting $F_{ST}$ values put forth by Hartl and Clark (1997), Frankham et al. (2002), and Frankham (2010), the five entities are as follows: the Youghiogheny River, the Clarion River, the upper Allegheny River (above the confluence of the Clarion River), the lower Allegheny River (below the confluence of the Clarion River), and the Red Bank Creek (Figure 1). Each are supported by pairwise $F_{ST} > 0.15$ between populations (Figure 2); this pattern is also supported by the DAPC and sNMF analyses.

The identification of two demes along the Allegheny River presents an interesting challenge for managing *Baptisia australis* in this river system. While *B. australis* populations from each of the other waterways (Clarion River, Red Bank Creek, Youghiogheny River) are distinct groups in the clustering analyses, the populations from the Allegheny River comprise two genetically distinct groups corresponding to upstream and downstream regions, respectively. While the genetic difference between upstream and downstream populations is not completely clear cut, a reduction in gene flow is apparent, with populations around the upstream–downstream divide showing admixture (e.g., Clarion Island; see Supplemental Table S3).

Genetic differences across these *Baptisia* populations could be caused, in part, by habitat differences such as development of surrounding areas. All upstream populations have less nearby development (e.g., homes, bank erosion) than those populations that are downstream, potentially reducing fragmentation and degradation of upstream populations. More developed areas may also experience a reduction in the ecological disturbance necessary to sustain existing *Baptisia* populations, such as annual ice scour. Variation in total population size could also influence $F_{ST}$ values with large populations like Grassy Flats (GF), Bear Creek (BC), and River's Edge (RE) having high $F_{ST}$ values between themselves and the majority of populations (Figure 2).

Supporting our second hypothesis, our clustering analyses and *F* statistics identify the Youghiogheny River population as the most genetically distinct population of *Baptisia australis* in Pennsylvania, sharing little to no current gene flow with other populations in the state ($F_{ST} = 0.292$–$0.354$). The 11 plants sampled there were the only remaining individuals in the last remaining population along the Youghiogheny River, a cause for concern regarding recruitment of new individuals at this location. Only three of these plants had fruits when the population was surveyed in July 2019, and rocky habitat at this location might not be conducive to rhizomatous reproduction. In addition to its small size and isolation, the population has low levels of genetic diversity ($H_O = 0.074$; Cheng et al. 2020; Tarazona-Pulido et al. 2024) and high levels of inbreeding ($F_{IS} = 0.151$; Table 2). These factors make it a population of distinct conservation concern and value (e.g., Ellstrand and Elam 1993). As an isolated population, the Youghiogheny River population follows the expectation of lower levels of genetic diversity and higher measures of genetic differentiation, perhaps related to genetic drift, founder effects, inbreeding, and other bottlenecks in the future (Antonovics et al. 2002; Eckert et al. 2008). Therefore, it is especially important to protect and conserve the population (including whatever seedbank may be present). Banking of seeds from this site could also be important if ex situ conservation of these genotypes is required. Potential future conservation action could include ex situ propagation and outplanting. However, if offspring seem to show low fitness,

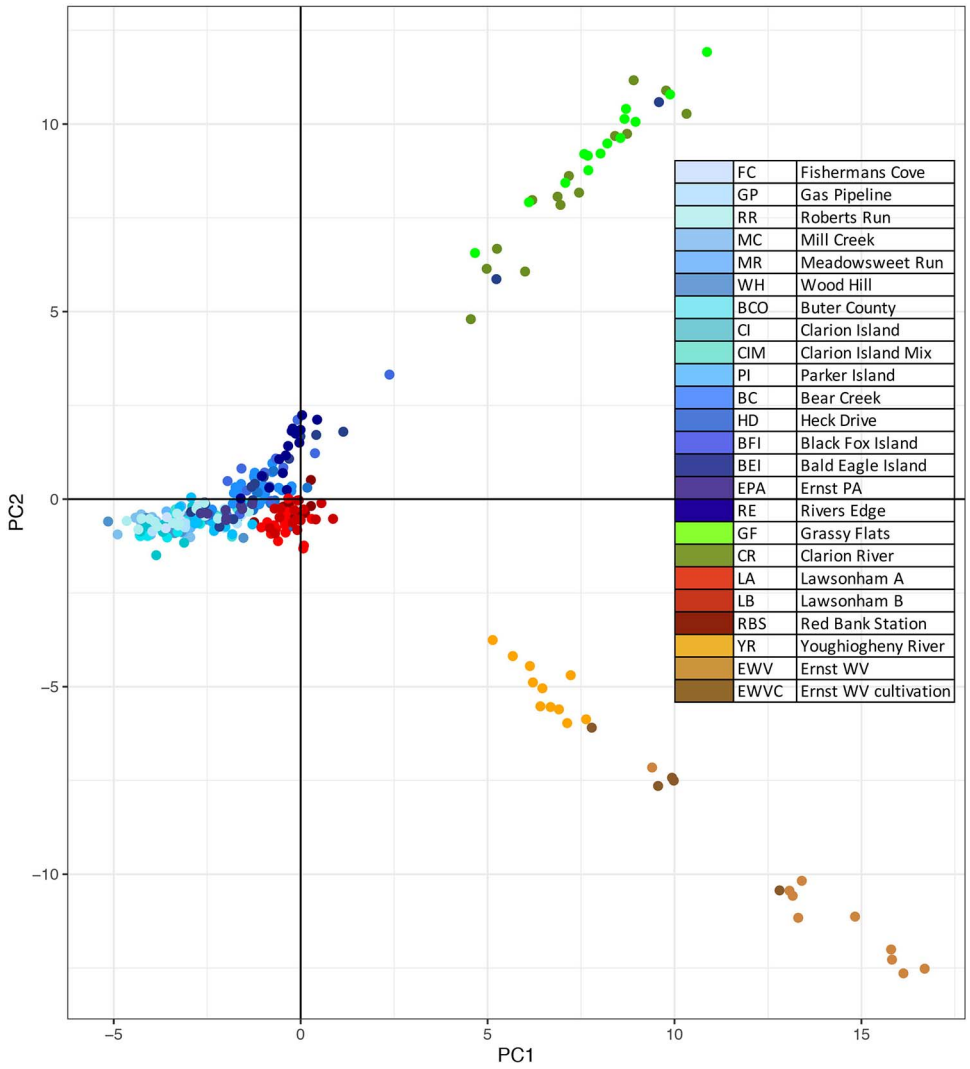


**Figure 3.**—Principal components analysis (PCA) of SNPs of sampled *Baptisia australis* var. *australis*. Shows distinct separation between West Virginia (brown), Youghiogheny (orange), and Clarion River (green) populations. Allegheny River samples (blue) appear to separate from upstream to downstream (light to dark), with the Red Bank Creek (red) sites clustering closely. PCoA Axis 1 explains 7.6% of the total variance; PCoA Axis 1 and 2 combined explain 12.9% of the total variance; PCoA Axis 1–3 combined explain 15.7% of the total variance.

this isolated population is in danger of further inbreeding and genetic drift and could be a candidate for facilitated gene flow, through pollination or seeding from outside sources (Frankham et al. 2017).

**Riparian Population Models**

The $F_{ST}$ values from this study indicate there is some support for a pattern of the classic metapopulation gene flow model in the Allegheny River, Clarion River, and Red Bank Creek system similar to what Tero et al. (2003) found in endangered, riparian *Silene tatarica* (L.) Pers. We did not include the Youghiogheny River population when trying to infer gene flow patterns, focusing instead on the groups and populations within the Allegheny River, Clarion River, and Red Bank Creek system. Despite some support for a classic metapopulation, the notion of stepping-stone populations and riparian unidirectional gene flow could be partially supported based on genetic clustering of the Allegheny River populations into two demes (upper and lower). Where Tero et al. (2003) found no structure among populations, we found evidence for an upstream to downstream unidirectional structure in clustering analyses with the PCA with particular support for riparian unidirectional gene flow along the Allegheny River (Figure 3). Within each of the two demes along the Allegheny River, there seems to be fragmented subpopulations with low to moderate genetic differences.

Our data also provide some modest support for the stepping-stone gene flow model, where populations are likely to exchange genetic material with populations next to them (Figures 2 and 3). The $F_{ST}$ and sNMF analyses suggest that the Red Bank Creek population is more similar to the Allegheny River populations than the Clarion River populations are, with particular affinity of Red Bank Creek to the lower Allegheny River populations. However, we also see a pattern of isolation by distance (IBD)

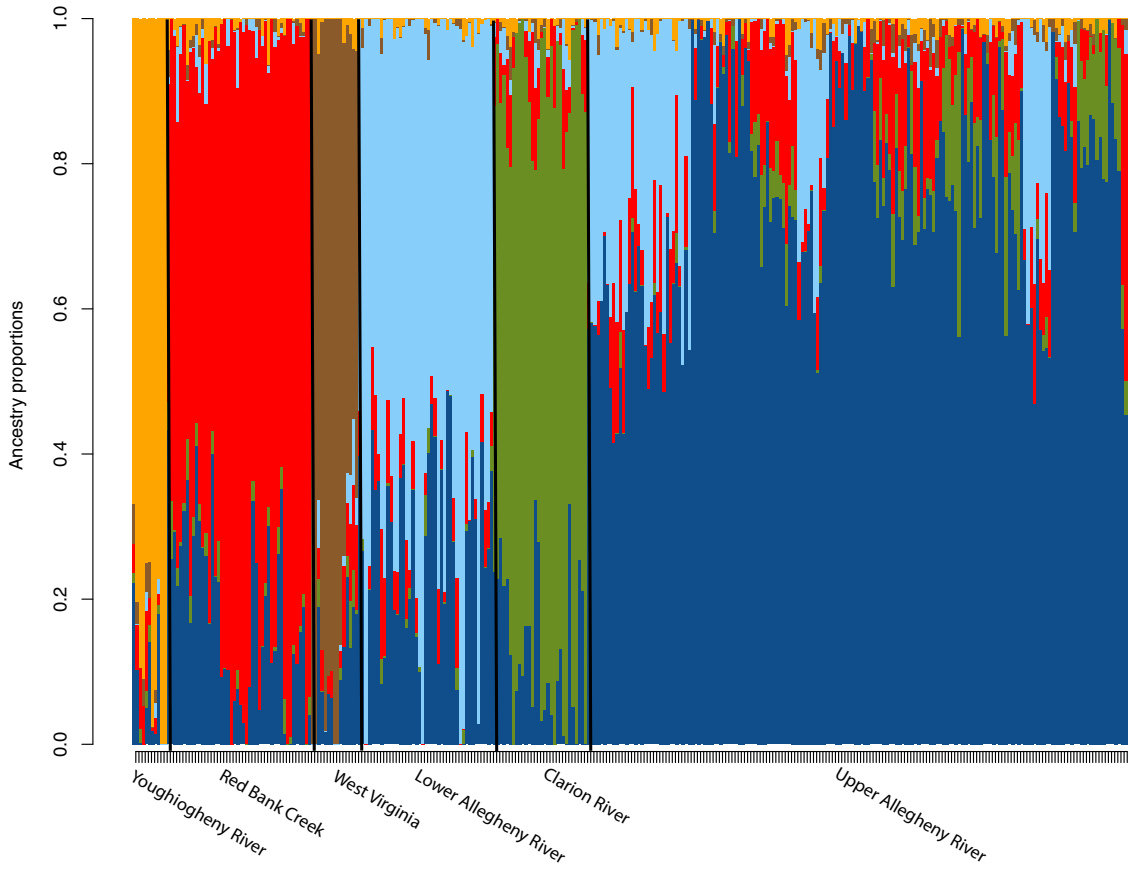


**Figure 4.**—Plot of sNMF ancestry coefficient proportions derived from ancestry plot for 24 sampled populations of *Baptisia australis* var. *australis* with $K = 6$. Five distinctive genetic populations composed of West Virginia (brown), Clarion River (green), Youghiogheny River (orange), and the Allegheny River and Red Bank Creek. Light blue represents upstream populations found on the Allegheny River, while dark Blue represents the downstream section of the Allegheny as well as Red Bank Creek.

both among all populations and along the Allegheny River alone. This supports a stepping-stone pattern of gene flow (Allendorf and Luikart 2007).

## Genetic Health

Most sampled populations are what would be considered to have low genetic diversity considering the life history of the species and high inbreeding (Cheng et al. 2020; Aguirre-Dugua et al. 2023; Wu et al. 2023; Tarazona-Pulido et al. 2024). This raises some concern about the future of the genetic integrity of the populations. $H_O$ and $H_S$ vary significantly across most populations with $H_S$ greater than $H_O$ for all populations (Supplemental Table S2). The few populations that do not have significant variance in $H_O$ and $H_S$ are those that are relatively isolated: Grassy Flats (the most upstream Clarion River population) and Fisherman's Cove (the most upstream population). These two isolated populations may be promising as sources of genetic diversity if facilitated gene flow is chosen as a management strategy.

Management strategies aimed at increasing overall genetic diversity in *B. australis* may be critical as it appears that small, isolated populations of this species can become inbred rather quickly. As an example, West Virginia ecotype *Baptisia* plants growing at Ernst Seeds for seed production (EWV) had an $F_{IS} = 0.154$, while the progeny from the seed crop collected from those plants 7 y after they started producing seeds (C. Ernst, pers. comm., 11 July 2019) had an $F_{IS} = 0.320$. The trend in inbreeding seen in this Ernst population parallels what is known about increased inbreeding in small, isolated populations (Ellstrand and Elam 1993; Honnay and Jacquemyn 2007; Frankham et al. 2017).

## Other Conservation Threats

While the genetic status of *B. australis* in Pennsylvania may not currently appear to be dire, there are other confounding factors that could easily diminish these valuable edge-of-range populations. The species currently faces a depletion of its typical habitat in Pennsylvania, threatened riparian prairies which rely primarily on ice scour—an important disturbance regime considered an analog to fire in other grass-dominated habitats (Prowse and Culp 2003; Rood et al. 2007; Scrimgeour et al. 1994). The decline of riparian grasslands is a trend around the globe, as the disruption of disturbances like burning and grazing facilitate the encroachment of woody species that displace both perennial herbs and grasses and potentially alter ecosystem functions (Veach et al. 2014).

Breakdowns or changes in disturbance regimes can also correspond with incursions of invasive species. In Pennsylvania river scour wet meadows, invasive species such as *Reynoutria* × *bohemica* Chrtek & Chrtková (Bohemian knotweed, Polygonaceae) and *Phalaris arundinacea* L. (reed canary grass, Poaceae) are becoming increasingly prevalent and are impacting other native riparian plants in the region (e.g., Walters and Williams 1999; Wilson et al. 2017). These aggressive invaders were observed to be actively competing with *Baptisia australis* populations sampled and are likely responsible for outcompeting several previously known populations.

Climate change may also pose a threat to riverscour wet meadow communities by reducing the frequency of important scour events (Perles 2019; C. Kentzel, pers. comm., June 2018) and increasing the frequency and severity of flooding events along rivers (Arnell and Gosling 2016). During our study many sites supporting *B. australis* populations along the Allegheny River were underwater for most of June 2019 (C.L. Moore, pers. obs.; USGS), when the species normally flowers and begins to set seed (Rhoads and Block 2007). Inundation likely reduces pollination and ultimately fruit set, thus reducing gene flow and reproductive potential in these populations.

The downstream section of the Allegheny River (especially below the Clarion River confluence) has experienced more human development than upstream reaches, further reducing suitable habitat. Work by Cowell and Stoudt (2002) provides one example along the Allegheny River at the Kinzua Dam near the New York and Pennsylvania border, where a recent history of reduced ice formations below the dam has caused shifts in riparian habitats. The compounded effects of human development and climate change can also be a powerful force in riparian systems, the two factors combining to reduce ice scour and increase flooding and bank erosion. These combined effects, along with erosion from boat wakes due to increased river traffic, could lead to depletion of suitable riverscour wet meadow habitat and loss of *B. australis* seed banks along the Allegheny River (Camfield et al. 1980; Nanson et al. 1994).

**Conservation Importance and Implications**

Effective conservation of rare plants requires that multiple management units or metapopulations be maintained in order to facilitate the long-term persistence of species (Hanski and Gilpin 1997; De Campos Telles et al. 2003; Funk et al. 2012; Coates et al. 2018). Likewise, genetic viability within and among populations may depend on the maintenance of gene flow—perhaps even via facilitated gene flow mediated by conservation practitioners (Allendorf and Luikart 2007; Frankham et al. 2017). Natural Heritage Programs tasked with assessing and protecting rare plants consider these concepts when assigning management units called Elements of Occurrences (EO). At present, the Clarion River, Red Bank Creek, and Allegheny River populations of *Baptisia australis* are considered one combined EO by the Pennsylvania Natural Heritage Program. The guidelines for delimiting EOs are defined as often corresponding with the local population, a portion of a population, or a group of nearby populations (e.g., metapopulation; NatureServe 2021). These definitions seem to imply the genetic distinctiveness of a population or metapopulation. We suggest, based on the results presented here, that the number of *B. australis* EOs in these waterways be divided according to their genetic affinities, thus creating two EOs corresponding to the upper and lower Allegheny River populations, one EO for the Clarion River populations, and one EO for the Red Bank Creek populations. These four EOs along with the Youghiogheny River EO would increase the total number of EOs in Pennsylvania from five to eight.

Recognizing the utility of genetic entities as EOs strengthens the chances that genetic diversity of management units (and across a taxon's distribution) is maintained and available for genetic rescue in the future (e.g., Frankham et al. 2017)—and should be part and parcel of conservation management strategies based on conservation units (as per Coates et al. 2018) as a means to best preserve biodiversity. The direct preservation of genetic diversity in *B. australis* likely has indirect effects, as well, including mitigation of invasive species establishment (Relva et al. 2010; Schuster et al. 2018), maintenance of pollinator habitat to facilitate gene flow (Van Geert et al. 2010; Menz et al. 2011), and minimization of the risks of observed, naturally occurring seed mortality (Petersen and Sleboda 1994; Moore et al. 2021).

An additional conservation implication is at the level of plant community. The Allegheny River, Red Bank Creek, and Clarion River populations are associated with a Willow/Big Bluestem – Indiangrass Riverscour Wet Meadow that is considered a globally imperiled community type with a very restricted range in the Midwest Great Lakes and Ohio River Valley (NatureServe 2021). Likewise, the Youghiogheny River population is associated with a Big Bluestem – Switchgrass – Blue Wild Indigo Riverscour Wet Meadow that is considered a globally vulnerable community with a restricted central Appalachian range (NatureServe 2021). The fact that genetic differentiation between *Baptisia australis* populations correspond with these distinct, globally rare community types suggests that community differentiation could be used as a surrogate for potentially identifying other species that harbor the diversity useful for genetic rescue and preservation of biodiversity.

## CONCLUSIONS

We have applied a population genomics approach to classic conservation genetics questions related to gene flow, genetic structure, and genetic health of a globally rare species. The result of this work impacts the conservation and management of *Baptisia australis* in Pennsylvania and the mid-Atlantic region of the United States through an agency–academic partnership seeking to incorporate genetic information to best define and prioritize units of management. We provide evidence to show that populations of *Baptisia australis* in Pennsylvania are composed of five genetically distinct groups that mainly align with the major waterways they occur along. These systems model classic metapopulations in some ways such as structure of

gene flow among populations on the Allegheny River, but also have characteristics of various other metapopulation models such as sub-structuring within populations. Our results indicate that each of the five genetic groups should be considered an individual EO by the Pennsylvania Natural Heritage Program. Pennsylvania populations of *B. australis* are largely genetically healthy and not currently inbred, but some local populations should receive special consideration as they face acute challenges such as habitat degradation, encroachment by invasive species, declining numbers of individuals, and flooding events likely becoming more frequent and more severe because of climate change. Small and declining populations, especially those containing less-common genotypes, should be targeted for collection and banking of seeds, both as a means to preserve genetic diversity within *B. australis* and to contribute to human-assisted maintenance of gene flow in real time. Facilitated gene flow within and among genetic groups is recommended to maintain the genetic health of rare species, especially when populations are isolated (Frankham et al. 2017).

## ACKNOWLEDGMENTS

Thanks to Mason Heberling and Bonnie Isaac at Carnegie Museum of Natural History and the New York Botanical Garden and Steere Herbarium for access to their specimens. We also thank Carnegie Museum of Natural History for the Buker Travel Award to fund the trip. Much appreciation for genomics advice and support from Dr. Ken Field and Dr. Sarah Lower at Bucknell University. Thanks to Barry Baldwin for assistance in the field. Finally, thanks to our reviewers for your insightful comments and feedback. Funding for this work came from the PA DCNR Wild Resources Conservation Program (project 17571), with additional support from the David Burpee Endowment and the Wayne Manning Student Intern Fund at Bucknell University. We also thank the Botanical Society of America Graduate Student Research Award and Southern Appalachian Botanical Society Earl Core Student Research Award for funding a portion of the fieldwork.

**Data Accessibility:** Sequence data can be found on the National Center for Biotechnology Information (NCBI) website, BioProject number PRJNA574847. The Sequence Read Archive (SRA) is available at https://www.ncbi.nlm.nih.gov/sra/PRJNA574847.
R code used for analyses can be found in the author's github: https://github.com/cheyennelmoore/Baptisia_pop_gen.

**Benefits Generated:** Benefits from this research include the sharing of our data and results on public databases described above.

**Author Contributions:** C.L.M. aided in collecting samples and completed sample preparation, as well as analyzing genomic data and preparing the manuscript draft. A.J.M. aided in sample preparation, provided intellectual contribution regarding appropriate analyses and critically revised the manuscript. S.S. contributed to study design and project coordination in addition to coordinating and participating in sample acquisition. C.T.M. contributed to study design and project coordination, oversaw the project, and provided critical revision and input to the manuscript.

*Cheyenne Moore is the Pennsylvania Plant Conservation Alliance coordinator for the Pennsylvania Department of Conservation and Natural Resources and PA Natural Heritage Program. She earned her MS in Biology studying rare plant population genetics at Bucknell University and her BS from Dickinson College. In her current role, Cheyenne facilitates the conservation and stewardship of Pennsylvania's rare, threatened, and endangered plant species, particularly those that are globally rare.*

*Angela McDonnell earned her PhD at Oklahoma State University and completed Postdoctoral research appointments at Bucknell University and the Chicago Botanic Garden. She is now an assistant professor at St. Cloud State University in St. Cloud, Minnesota, and serves as the director of the Biology Greenhouses and the SCL Herbarium. Her research is focused on phylogenomics and population biology of flowering plants.*

*Scott Schuette is the Botany Program Manager with the Pennsylvania Natural Heritage Program at the Western Pennsylvania Conservancy. He earned his PhD in plant biology from Southern Illinois University, Carbondale. He supervises the collection of rare plant occurrence data and works with his team of botanists to provide accurate and up to date population and habitat information to the Pennsylvania Department of Conservation and Natural Resources. His work involves working with university partners to train graduate students in applied plant conservation using natural heritage methodologies.*

*Chris Martine holds a PhD in Botany from the University of Connecticut and is currently the David Burpee Professor at Bucknell University, where he also serves as Director of the Wayne E. Manning Herbarium. He is a Fellow of the Linnean Society of London and the host and co-producer of the "Plants are Cool, Too!" YouTube series.*

## LITERATURE CITED

Aguirre-Dugua, X., J. Barrera-Redondo, J. Gasca-Pineda, A. Vázquez-Lobo, A. López-Camacho, G. Sánchez-de la Vega, G. Castellanos-Morales, E. Scheinvar, E. Aguirre-Planter, R. Lira-Saade, and L.E. Eguiarte. 2023. Population genomics of domesticated *Cucurbita ficifolia* reveals a recent bottleneck and low gene flow with wild relatives. Plants 12:3989.

Allendorf, F.W., and G.H. Luikart. 2007. Conservation and the Genetics of Populations. Blackwell Publishing, Oxford, UK.

Allendorf, F.W., P.A. Hohenlohe, and G.H. Luikart. 2010. Genomics and the future of conservation genetics. Nature Reviews Genetics 11:697–709.

Alpers, D.L., B.J. Van Vuuren, P. Arctander, and T.J. Robinson. 2004. Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. Molecular Ecology 13:1771–1784.

Andrews, K.R., J.M. Good, M.R. Miller, G. Luikart, and P.A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics 17:81–92.

Antonovics, J., T.J. Newman, and B.J. Best. 2002. Spatially explicit studies on the ecology and genetics of population margins. Pp. 97–117 *in* J. Silvertown and J. Antonovics, eds. Integrating Ecology and Evolution in a Spatial Context. Blackwell, Oxford, UK.

Arnell, N.W., and S.N. Gosling. 2016. The impacts of climate change on river flood risk at the global scale. Climate Change 134:387–401.

Assis, J., N. Castilho Coelho, F. Alberto, M. Valero, P. Raimondi, D. Reed, and E. Álvares Serrão. 2013. High and distinct range-edge genetic diversity despite local bottlenecks. PLOS One 8:e68646.

Bowen, B.W., and S.A. Karl. 2007. Population genetics and phylogeography of sea turtles. Molecular Ecology 16:4886–4907.

Broyles, P.J. 2006. *Baptisia australis*. USDA NRCS PLANTS database. Accessed 28 October 2018 from <https://plants.usda.gov/plantguide/pdf/cs_baau.pdf>.

Camfield, F.E., R.E.L. Ray, and J.W. Eckert. 1980. The possible impact of vessel wakes on bank erosion. Report CG-W-1-8. U.S. Department of Transportation, Washington, DC.

Cheng, J., H. Kao, and S. Dong. 2020. Population genetic structure and gene flow of rare and endangered *Tetraena mongolica* Maxim. revealed by reduced representation sequencing. BMC Plant Biology 20:391.

Cipollini, K., P. Lavretsky, D. Cipollini, and J.L. Peters. 2017. Genetic population structure across the range of endangered northeastern bulrush, *Scirpus ancistrochaetus*. International Journal of Plant Science 178:67–78.

Coates, D.J., M. Byrne, and C. Moritz. 2018. Genetic diversity and conservation units: Dealing with the species–population continuum in the age of genomics. Frontiers in Ecology and Evolution 6:165.

Cowell, C.M., and R.T. Stoudt. 2002. Dam-induced modifications to upper-Allegheny River streamflow patterns and their biodiversity implications. Journal of the American Water Resource Association 38:187–196.

Danecek, P., A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, R.E. Handsaker, G. Lunter, G.T. Marth, S.T. Sherry, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

De Campos Telles, M.P., A.S. Guedes Coelho, L.J. Chaves, J.A. Felizola Diniz-Filho, F. D'Ayala Valva. 2003. Genetic diversity and population structure of *Eugenia dysenterica* DC. ("cagaiteira" - Myrtaceae) in Central Brazil: Spatial analysis and implications for conservation and management. Conservation Genetics 4:685–695.

Delgado, P., L.E. Eguiarte, F. Molina-Freaner, E.R. Alvarez-Buylla, and D. Piñero. 2008. Using phylogenetic, genetic and demographic evidence for setting conservation priorities for Mexican rare pines. Biodiversity and Conservation 17:121–137.

Diekmann, O.E., and E.A. Serrão. 2012. Range-edge genetic diversity: Locally poor extant southern patches maintain a regionally diverse hotspot in the seagrass *Zostera marina*. Molecular Ecology 21:1647–1657.

Eaton, D.A.R. 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30:1844–1849.

Eckert, C.G., K.E. Samis, and S.C. Lougheed. 2008. Genetic variation across species' geographical ranges: The central–marginal hypothesis and beyond. Molecular Ecology 17:1170–1188.

Ellstrand, N.C., and D.R. Elam. 1993. Population genetic consequences of small population size: Implications for plant conservation. Annual Review of Ecology, Evolution, and Systematics 24:217–242.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, S.E. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLOS One 6: e19379.

ESRI. 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

Frankham R. 2010. Challenges and opportunities of genetic approaches to biological conservation. Biological Conservation 143:1919–1927.

Frankham, R., J.D. Ballou, and D.A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK.

Frankham, R., J.D. Ballou, M.D. Eldridge, R.C. Lacy, K. Ralls, M.R. Dudash, and C.B. Fenster. 2011. Predicting the probability of outbreeding depression. Conservation Biology 25:465–475.

Frankham, R., J.D. Ballou, K. Ralls, M.R. Dudash, C.B. Fenster, and P. Sunnucks. 2017. Genetic Management of Fragmented Animal and Plant Populations. Oxford University Press, Oxford, UK.

Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973– 983.

Frichot, E., and O. François. 2015. LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution 6:925– 929.

Funk, W.C., J.K. McKay, P.A. Hohenlohe, and F.W. Allendorf. 2012. Harnessing genomics for delineating conservation units. Trends in Ecology and Evolution 27:489–496.

Goudet, J. 2005. hierfstat, a package for R to compute and test hierarchical *F*-statistics. Molecular Ecology Resources 5:184–186.

Hale, H., E.M. Gardner, J. Viruel, L. Pokorny, and M.G. Johnson. 2020. Strategies for reducing per-sample costs in target capture sequencing for phylogenomics and population genomics in plants. Applied Plant Science 8:e11337.

Hanski, I. and M.E. Gilpin, eds. 1997. Metapopulation Biology: Ecology, Genetics, and Evolution. Academic Press, San Diego, CA.

Hartl, D.L., and A.G. Clark. 1997. Principles of Population Genetics. Sinauer, Sunderland, MA.

Hayes, J.D. 2021. Genetic diversity & connectivity of *Chasmanthium latifolium* (Poaceae) in Pennsylvania & the effect on conservation status of a rare species. Honors thesis, Bucknell University, Lewisburg, PA. <https://digitalcommons.bucknell.edu/honors_theses/577>

Honnay, O., and H. Jacquemyn. 2007. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. Conservation Biology 21:823–831.

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

Jombart, T., S. Devillard, and F. Ballou. 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics 11:94.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071.

Kamvar, Z.N., J.F. Tabima, and N.J. Grünwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281.

Kamvar, Z.N., J.C. Brooks, and N.J. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in Genetics 6:208.

Kartesz, J.T. 2003. A synonymized checklist and atlas with biological attributes for the vascular flora of the United States, Canada, and Greenland. 2nd ed. *In* J.T. Kartesz. Synthesis of the North American Flora, Version 2.0.

Knaus, B.J., and N.J. Grünwald. 2017. VCFR: A package to manipulate and visualize variant call format data in R. Molecular Ecology Resources 17:44–53.

Lee, S.L., K.K.S. Ng, L.G. Saw, C.T. Lee, N. Muhammad, N. Tani, Y. Tsumura, and J. Koskela. 2006. Linking the gaps between conservation research and conservation management of rare dipterocarps: A case study of *Shorea lumutensis*. Biological Conservation 131:72–92.

McDonnell, A.J., C.M. Moore, S. Schuette, and C.T. Martine. 2021. Population genomics and conservation of *Erigenia bulbosa* (Apiaceae), an edge-of-range species in Pennsylvania. International Journal of Plant Science 182:344–355.

McLennan, E.A., B.R. Wright, K. Belov, C.J. Hogg, and C.E. Grueber. 2019. Too much of a good thing? Finding the most informative genetic data set to answer conservation questions. Molecular Ecology Resources 19:659–671.

Menz, M.H.M., R.D. Phillips, R. Winfree, C. Kremen, M.A. Aizen, S.D. Johnson, and K.W. Dixon. 2011. Reconnecting plants and pollinators: Challenges in the restoration of pollination mutualisms. Trends in Plant Science 16:4–12.

Moore, C.L., A.J. McDonnell, S. Schuette, and C.T. Martine. 2021. Lepidopteran granivory reduces seed counts in a rare species of riparian scour prairies. Natural Areas Journal 41:47–54.

Nanson, G.C., A. Von Krusenstierna, E.A. Bryant, and M.R. Renilson. 1994. Experimental measurements of river-bank erosion caused by boat-generated waves on the Gordon River, Tasmania. Regulated Rivers Research and Management 9:1–14.

NatureServe. 2021. NatureServe Explorer: An online encyclopedia of life. Version 7.1. NatureServe, Arlington, VA. Accessed 12 November 2021 from <http://explorer.natureserve.org>

Noss, R.F., J.M. Cartwright, D. Estes, T. Witsell, G. Elliott, D. Adams, M. Albrecht, R. Boyles, P. Comer, C. Doffitt, et al. 2021. Improving species status assessments under the U.S. Endangered Species Act and implications for multispecies conservation challenges worldwide. Conservation Biology 35:1715–1724.

Palsbøll, P.J., M. Bérubé, and F.W. Allendorf. 2007. Identification of management units using population genetic data. Trends in Ecology and Evolution 22:11–16.

Pennsylvania Department of Conservation and Natural Resources. 2017. Native wild plant species accounts. Department of Conservation and Natural Resources Conservation of Native Wild Plants (Chapter 45 Subchapter B) Rulemaking Change Data. Pennsylvania Department of Conservation and Natural Resources.

Perles, S. 2019. Protecting rare riparian prairies using science-based management. Paper presented at Natural Areas Conference 2019, Pittsburgh, PA, 9 October 2019.

Petersen, C.E., and J.A. Sleboda. 1994. Selective pod abortion by *Baptisia leucantha* (Fabaceae) as affected by a curculionid seed predator, *Apion rostrum* (Coleoptera). Great Lakes Entomologist 27:143–147.

Peterson, B.K., J.N. Weber, E.H. Kay, H.S. Fisher, and H.E. Hoekstra. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model Species. PLOS One 7:e37135.

Prowse, T.D., and J.M. Culp. 2003. Ice breakup: A neglected factor in river ecology. Canadian Journal of Civil Engineering 30:128–144.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Relva, M.A., M.A. Nuñez, and D. Simberloff. 2010. Introduced deer reduce native plant cover and facilitate invasion of non-native tree species: Evidence for invasional meltdown. Biological Invasions 12:303–311.

Resh, C.A., M.P. Galaska, and A.R. Mahon. 2018. Genomic analyses of northern snakehead (*Channa argus*) populations in North America. PeerJ 6:e4581.

Rhoads, A.F., and T.A. Block. 2007. *Baptisia*. P. 574 *in* The Plants of Pennsylvania: An Illustrated Manual. 2nd ed. University of Pennsylvania Press, Philadelphia.

Roberts, D.G., D.J. Ayre, and R.J. Whelan. 2007. Urban plants as genetic reservoirs or threats to the integrity of bushland plant populations. Conservation Biology 21:842–852.

Rood, S.B., L.A. Goater, J.M. Mahoney, C.M. Pearce, and D.G. Smith. 2007. Floods, fire, and ice: Disturbance ecology of riparian cottonwoods. Canadian Journal of Botany 85:1019–1032.

Rowe. H.C., S. Renaut, and A. Guggisberg. 2011. RAD in the realm of next-generation sequencing technologies. Molecular Ecology 20:3499–3502.

Schilling, M.P., P.G. Wolf, A.M. Duffy, H.S. Rai, C.A. Rowe, B.A. Richardson, and K.E. Mock. 2014. Genotyping-by-sequencing for *Populus* population genomics: An assessment of genome sampling patterns and filtering approaches. PLOS One 9:e95292.

Schuster, M.J., P.D. Wragg, and P.B. Reich. 2018. Using revegetation to suppress invasive plants in grasslands and forests. Journal of Applied Ecology 55:2362–2373.

Scrimgeour, G.J., T.D. Prowse, J.M. Culp, and P.A. Chambers. 1994. Ecological effects of river ice break-up: A review and perspective. Freshwater Biology 32:261–275.

Seeb, J.E., G. Carvalho, L. Hauser, K. Naish, S. Roberts, and L.W. Seeb. 2011. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11:1–8.

Sexton, J.P., P.J. McIntyre, A.L. Angert, and K.J. Rice. 2009. Evolution and ecology of species range limits. Annual Review of Ecology, Evolution, and Systematics 40:415–436.

Sexton, J.P., S.Y. Strauss, and K.J. Rice. 2011. Gene flow increases fitness at the warm edge of a species' range. Proceedings of the National Academy of Sciences 108:11704–11709.

Silliman, K. 2019. Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. Evolutionary Applications 12:923–939.

Tarazona-Pulido, L., P.A. Rugeles-Silva, K.M. Cardona Tobar, L.A. Díaz-Ariza, J.E. Muñoz Florez, and D. López-Álvarez. 2024. Approach of genetic diversity of *Lippia alba* (Mill) and *Petiveria alliacea* L.: Medicinal plants of Colombia. Plant Molecular Biology Reporter. <https://doi.org/10.1007/s11105-024-01484-z>

Tero, N., J. Aspi, P. Siikamaki, A. Jakalaniemi, and J. Tuomi. 2003. Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. Molecular Ecology 12:2073–2085.

Tero, N., J. Aspi, P. Siikamäki, and A. Jäkäläniemi. 2005. Local genetic population structure in an endangered plant species, *Silene tatarica* (Caryophyllaceae). Heredity 94:478–487.

Van Geert, A., V. Rossum, and L.Triest. 2010. Do linear landscape elements in farmland act as biological corridors for pollen dispersal? Journal of Ecology 98:178–187.

Veach, A.M., W.K. Dodds, and A. Skibbe. 2014. Fire and grazing influences on rates of riparian woody plant expansion along grassland streams. PLOS One 9:106922.

Walters, G.L., and C.E. Williams. 1999 Riparian forest overstory and herbaceous layer of two upper Allegheny River islands in northwestern Pennsylvania. Castanea 64:81–89.

Weakley, A.S. 2024. Flora of the Southeastern United States. University of North Carolina Herbarium, North Carolina Botanical Garden, Chapel Hill, NC. <https://ncbg.unc.edu/research/unc-herbarium/floras/>

Weir, B.S., and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370.

Wilson, M.J., A.E. Freundlich, and C.T. Martine. 2017. Understory dominance and the new climax: Impacts of Japanese knotweed (*Fallopia japonica*) invasion on native plant diversity and

recruitment in a riparian woodland. Biodiversity Data Journal 5: e20577.
Woods, M., and A.R. Diamond. 2014. The genus *Baptisia* (Fabaceae) in Alabama. Phytoneuron 83:1–11.
Wright, B.R., K.A. Farquharson, E.A. McLennan, K. Belov, C.J. Hogg, and C.E. Grueber. 2020. A demonstration of conservation genomics for threatened species management. Molecular Ecology Resources 20:1526–1541.
Wu, Q., S. Dong, Y. Zhao, L. Yang, X. Qi, Z. Ren, S. Dong, and J. Cheng. 2023. Genetic diversity, population genetic structure and gene flow in the rare and endangered wild plant *Cypripedium macranthos* revealed by genotyping-by-sequencing. BMC Plant Biology 23:254.
Zimmerman, E., T. Davis, G. Podniesinski, M. Furedi, J. McPherson, S. Seymour, B. Eichelberger, N. Dewar, J. Wagner, and J. Fike, eds. 2012. Terrestrial and Palustrine Plant Communities of Pennsylvania. 2nd ed. Pennsylvania Natural Heritage Program, Pennsylvania Department of Conservation and Natural Resources, Harrisburg, PA.