The Genetic Diversity And Population Structure Of *Geum radiatum*: Effects Of A Past Augmentation Of An Endangered Hexaploid

By: **Nikolai M. Hay**, Chris Ulrey, Gary Kauffman, **Zack E. Murrell**, and **Matt C. Estep**

Abstract

Geum radiatum is a federally endangered high-elevation rock-outcrop endemic herb that is widely recognized as a hexaploid and a relic species. Little is known about G. radiatum genetic diversity, population interactions, or the effect of past augmentations of populations. This study sampled every known population of G. radiatum and used microsatellite markers to measure genetic diversity and population structure. The analysis demonstrates that there is interconnectedness and structure among populations. In addition, the analysis was able to differentiate transplanted individuals and identify putative anthropogenically admixed individuals within augmented populations. Geum radiatum exhibits diversity within and among populations and current gene flow connects the northern populations. This information provides a greater understanding of the genetic sustainability of G. radiatum and what conservation efforts will most help this imperiled species to survive.

Hay NM, Ulrey C, Kauffman G, **Murrell ZE, Estep MC**. The Genetic Diversity and Population Structure of Geum radiatum : Effects of a Past Augmentation of an Endangered Hexaploid. Castanea. 2019;84(2):273-288.
Publisher version of record available at: https://www.jstor.org/stable/26865725

The Genetic Diversity and Population Structure of *Geum radiatum*: Effects of a Past Augmentation of an Endangered Hexaploid

Nikolai M. Hay,[1] Chris Ulrey,[2] Gary Kauffman,[3] Zack E. Murrell,[1] and Matt C. Estep[1*]

[1]Department of Biology, Appalachian State University, Rankin Science Building, 572 River St, Boone, NC 28608

[2]National Park Service, 199 Hemphill Knob Rd, Asheville, NC 28803

[3]U.S. Forest Service, 160 Zillicoa St, Asheville, NC 28801

ABSTRACT

Geum radiatum is a federally endangered high-elevation rock-outcrop endemic herb that is widely recognized as a hexaploid and a relic species. Little is known about *G. radiatum* genetic diversity, population interactions, or the effect of past augmentations of populations. This study sampled every known population of *G. radiatum* and used microsatellite markers to measure genetic diversity and population structure. The analysis demonstrates that there is interconnectedness and structure among populations. In addition, the analysis was able to differentiate transplanted individuals and identify putative anthropogenically admixed individuals within augmented populations. *Geum radiatum* exhibits diversity within and among populations and current gene flow connects the northern populations. This information provides a greater understanding of the genetic sustainability of *G. radiatum* and what conservation efforts will most help this imperiled species to survive.

Key words: augmentation, cliff-face, conservation, endangered species, *Geum radiatum*, hexaploid, population genetics

INTRODUCTION

Geum radiatum Michx, (Rosaceae), commonly known as Spreading Avens or Appalachia Avens, is a rare Appalachian endemic perennial herb found on high-elevation rock-outcrops and a grassy bald above 1500 m in western North Carolina (NC) and eastern Tennessee (TN) in the eastern United States (Weakley 2015). *Geum radiatum* has showy yellow flowers contained within a cymose inflorescence, typically fewer than five larger basal leaves with 2–5 small sessile orbicular stem leaflets that are pinnately divided per rosette, with a large terminal reniform serrate-margined leaflet. The plants typically grow in dense mats from a horizontal rhizome occupying cracks and crevasses of rock outcrops and cliff faces, though one population is known from an open grassy area. *Geum radiatum* grows in close association with several other rare mountaintop pseudo-alpine species, many of which are also considered in peril, including *Houstonia montana* Small, *Carex misera* Buckley, *Calamagrostis cainii* Hitchc., *Juncus trifidus* L., *Gymnoderma lineare* (A. Evans) Yoshim. & Sharp, and *Solidago spithamaea* M.A. Curtis ex A. Gray (Wiser 1994, Weakley 2015).

Geum radiatum has been federally listed as endangered under the Endangered Species Act since 1990 (Murdock 1993) and is only known from 14 sites in NC and TN (Figure 1), of which 10 are managed by state and federal agencies. Several of the populations are small and have been damaged by trampling or other recreational activity and/or development. *Geum radiatum* is at a high risk of extinction from both global warming (Ulrey et al. 2016) and the continued residential and recreational development of high elevation sites in the Southern Appalachians (Godt et al.

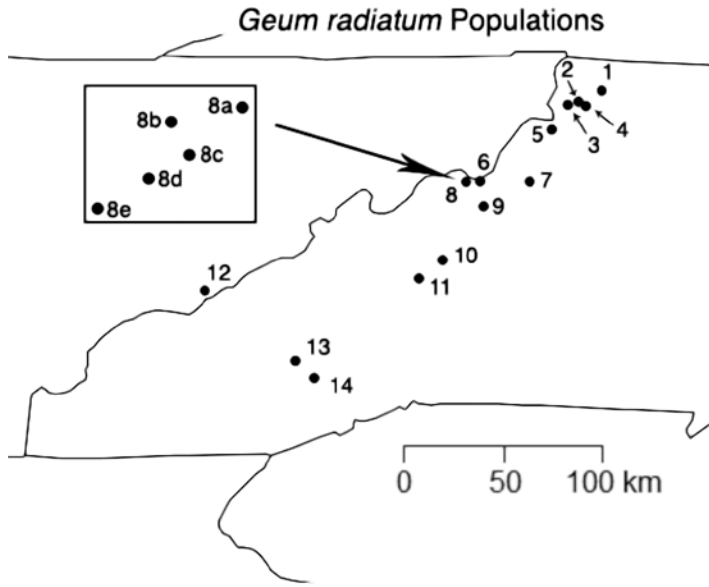

Figure 1. The geographic distribution of *G. radiatum*. The outline of North Carolina, Tennessee, and Virginia are shown. Populations are labeled with the sample name used in the study to protect the location information.

1996, Wiser 1998). Population 8 (see above) is in the center of the range, one of the largest remaining populations, and comprises multiple subpopulations that are found on close but discontinuous cliff systems surrounded by dense spruce-fir forest. Individuals from Population 1, from the north of the range, were transplanted to Population 8 in the center of the range to bolster failing subpopulations (demographic rescue) effected by trampling and recreational activity (correspondence with USFS). In the present study, Population 8 was heavily sampled across all subpopulations to try to elucidate the impact this augmentation may have had on genetic diversity.

The current range of *Geum radiatum* extends from northwest North Carolina to southwest North Carolina following the highest peaks of the Blue Ridge Mountains close to the border with TN, with all populations within NC, except one isolated population in TN near the southern end of the range. These populations are between 1341 to 2012 m in elevations. Populations on the lowest elevations are in the northern end of the range found within the Amphibolite Mountains Macrosite (AMM) (Poindexter and Murrell 2008). These northern populations are much closer to each other geographically compared to southern populations. In contrast with the southern populations there is a conspicuous lack of spruce-fir forest in the AMM and the unique geology creates a relatively neutral pH mafic soil (Peet et al. 2003).

Geum radiatum is thought to be a relic alpine species that was more widespread at the end of the last ice age (Wiser 1994). The species retreated to the cooler mountain tops of the Southern Appalachians and became stranded when the earth warmed, unable to retreat further north (Weakley, 2015). Its closest relatives include *Geum peckii* Pursh, known from alpine communities in the Presidential Range in New Hampshire and from one coastal community in Nova Scotia (Paterson and Synder 1999), and *G. calthifolium* Smith, an alpine herb from the southern coast of Alaska and British Colombia to Russia and Japan and throughout the Aleutian Islands (Rohrer 2014). The distribution of above timberline alpine communities like those found in the Presidential Range are hypothesized to have covered large areas of the Southern Appalachians during the Pleistocene, and these remnant populations represent relictual species distributions (Delcourt and Delcourt 1998). The outcrop habitat occupied by *G. radiatum* is the closest analog in the Southern Appalachians to an alpine habitat and is often found within spruce-fir mountain top forests. These

rock faces have high sun exposure, low annual temperatures, and are some of the wettest high-elevation cliffs in the Southern Appalachians (Ulrey et al. 2016). Similar rock outcrops in the same region are ecologically maintained by periodic fire and are devoid of *G. radiatum*. Those rock outcrops with *G. radiatum* are lacking in rare fire-dependent rock outcrop specialists such as *Liatris helleri* Porter and *Hudsonia montana* Nutt. suggesting the lack of fire plays a key role in defining habitat preferences.

Geum radiatum as well as many other members of the genus *Geum,* including *G. peckii* and *G. calthifolium,* are hexaploid (2n=6x=42) as reported by Gajewski (1957). Microsatellite markers work especially well in diploids with only two sets of chromosomes and therefore generally a maximum of two alleles per genetic locus. The possible number of allele sites per locus equals the number of chromosome sets or the ploidy of the individual, for a hexaploid, such as *G. radiatum,* to be a homozygote all six alleles would be identical. In polyploids, the number of possible heterozygous states increases dramatically. This problem compounds as ploidy increases the number of possible alleles at each site, known as allele dosage, and obscures the ratios of alleles at each locus (De Silva et al. 2005). Thus, allelic dosage greatly increases the difficulty of interpreting heterozygosity as a measure of genetic diversity. Though complicating analysis for scientists, polyploid genetic diversity may more robustly resist fixation (Glendinning 1987) and may lessen the effects of inbreeding depression (Husband and Schemske 1997, Rausch and Morgan 2007). In rare plants and narrowly distributed plants, such as *G. radiatum*, polyploidy may help preserve genetic diversity by reducing the occurrence of homozygosity of deleterious alleles from inbreeding (Buza et al 2000).

Publicly available data on the distribution of *Geum radiatum,* range from rapid plant inventories by conservation agencies such as United States Fish and Wildlife Service (USFWS), United States Forest Service (USFS), Blue Ridge Parkway (BRP), Grandfather Mountain Inc., North Carolina Heritage Program (NCHP), and North Carolina State Parks (NCSP), to more extensive floristic studies that provide an understanding of community structure and type (Tucker 1972, Jenkins 2011). Other studies on *G. radiatum* include a prediction of tree-line shifts in the Southern Appalachians, which identifies *G. radiatum* as an herbaceous floral component of a Pleistocene Southern Appalachian alpine community (Cogbill et al. 1997). More recently, an extensive 10-year demographic study and species distribution model that projected different possible future CO_2 concentrations and the climatic implications for *G. radiatum* and its survival (Ulrey et al. 2016). Demographic investigations suggest that *G. radiatum* is a very long-lived perennial plant with high survivorship in undisturbed habitats and low seedling recruitment (Ulrey et al. 2016). Monitoring conducted over 10 years found that most populations had no seedling recruitment, but that most populations did not lose a single adult plant during that period of time (Ulrey et al. 2016). *Geum radiatum* resides close to its thermal maximum and is therefore difficult to grow from seed even within close proximity to extant populations. In a previous study, the establishment of seedlings required the use of air-conditioned greenhouses (Johnson 1995). The lack of seedling recruitment and difficulties in seed germination suggests that *G. radiatum* is a species that cannot naturally or quickly recover from anthropogenic damage. In some cases, it may take centuries for populations that are at their climatic limit to regenerate from even modest disturbance, like trampling (Johnson 1995). Although there has been a great deal of locality information collected on *G. radiatum* populations and also some interspecific morphological comparisons (Medford 2001), there has been only one population genetics study, which found very low genetic diversity and very small genetic distances between individuals among populations (Godt et al. 1996). The study tested diversity and richness within four different high elevation rock outcrop species and found that within *G. radiatum* the level of diversity was lower than expected for a narrowly endemic plant. Diversity at the population level was comparable to what has been seen in other narrowly endemic plants in the study (Godt et al 1996).

Geum radiatum is a charismatic plant and has long been sought by botanists such as Asa Gray (Gray 1889). This fascination has led to over-collecting of the species and has complicated the history of conservation actions. Past attempts to rescue failing populations of *G. radiatum* resulted in

some subpopulations being augmented with plants from alternative source populations, without consideration of genetic information (correspondence with USFS). These augmentations may have yielded anthropogenically admixed descendants. No research has been performed on the effects such conservation strategies could have on the genetic makeup and viability of native populations.

The goals of this study were to describe the genetic diversity contained within the 14 extant populations, evaluate how this diversity is distributed within the species, and to examine our ability to detect transplanted individuals and admixed descendants in Population 8. The role of these augmentations in the long-term survival of *G. radiatum* could be key to our understanding what effects human mediated gene flow can have on rare plant conservation in the face of climate change.

MATERIALS AND METHODS

Collection

The federal status of *Geum radiatum* guarantees protection and management on State and Federal lands. Populations on private sites are the most at risk for destruction and eventual extirpation by landowners. Collection permits were obtained from USFWS (agent Dale Suiter's permit; TE178876-1), North Carolina Department of Agriculture's Plant Conservation Program Endangered (Permit #355), NCSP (R14-45), Great Smoky Mountains National Park (GRSM-2014-SCI-1184), and North Carolina Game Lands (14-ES00404). On other federal lands collections were made with agency personnel to ensure collection guidelines. In accordance with our permit agreements, no specific locality data can be provided. All known extant populations of *G. radiatum* were sampled from 2014–2016 (Figure 1). An approximately 100 mm^2 piece of leaf tissue was harvested from each individual for DNA extraction. Plant tissue samples were placed on silica gel (Sigma-Aldrich 294316) when collected and stored in an -80° C freezer for long-term storage. An individual was defined as having at least 25 cm separation from other rhizomatous clumps, although these clumps may represent more than a single genetic individual. Twenty individuals were sampled from each population when available. In those cases where populations were smaller, all individuals found at the site were sampled. Larger populations with over 20 individuals were sampled at random intervals along the cliff face or topography of the population with the goal of maximizing the area sampled while minimizing damage to the population from trampling. Only plants that were deemed healthy and large enough to be sampled without long-term damage were sampled.

DNA Extraction

Dried tissue was disrupted and ground to a fine powder using a micro-pestle and sterile sand in a microcentrifuge tube. DNA was extracted from the powder using an Invitrogen PureLink Plant Total DNA Purification Kit (Invitrogen, Carlsbad, California) or a Qiagen DNeasy Plant Mini Kit (Qiagen Valencia, California) following the manufacture's protocol. A Nano-drop 1000 (Thermo Fisher Scientific, Waltham, Massachusetts) was used to assess the quantity of DNA. The DNA was examined for quality using a 1% TBE agarose gel. Low concentration samples were concentrated with a traditional NaOAc and Ethanol precipitation.

Genotyping

DNA was diluted to 20 ng/ul and arrayed into multiple 96 well plates. Each 96 well plate array contained two individual samples chosen as controls that were placed in different positions on each plate to ensure uniform scoring. Microsatellite markers from closely related species *Geum urbanum* and *G. reptans* were tested for transferability to *G. radiatum* (Arens et al. 2004, Hamann et al. 2014). Five microsatellite markers were selected from *G. urbanum* (WGU5-12, WGU8-1, WGU6-23, WGU6-1, and WGU3-15) (Arens et al. 2004). Three were selected from *G. reptans* (003651, 011534, and 014769) (Hamann et al. 2014). Each forward primer was modified to include a 5' M13 tag (5'-CACGACGTTGTAAAACGAC-3') for fluorescent labeling of PCR products using a third primer labeled with FAM, VIC, NED, or PET (Life Technologies, Grand Island, New York) (Schuelke, 2000). Polymerase chain reactions were prepared in 10μL volumes consisting of GoTaq Flexi Buffer, 2.5 mM $MgCl_2$, 800 μM dNTPs, 0.5 μM of reverse primer, 0.25 μM of tagged forward primer, 0.25 μM

of a M13 fluorescent labeled primer, 0.5 units of GoTaq Flexi DNA Polymerase, and ~20ng of DNA (Promega, Madison, Wisconsin). PCR was completed using a touchdown thermal cycling program on an Eppendorf Mastercyler thermal cycler (Eppendorf, Hauppauge, New York). Initial denaturation was at 94°C for five minutes, followed by 13 cycles at 94°C for 45 seconds, 68°C for two minutes descending 1°C in temperature per cycle, and 72°C for one minute. These were then followed by 25 cycles at 94°C for 45 seconds, 55°C for one minute, and 72°C for one minute, and a final extension of 72°C for five minutes. Different fluorescently tagged PCR products from the same individual were combined to pseudo-multiplex four markers that were added to HI-DI (Applied Biosystems, Foster City, California) with a GeneScan Liz 500 size standard (Applied Biosystems). Samples were shipped to Georgia Genomics (UGA, Athens, Georgia) and were separated using an ABI 3730 Sequencer (Applied Biosystems). The resulting chromatograms were scored in Geneious 9.1 using the microsatellite plug-in (Biomatters, Auckland, NZ). Individuals were scored with the potential of six distinct peaks, as *G. radiatum* is a hexaploid (Gajewski 1957).

Statistical Analysis

Basic descriptive statistics, including the number of alleles per locus, number of alleles per locus per population, total number of alleles, and the allelic ranges were calculated in Microsoft Excel (Redmond, Washington).

The allelic frequency was estimated with the function simpleFreq, which assumes partial heterozygosity, in polysat (version 1.6.0, Clark and Jasieniuk 2011) in the R statistical language (version 3.3.2, R Core Team, 2016). Genetic distances were calculated using the Bruvo method (Bruvo et al. 2004) and the Lynch method (Lynch 1990). Principal Coordinate Analyses was performed in polysat. Deviations in total heterozygosity (HT) and inbreeding coefficients of (F_{IS}) and (G_{IS}) from Hardy-Weinberg equilibrium were tested in GENODIVE (version 2.0 b27, Meirmans and Van Tienderen 2004). Fixation and population structure where estimated with Wright's F_{ST} statistics (Wright 1943, 1965), Nei's G_{ST} (Nei 1973), and Jost's D (Jost 2008) using polysat.

Population structure was investigated using the Bayesian clustering software STRUCTURE (version 2.3.4, Pritchard et al. 2000, Falush et al. 2003) assuming an admix model with a ploidy of six. A K-value analysis was run 100,000 times with a burn-in period of 25,000 in replicates of five from k=1 to k=13. The appropriate K-value was determined using the Evanno et al. method (2005) in STUCTURE HARVESTER (version 0.6.94, Earl and von Holdt 2012). Bar graphs of genetic clusters were generated using POPHELPER (version 2.2.0, Francis 2017). A final STRUCTURE analysis was run 5,000,000 times with a burn-in period of 1,000,000 at K=4. A map of population structure displayed as pie graphs was performed with Excel and R the statistical language using the Q-matrix output from STRUCTURE. Mapping of the predicted ancestral coefficients was performed using POPS R scripts, the Q-matrix produced by STRUCTURE and the GPS coordinates of sampled populations (Jay et al. 2011).

Hybrid Index and Augmentation Analysis

Populations known to have been augmented with plants from a distant population were tested in GenoDive using the Hybrid Index function (version 1.2.3, Gompert and Buerkle 2009) and a maximum-likelihood hybrid index method with an admix model, where the genotypes of non-augmented and putative native plants were defined. The putative native genotypes were identified using STRUCTURE results and prior knowledge of augmented plants nativity. The results of the hybrid index were mapped using the R package maps and mapplots (version 3.1.1, Becker et al., 2016, version 1.5, Gerritsen and Gerritsen 2014).

RESULTS

Genetic Diversity

A total of 310 individuals were genotyped from all 14 of the known populations of *G. radiatum* (Figure 1). A total of 141 alleles were identified across eight microsatellite loci. The number of alleles per locus ranged from six to 27, with an average of 17.6 (Table 1). The effective number

Table 1. Descriptive statistics of the loci across all populations.

Locus	Author	N	A	Ae
003651	Hamann et al.	227	27	3.286
011534	Hamann et al.	264	16	1.698
014769	Hamann et al.	219	6	1.934
WGU3-15	Arens et al.	214	24	2.999
WGU5-12	Arens et al.	272	16	4.746
WGU6-1	Arens et al.	290	15	2.622
WGU6-23	Arens et al.	307	22	3.574
WGU8-1	Arens et al.	280	15	3.114
Mean		259.125	17.625	2.997

Note: N = the total number of genotyped individuals at each loci, A = the total number of alleles at each locus, Ae = the effective number of alleles at each locus calculated by reciprocal of the expected homozygosity.

of alleles per locus (calculated as the reciprocal of expected homozygosity) ranged from 1.698 to 4.746, with an average of 2.997.

The populations were numbered in order from north to south to protect the location identification. The number of individuals scored per population ranged from three to 81, with an average of 22.143 (Table 2). The small number of sampled individuals in Populations 4 and 9 were the result of small population sizes. The eight microsatellite markers were scored with an average allelic richness of 5.666, ranging from 2.571 to 9.875 (Table 2). The total number of alleles identified in a single population ranged from 19 in Population 9 to 79 in Population 8 with a mean of 43.857. The effective number of alleles ranged from 2.138 in Population 9 to 3.611 in Population 5 with a mean of 2.933 (Table 2). None of the markers were monomorphic for any population. The total observed heterozygosity ranged from 0.489 to 0.783, with an average of 0.627. The estimated F_{IS} values range from -0.231 to -0.592, with an overall F_{IS} -0.413 across all populations. The estimated G_{IS} values range from -0.199 to 0.104, with an overall G_{IS} -0.058 across all populations (Table 2).

An individual pairwise Bruvo distance and a Lynch distance were calculated to estimate pairwise genetic distance between individuals. A PCA was performed using the pairwise table calculated for each distance. The first component of the Bruvo distance explained 21.5% of the variation and the second component explained another 15% of the variation (Figure 2). The Bruvo distance loosely clustered the northern AMM populations (1–5) separate from the remaining populations with some overlap. Individuals from Population 8 displayed the broadest distribution with overlap of all other populations. A second PCA generated using the Lynch distance method of calculating genetic distance did not produce any discernable geographic or population clustering (Supplemental Figure 1). A Mantel's test using the Bruvo distance was run to assess the relationship of the genetic and geographic distances between the individuals. The analysis was run with 1000 replicates. The resulting r^2 value was 0.181 (p=0.001).

F_{ST}, G_{ST}, and Jost's D pairwise values were calculated for each of the populations using a partial heterozygote model for inferred allele frequency (Clark and Jasieniuk 2011). The three tests were run to provide cross validation of the results of each method and all three methods yielded similar results. The ranges for the F_{ST} values are 0.011 to 0.126 (Supplemental Table 1). The G_{ST} values ranged from 0.006 to 0.247 (Supplemental Table 2). The Jost's D values ranged from 0.034 to 0.699 (Supplemental Table 3). The very small size of Populations 4 and 9 can influence the calculation of differentiation, generally increasing estimated fixation.

Population Structure

Analysis of the output of the Bayesian clustering program STRUCTURE using the Evanno method strongly suggests four clusters (K=4) but also showed support for two clusters (K=2) and seven clusters (K=7) (Figure 3). The K=7 analysis made little biological sense and the output had most

Table 2. Genetic diversity of *Geum radiatum* as revealed by eight microsatellite markers.

Population	N	A	Rs	Ae	H_O	F_{IS}	G_{IS}
Population 1	20	51	6.375	3.436	0.675	-0.343	-0.038
Population 2	20	56	6.375	2.638	0.597	-0.41	-0.199
Population 3	19	61	7.625	2.873	0.605	-0.356	-0.127
Population 4	7	37	4.625	3.394	0.783	-0.39	0.104
Population 5	19	30	7	3.611	0.687	-0.316	0.051
Population 6	18	42	5.25	3.104	0.608	-0.397	-0.009
Population 7	14	41	5.125	2.924	0.628	-0.431	-0.097
Population 8	81	79	9.875	3.372	0.645	-0.404	-0.036
Population 9	3	19	2.571	2.138	0.685	-0.231	0.1
Population 10	12	34	4.25	2.467	0.537	-0.543	-0.107
Population 11	18	48	5.75	3.171	0.665	-0.382	-0.041
Population 12	23	38	4.75	2.709	0.573	-0.517	-0.135
Population 13	27	39	4.875	2.487	0.489	-0.472	-0.178
Population 14	29	39	4.875	2.744	0.597	-0.592	-0.101
Mean	22.143	43.86	5.666	2.933	0.627	-0.413	-0.058

Note: N = number of individuals genotyped per population, A = number alleles per population, Rs = allelic richness per population, Ae = effective number of alleles per population, H_O = observed heterozygosity per population, F_{IS} = Wright's inbreeding coefficient, G_{IS} = Nei's inbreeding coefficient

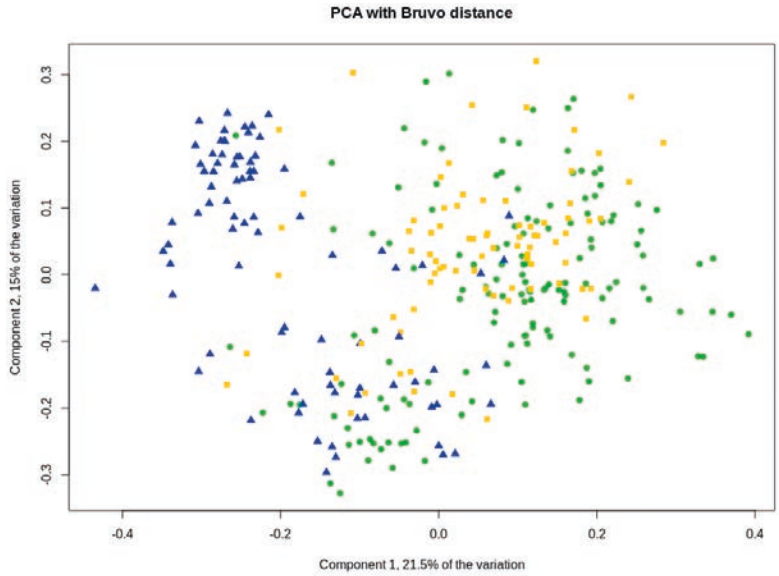

Figure 2. PCA using the Bruvno genetic distance calculation. The blue individuals are members of the geographically close and geological similar AMM populations (1-5). The yellow individuals are members of population 8. The green individuals are members of remaining populations. Component 1 explains 21.5% of the variation and component 2 explains 15% of the variation.

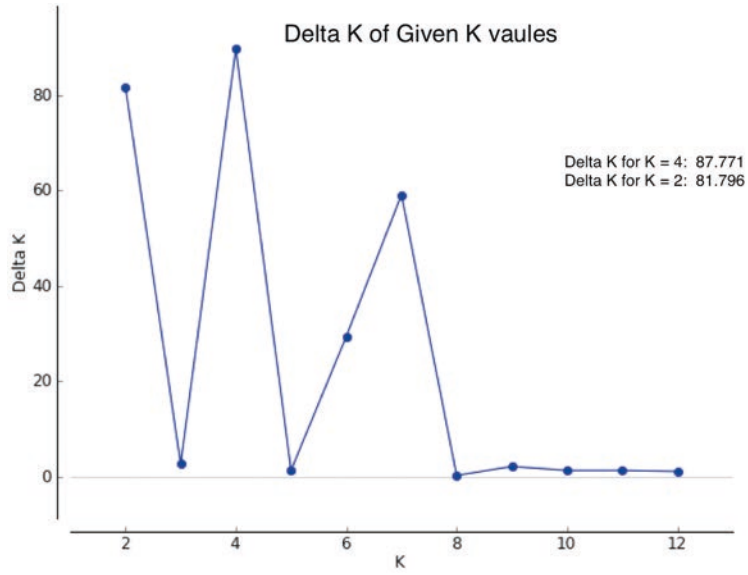

Figure 3. The delta K values for each possible K=1-13, using the Evanno correction. K=4 has the strongest support with a delta K of 87.771 and K=2 has the second strongest support with a delta of 81.796.

individuals heavily admixed between multiple genetic clusters; it also has the weakest delta K among putative K values calculated. As K value increases, southern populations are sub-divided into unique genetic clusters (Figure 4). A map displaying the Q-matrix values from STRUCTURE as pie graphs allows visualization of population structure in relation to geographic distribution (Figure 5). A second approach to visualize population structure on a landscape was employed in TESS to infer the ancestor coefficients using current population locality data and the Q-matrix for K=4 generated in STRUCTURE (Figure 6). Population 8 had the largest number of samples collected (N=81), because of its known history of past augmentations (correspondence with USFS). Based on the STRUCTURE output this population contained individuals of all four genetic clusters and was further analyzed to identify transplanted individuals and admixed descendants with parentage from both native and transplanted genotypes, from here on referred to as anthropogenically admixed individuals.

Hybrid Index and Augmentation Analysis

Individuals from Population 8 that contained multi-locus genotypes (both pure and admixed) corresponding to Population 1 (known source of transplants) were removed and the Q-matrix and ancestor coefficients were remapped to show the effect on clustering (Supplemental Figure 2). The only discernable change was the increases in intensity of the clusters within Population 8. To further Examine Population 8 individuals, a hybrid index was run. Four of the five subpopulations within Population 8 were augmented with transplanted plants and an isolated subpopulation (8e) was known to not be augmented because of the difficulty in accessing the site. (correspondence with USFS). The analysis revealed 14 plants with strong affinity for Population 1, the original source population (referred to as transplants from here forward). The hybrid index also suggested 13 more individuals consistent with having an anthropogenically admixed origin (Table 3). These results were mapped showing the geographic relationships of the anthropogenically admixed individuals, transplanted individuals and native individuals within Population 8 (Figure 7).

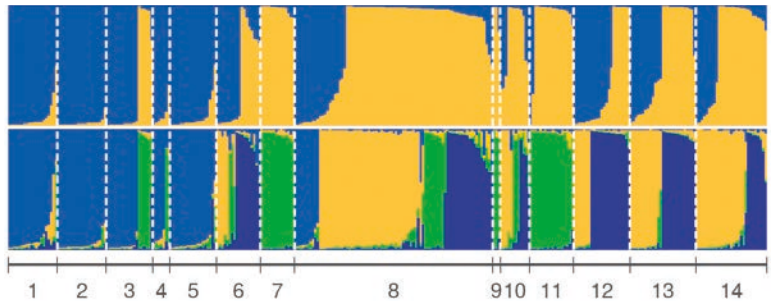

Figure 4. Bar plot output of analysis from STRUCTURE with both K=2 and K=4. K=4 clusters are blue = northern AMM populations, green = eastern affinity cluster, purple = western affinity cluster, and yellow = central Appalachian high peak cluster. The populations are separated by a white dotted line and are numbered according to the population labels in Figure 1.

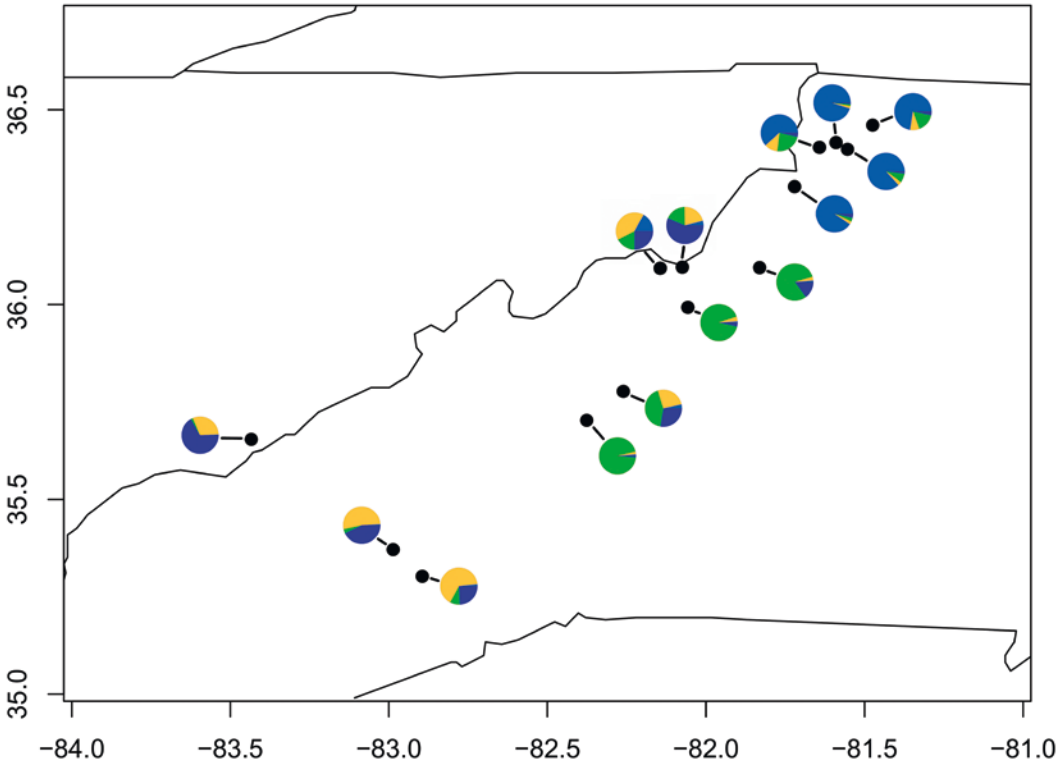

Figure 5. Map geographically displaying the Q-marix from the K=4 STRUCTURE analysis represented as pie graphs. The colors correspond to Figure 4.

DISCUSSION

Geum radiatum has a complicated genetic structure due to past conservation actions, restrictive habitat requirements, and a natural history that likely includes a past distribution in widespread alpine and tundra habitat. The Amphibolite Mountains Macrosite (AMM) populations (1–5) have little differentiation, even over multiple mountain tops, likely representing a meta-population. Population 8, in large part due to past augmentations, now represents the most diverse population

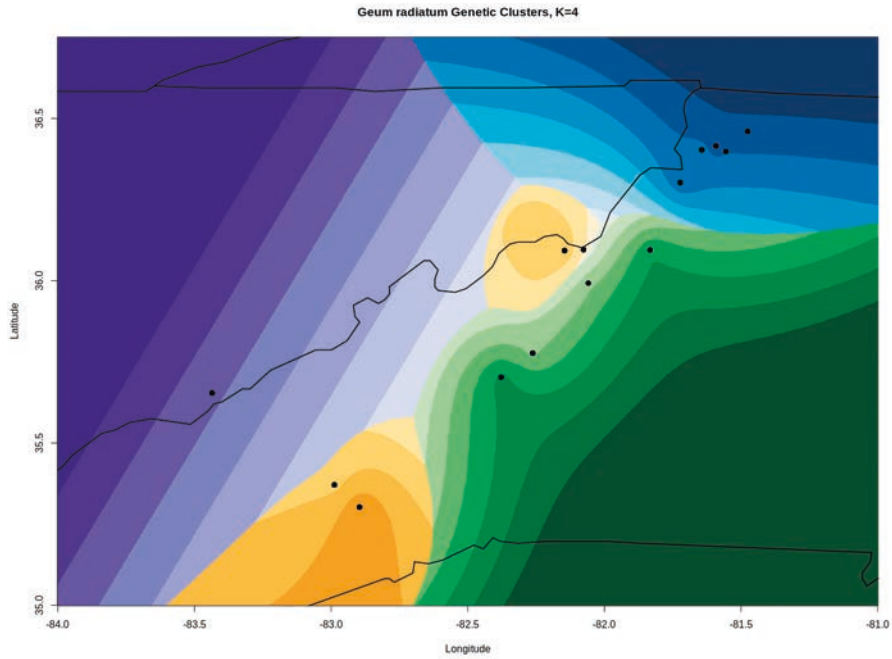

Figure 6. Map geographical displaying the predicted ancestor coefficients from the K=4 STRUCTURE analysis. The colors correspond to Figure 4.

Table 3. Hybrid individuals identified in GENODIVE

Subpopulation	h-value
8a	0.524
8a	0.51
8a	0.473
8a	0.428
8a	0.444
8b	0.571
8b	0.455
8c	0.59
8c	0.57
8c	0.513
8d	0.537
8d	0.486
8d	0.595

Note: h-vaule = percent identity that the individual shares with Subpopulation 8e. Subpopulations location is shown in Fig 1.

and contains all the genetic groups that where identified by the STRUCTURE analysis. The transplanted individuals in Population 8 have survived for ~25 years and have yielded putative anthropogenically admixed descendants.

Genetic Diversity

Geum radiatum has considerably higher genetic diversity than previously reported based on four allozyme loci (Godt et al. 1996). This is likely due to the higher mutation rate of microsatellite loci, which have been shown experimentally to contain significantly more diversity then allozyme loci (Estoup et al. 1998). *Geum radiatum* is a stable hexaploid (2n=6x=42), as is much of the genus *Geum*, which could also influence the amount of diversity maintained by neutral (microsatellite) vs coding (allozyme) loci, especially when inbreeding is occurring (Husband and Schemske 1997, Rausch and Morgan 2007). It has been shown that polyploids lose coding sites more quickly than non-coding sites (Liu et al., 1998). The long-life span of *G. radiatum* (Ulrey et al. 2016) allows individuals to survive indefinitely via clonality allowing populations to maintain higher diversity at neutral loci. The number of effective alleles was low in comparison to number of alleles observed, with an average of 2.66 effective alleles per locus. Many alleles were only found in low frequency, likely due to incomplete fixation or may have arisen as novel mutations in long-lived individuals.

Heterozygosity statistics and measures of Hardy-Weinberg equilibrium must be interpreted cautiously in *G. radiatum* because it is a polyploid. A heterozygous locus should be more common in

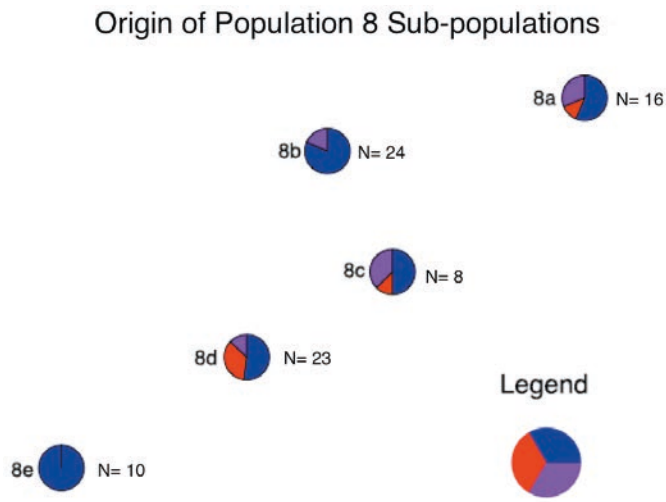

Figure 7. Map of the anthropogenically admixed individuals (Purple), transplanted individuals (Red) and native individuals (Blue) within the subpopulations of Population 8 as determined by GENODIVE hybrid index. The locations of the pie charts are geo-referenced.

a hexaploid than in a diploid species, with the increase in the possible number of alleles per locus. It should also take more extended periods of inbreeding to create homozygous individuals where all six genomic locations carry the same allele. Given the observed heterozygosities, it is not surprising that all populations have negative inbreeding coefficients for both F_{IS} and G_{IS} (Table 2), except for Population 9 with a positive G_{IS}, likely caused by its small population size (N=3). Our results suggest there is genetic diversity both within and between populations of *G. radiatum,* and that the species has not been reduced to a handful of genetic clones that are fixed across all loci.

Genetic Structure

Different analyses across the populations of *G. radiatum* show that populations are differentiated from each other, but how populations are structured changes from north to south across the range. Biological structure was likely caused by the lack of long-distance dispersal of pollen or seeds and the past geographic histories. The genetic structure of populations of *G. radiatum* indicates geographical division generally into a northern AMM group, a central group between Boone, NC and Asheville NC, and a southern group south of Asheville, NC. The northern AMM populations (1–5) consistently show the strongest relationship with genetic distance, Bayesian clustering, and F_{ST} values reliably grouping these populations together. The southern and middle parts of the range (Populations 6–14) contain three different genetic clusters and greater differentiation between populations when compared to the northern AMM populations. These populations also tended to cluster in more complex ways with an east, west, and central breakdown (see below).

The Bruvo genetic distance PCA clustered the populations with some resolution of groups but has significant overlap in the ranges of different populations. Population 8 displayed the largest spread of ordination of any population, suggesting it contains the greatest diversity, but this could be influenced by the increased sample size and history of augmentation. The Mantel's test, with a low r^2 value, suggests that the genetic diversity is not correlated with a simple linear geographic relationship, indicating a more complex isolation pattern than by distance alone. In some cases, for example Populations 11 and 10 found in close geographic proximity, have distinctly different genotypes with strong affinities to populations at greater distances (Figure 3).

Each approach at measuring population differentiation, pairwise F_{ST}, G_{ST}, and Jost's D, follow similar trends in the species, where the strongest relationships seems to be between populations

that fall into the same geographic region. Oddly, the closest geographic populations do not always seem to be the most related, corroborating the results shown by the PCA approach and the Mantel's test (Figure 2). Based on F_{ST} values, the northern AMM populations appear to be the most connected by gene flow. For example, all pairs of populations north of Population 6 have F_{ST} values less than 0.1, suggesting that there is more natural gene flow between the close peaks or a more recent separation in the northern part of the range.

The northern AMM populations (1–5) have F_{ST} values that are similar to subpopulation F_{ST} values of Population 8. The F_{ST} values among subpopulations of Population 8 range from 0.019 to 0.097, with a mean of 0.045, which is comparable to the northern AMM populations values (Supplemental Table 4). This suggests the northern AMM populations are as interconnected via gene flow as the subpopulations of 8, with similar levels of differentiation. If Population 8 is to be managed as one population, perhaps the northern AMM populations should also be managed as a single interconnected population as well.

The Bayesian STRUCTURE analysis and Evanno correction's delta k values suggest 2, 4, or 7 genetic clusters. The results of the K=4 indicates similar groups to the hierarchical F_{ST} pairings and the Bruvo distance PCA groups for the northern populations (Figure 4). The middle populations (6, 7, 9, 10, and 11) cluster and contain three different genetic groups. The southern populations (12–14) cluster and contain two different genetic groups. A map of the ancestor coefficients from the STRUCTURE analysis suggests the affinities of these groups have a geographic origin (Figure 6). The blue cluster is strongly associated with the northern populations and is only found within the northern AMM populations with the exception of the augmented plants within Population 8. The yellow cluster has a central and southern affinity and may represent the ancestral high peak genetic group. The purple cluster has a western affinity and the green cluster has an eastern affinity, each becoming more common within populations moving father west or east. These ancestral coefficients have even stronger geographic affinities, when the augmented plants in Population 8 are removed from the analysis (Supplemental Figure 2). These four different genetic clusters may represent ancient genetic partitioning from the Pleistocene when true above timberline alpine communities and permafrost existed more commonly in the Southern Appalachians (Delcourt and Delcourt, 1998).

The directional affinities of the genetic clusters may represent the remnants of diversity from a time when *G. radiatum* was more widespread with multiple large interbreeding populations. At the end of the Pleistocene, when the earth's climate was warming, the populations of *G. radiatum* appear to have retreated up mountains to the tops of the highest peaks where they became stranded on cliff faces. In the Southern Appalachians there is a fairly narrow band of peaks that reach over 1500 m, which is widest south of Asheville, NC, where the Smoky Mountains and the Balsam Mountain have multiple high peaks around the same latitude. The width of the high Appalachian peaks narrows to a single mountain in northwest NC. The data supports the population differentiation seen today is a geographically condensed relic of past population structure where distinct genetic populations that were once geographically separated by great distance retreated into the only remaining suitable habitat and are now close neighbors with their once distant relatives.

Hybrid Index Analysis

Subpopulation 8e from Population 8 was thought to be free of augmented plants and did not contain any admixed genotypes or transplants from northern populations in the STRUCTURE analysis. These individuals were used as the native genotype for the hybrid index analysis in GENODIVE. A group of individuals from Population 1 were used for augmentation of Population 8 and therefore the source for the other parental genotype in the hybrid index analysis. This analysis suggested several putative anthropogenically admixed individuals among the subpopulations that were augmented (Figure 7). All subpopulations where admixed individuals were identified using the hybrid index also have transplanted individuals, except Subpopulation 8b. The admixed individuals identified in Subpopulation 8b either represents gene flow from transplanted individuals at another

subpopulation or that transplanted individuals were missed in the collections from that subpopulation. These admixed individuals represent a unique anthropogenically mediated gene flow event that has increased the diversity within Population 8 and lowered its pairwise differentiation values (F_{ST}) when compared to every other population of *G. radiatum* (Supplemental Table 1). Population 8 has the least consistent population assignment of any of the STRUCTURE clusters and the lowest amount of clustering in the Bruvo genetic distance PCA. The past augmentations of four subpopulations within Population 8 were successful and F1 admixed individuals have been formed. This result demonstrates that artificial gene flow by augmentation can be a successful strategy for conservation. Due to this augmentation event, Population 8 serendipitously contains all of the genetic clusters identified in this analysis, with the highest amount of diversity and the least differentiation from other populations, perhaps cementing the importance of Population 8 to the long-term conservation of *G. radiatum*.

Conservation Implication

Geum radiatum contains more genetic diversity and population structure then previously reported; like all species preserving and protecting every single extant individual is the most reliable strategy to maintain diversity in the species as a whole. *Geum radiatum* is long lived and has a rhizomatous growth pattern, but also has exceptionally low seedling recruitment especially in the more southern and imperiled populations (Ulrey et al. 2016). Considering these factors, the loss of a single adult *G. radiatum* plant may take hundreds of years to regenerate, especially in the face of modern climate change, and the loss of a single plant could mean the loss of a unique genetic make-up for that population or the species as a whole. The average population size of *G. radiatum* is very low with some populations only containing three individuals, where each individual may have arisen millennia ago. *Geum radiatum* retains many different genetically unique individuals and has not been reduced to a handful of clones. To ensure long-term species survival, this genetic diversity should be maintained by continuing current protections and strictly limiting recreational development around populations.

If the end goal of conservation strategies is to increase local genetic diversity, then the past augmentations of *G. radiatum* in Population 8 have been successful. The authors highlight that the risk of augmenting populations with unknown genotypes can result in outbreeding depressions by breaking locally adapted traits (Storfer 1999) or the introduction of non-native invasive genotype (Montalvo et al. 2007), can make genetic augmentation a risky conservation action. The augmentations conducted in the early 1990s increased diversity but it is still unknown if they increased the chances of the long-term survival of the species. Population 8 has a complex genetic structure formed by its geographic location at the center of the range on the highest peaks, the large size of the population, and its organization of connected but distinct cliff-faces that form the greater metapopulation. The anthropogenic gene flow event examined in this study may have completed the natural processes of geographic condensation to higher peaks likely caused by climate change since the last glacial maximum. This process has been occurring naturally and has migrated the species to higher elevational points where the genetic diversity has been consolidated. The augmentation of Population 8 may be beneficial to the long-term genetic survival of population and the species as a whole.

The two most important targets for the conservation of *G. radiatum* are the AMM populations (1–5) that are genetically unique and are interbreeding between mountain tops and Population 8 which has five subpopulations and now contains the best overall representation of the genetic diversity within *G. radiatum* containing all 4 of the genetic clusters identified in the species.

ACKNOWLEDGMENTS

This research was funded using startup funds granted to MCE by Appalachian State University and the Department of Biology. The authors also acknowledge financial support from the Office of Student Research and the Graduate School at Appalachian State University. The authors thank

Dale Suiter with USFWS, Lesley Starke with NCDA&CS Plant Conservation Program, North Carolina State Parks, Great Smoky Mountains National Park, North Carolina Game Lands, and The Nature Conservancy for collection permits. The authors also thank James Wise for help with field collections and DNA extractions. The authors would also like to thank an anonymous reviewer.

LITERATURE CITED

Arens, P., W. Durka, J. Wernke Lenting, and M. Smulders. 2004. Isolation and characterization of microsatellite loci in *Geum urbanum* (Rosaceae) and their transferability within the genus *Geum*. Molec. Ecol. Notes.4:209–212.

Becker, R., A. Wilks, R. Brownrigg, T. Minka, and A. Deckmyn. 2016. Package maps: Draw Geographical Maps. (**https://cran.r-project.org/web/packages/maps/index.html**, 05 May 2017).

Bruvo, R., N.K. Michiels, T.G. D'Souza, and H. Schulenburg. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol. Ecol. 13:2101–2106.

Buza, L., A. Young, and P. Thrall. 2000. Genetic erosion, inbreeding and reduced fitness in fragmented populations of the endangered tetraploid pea *Swainsona recta*. Biol. Conservation 93:177–186.

Clark, L.V. and M. Jasieniuk. 2011. polysat: an R package for polyploid microsatellite analysis. Mol. Ecol. Resources 11:562–566.

Cogbill, C.V., P.S. White, and S.K. Wiser. 1997. Predicting treeline elevation in the Southern Appalachians. Castanea 62:137–146.

De Silva, H.N., A.J. Hall, E. Rikkerink, M.A. McNeilage, and L.G. Fraser. 2005. Estimation of allele frequencies in polyploids under certain patterns of inheritance. Heredity 95:327–334.

Delcourt, P.A. and H.R. Delcourt. 1998. Paleoecological insights on conservation of biodiversity: a focus species, ecosystems, and landscapes. Ecol. Applic 8:921–934.

Earl, D.A. and B.M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361.

Estoup, A., F. Rousset, Y. Michalakis, J. Cornuet, M. Adriamanga, and R. Guyomard. 1998. Comparative analysis of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). Mol. Ecol. 7:339–353.

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

Falush, D., M. Stephens, and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567.

Francis, R.M. 2017. Pophelper: an R package and web app to analyse and visualize population structure. Mol. Ecol. Resources 17:27–32.

Gajewski, W. 1957. A cytogenetic study on the genus *Geum* L. Polskie Towarzystwo Botaniczene 4:3–414.

Gerritsen, H. and M.H. Gerritsen. Package mapplots. (**https://cran.r-project.org/web/packages/mapplots/index.html**, 05 May 2017).

Glendinning, D.R. 1987. Some aspects of autotetraploid population dynamics. Theor. Appl. Genet. 78:233–242.

Godt, M.J.W., B.R. Johnson, and J. Hamrick. 1996. Genetic diversity and population size in four rare Southern Appalachian plant species. Conservation Biology 10:796–805.

Gompert, Z. and C.A. Buerkle. 2009. A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. Mol. Ecol. 18:1207–1224.

Gray, A. 1889. Scientific papers of Asa Gray. Houghton, Mifflin and Company, Boston, Massachusetts.

Husband, B.C. and D.W. Schemske. 1997. The effect of inbreeding in diploid and tetraploid populations of *Epilobium angustifolium* (Onagraceae): Implications for the genetic basis of inbreeding depression. Evolution 51:737–746.

Hamann, E., H. Kesselring, J. Stöcklin, and G.F. Armbruster. 2014. Novel microsatellite markers for the high-alpine *Geum reptans* (Rosaceae). Appl. Plant Sci. 2:1400021.

Jay, F., E. Durand, and R. All. The POPS web page. (**http://membres-timc.imag.fr/Olivier.Francois/pops.html**, 05 May 2017).

Jenkins, A.P. 2011. The vascular flora of Three Top Mountain Game Land Preserve, Ashe County, North Carolina. M.S. thesis, Appalachian State University, Boone, North Carolina.

Johnson, B.R. 1995. The ecology and restoration of a high montane rare plant community. Ph.D. dissertation, University of Georgia, Athens, Georgia.

Jost, L. 2008. GST and its relatives do not measure differentiation. Molecular Ecology 17:4015–4026.

Liu, B., J.M. Vega, and M. Feldman. 1998. Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences. Genome 41:535–542.

Lynch, M. 1990. The similarity index and DNA fingerprinting. Molecular Biology and Evolution 7:478–484.

Medford, D. 2001. The detection of morphological variation across time in two Roan Mountain endemics: *Geum radiatum* and *Houstonia montana*. M.S. thesis, East Tennessee State University, Johnson City, Tennessee.

Meirmans, P.G. and P.H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes 4:792–794.

Montalvo, A.M., S.L. Williams, K.J. Rice, S.L. Buchmann, C. Cory, S.N. Handel, G.P. Nabhan, R. Primack, and R.H. Robichaux. 1997. Restoration biology: a population biology perspective. Restoration Ecology 5:277–290.

Murdock, N.A. 1993. Recovery plan for spreading avens (*Geum radiatum*) Rafinesque. (**https://ecos.fws.gov/docs/recovery_plan/930428.pdf**, 05 May 2017).

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321–3323.

Paterson, I.G. and M. Snyder. 1999. Genetic evidence supporting the taxonomy of *Geum peckii* (Rosaceae) and *G. radiatum* as separate species. Rhodora 101:325–340.

Peet, R.K., J.D. Fridley, and J.M. Gramling. 2003. Variation in species richness and species pool size across a pH gradient in forests of the Southern Blue Ridge Mountains. Folia Geobotanica 38:391–401.

Poindexter, D.B. and Z.E. Murrell. 2008. Vascular flora of Mount Jefferson State Natural Area and environs, Ashe County, North Carolina. Castanea 73:283–327.

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

Rausch, J.H. and M.T. Morgan. 2007. The effect of self-fertilization, inbreeding depression, and population size on autopolyploid establishment. Evolution 59:1867–1875.

Rohrer, J.R. 2014. *Geum*. In: Flora of North America Editorial Committee, eds. 1993+. Flora of North America North of Mexico. 12+ vols. New York and Oxford. Vol. 9, pp. 58–69.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18:233–234.

Storfer, A. 1999. Gene flow and endangered species translocations: a topic revisited. Biological Conservation 87:173–180.

Tucker, G. 1972. The vascular flora of Bluff Mountain, Ashe County, North Carolina. Castanea 37:2–26.

Ulrey, C., P.F. Quintana-Ascencio, G. Kauffman, A.B. Smith, and E.S. Menges. 2016. Life at the top: Long-term demography, microclimatic refugia, and responses to climate change for a high-elevation Southern Appalachian endemic plant. Biological Conservation 200:80–92.

Weakley, A. 2015. Flora of the southern and mid-Atlantic states, working draft of May 2015. University of North Carolina Herbarium, Chapel Hill, North Carolina.

Wiser, S.K. 1994. High-elevation cliffs and outcrops of the Southern Appalachians: vascular plants and biogeography. Castanea 59:85–116.

Wiser, S.K. 1998. Comparison of Southern Appalachian high elevation outcrop plant communities with their Northern Appalachian counterparts. Journal of Biogeography 25:501–513.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138.

Wright, S. 1965. The Interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395–420.