Conservation Genetics
https://doi.org/10.1007/s10592-021-01404-8

RESEARCH ARTICLE

Landscape and stocking effects on population genetics of Tennessee Brook Trout

John S. Hargrove[1,6] · David C. Kazyak[2] · Barbara A. Lubinski[2] · Karli M. Rogers[2] · Olivia K. Bowers[1] · Kurt A. Fesenmyer[3] · Jim W. Habera[4] · Jason Henegar[5]

Received: 19 January 2021 / Accepted: 23 September 2021

Abstract

Throughout their range, Brook Trout (*Salvelinus fontinalis*) occupy thousands of disjunct drainages with varying levels of disturbance, which presents substantial challenges for conservation. Within the southern Appalachian Mountains, fragmentation and genetic drift have been identified as key threats to the genetic diversity of the Brook Trout populations. In addition, extensive historic stocking of domestic lineages of Brook Trout to augment fisheries may have eroded endemic diversity and impacted locally adapted populations. We used 12 microsatellite loci to describe patterns of genetic diversity within 108 populations of wild Brook Trout from Tennessee and used linear models to explore the impacts of land use, drainage area, and hatchery stockings on metrics of genetic diversity, effective population size, and hatchery introgression. We found levels of within-population diversity varied widely, although many populations showed very limited diversity. The extent of hatchery introgression also varied across the landscape, with some populations showing high affinity to hatchery lineages and others appearing to retain their endemic character. However, we found relatively weak relationships between genetic metrics and landscape characteristics, suggesting that contemporary landscape variables are not strongly related to observed patterns of genetic diversity. We consider this result to reflect both the complex history of these populations and the challenges associated with accurately defining drainages for each population. Our study highlights the importance of genetic data to guide management decisions, as complex processes interact to shape the genetic structure of populations and make it difficult to infer the status of unsampled populations.

Keywords *Salvelinus fontinalis* · Legacy effects · Hatchery introgression · Effective population size · Landscape genetics

Introduction

Conservation of freshwater fishes occupying headwater streams is a complicated task as fish community composition and structure are influenced by a complex set of interacting biotic (e.g., competition and predation) and abiotic factors (physical and chemical attributes; Greswell and Vondracek 2010). Adding to these challenges, anthropogenic changes to the aquatic and surrounding terrestrial landscape can elicit ecological changes that vary across both spatial and temporal scales (Fausch et al. 2002). Theory predicts that smaller populations in more isolated habitats are subject to greater risk of extinction via demographic and environmental stochasticity (Boyce 1992) and land use changes such as logging and agriculture can decrease available habitat (Lunn et al. 2017), reduce population sizes, and exacerbate the risk of extirpation. Efforts to relate habitat characteristics (e.g., elevation, temperature, stream flow) and land use

John S. Hargrove and David C. Kazyak contributed equally to this work.

✉ John S. Hargrove
john.hargrove@idfg.idaho.gov

✉ David C. Kazyak
dkazyak@usgs.gov

1 Tennessee Tech University, Cookeville, TN, USA

2 U.S. Geological Survey, Eastern Ecological Science Center at the Leetown Research Laboratory, Kearneysville, WV, USA

3 Trout Unlimited, Boise, ID, USA

4 Tennessee Wildlife Resources Agency, Morristown, TN, USA

5 Tennessee Wildlife Resources Agency, Nashville, TN, USA

6 Present Address: Pacific States Marine Fisheries Commission, Eagle, ID, USA

Published online: 03 December 2021

effects on fish communities commonly involve dissecting landscapes into homogenous 'patches' that differ from surrounding areas, and testing for associations between biological (e.g., abundance, distribution, etc.) and landscape variables of interest (Pringle et al. 1988). Studies applying a patch-dynamics framework have identified that coldwater stream fish persistence can be positively linked to patch area (e.g., Peterson et al. 2014) and the percentage of forested cover within a patch (i.e., minimal habitat degradation; Stranko et al. 2005). In addition to identifying the impacts of contemporary changes to the landscape (i.e., recent development or deforestation), landscape studies have shown that aquatic assemblages can be negatively affected by changes to the environment that occurred in the distant past (i.e., legacy effects; Harding et al. 1998). Identifying predictable relationships among habitat characteristics, land use practices, and demographic responses can help promote conservation efforts by identifying populations that are most likely to benefit from restoration efforts (EBTJV 2011).

Additional threats to freshwater fish may arise from ecological interactions with introduced fishes in the form of predation, competition, hybridization, and disease transmission (see Gozlan et al. 2010 for review). While studies of fish introductions are commonly focused on species introduced outside of their native range (e.g., Hargrove et al. 2017), salmonids have been reared in hatcheries and stocked into populations of native conspecifics throughout much of North America for supplementation purposes (e.g., Horak 1995). Supplementation programs are commonly implemented to offset losses due to harvest, enhance recruitment, or overcome habitat limitations (Trushenski et al. 2010). However, successful reproduction between hatchery and wild populations (hereafter 'genetic introgression') can be problematic as genetic dissimilarities between stocks can negatively affect wild populations. For example, the hatchery environment can elicit phenotypic and genetic changes which can be maladapted for survival in the wild (e.g., Heath et al. 2003; Sundström et al. 2004; Le Luyer et al. 2017). As a result, genetic introgression between hatchery and wild stocks may negatively affect fitness, resiliency, and adaptive potential via the disruption of co-adapted gene complexes (Hallerman 2003; Naish et al. 2007), introduction of maladaptive phenotypes (Bolstad et al. 2017; Gossieaux et al. 2020) or deleterious mutations (Ferchaud et al. 2018), and increased susceptibility to disease (Currens et al. 1997).

Assessments of genetic structure and diversity across the landscape are particularly well-suited to evaluate the evolutionary impacts of habitat characteristics, land use patterns, and hatchery introgression on native fish populations. For instance, water quality impairment due to agricultural land use (e.g., nutrient loading or water turbidity) has been shown to be a strong predictor of both genetic diversity and species diversity among stream fishes in the state of Ohio (Blum et al. 2012). Additionally, genetic diversity and population structure in Brook Trout (*Salvelinus fontinalis*) are linked to habitat fragmentation and patch area (e.g., Kanno et al. 2015; Nathan et al. 2019), with population-level metrics of genetic diversity and the number of adults contributing to annual reproduction being positively correlated with the area of patches above in-stream barriers (Whiteley et al. 2013). Molecular markers have also been used to quantify introgression between hatchery and native populations (e.g., Ozerov et al. 2016; Sanz et al. 2006) and to assess variables that explain observed levels of hatchery introgression (e.g., Harbicht et al. 2014). In addition to identifying the impacts of historical actions on contemporary populations, cataloging levels of genetic diversity provides insights into adaptive potential and the ability for populations to respond to future environmental change (Wade et al. 2017). Combined, genetic assessments can inform several conservation and management related decisions including identifying at-risk populations, prioritizing restoration efforts, and identifying suitable donor populations for translocation purposes (Pavlova et al. 2017; Malone et al. 2018).

Brook Trout are native to lakes and headwater streams of eastern North America (Georgia to Maine). They are subject to extensive conservation efforts because of their substantial ecological, recreational, and cultural importance. Brook Trout have experienced significant declines throughout their range (Hudy et al. 2008) and populations at the southern end of their distribution (southern Appalachian Mountains) are of particular concern for several reasons. First, Brook Trout native to the southern Appalachians occupy only 20–30% of their historic range (Bivens 1984; Habera and Strange 1993) as a result of habitat degradation, overharvest, and interactions with introduced non-native salmonids (Bivens 1984; Habera and Moore 2005). Additionally, many populations of Brook Trout at the southern end of their distribution occur near thermal maximum limits and are predicted to become more fragmented as climates continue to warm (Flebbe et al. 2006). Lastly, hatchery-reared Brook Trout of non-native ancestry (primarily northeastern US stocks) were introduced throughout the southern Appalachian Mountains during the last century (Lennon 1967; Habera and Moore 2005; Kazyak et al. 2018). Various facets of Brook Trout life history and ecology have been studied, including the distribution of populations across the landscape (i.e., habitat requirements and land use impacts; Hudy et al. 2008) and the extent of hatchery introgression (Guffey 1998; McCracken et al. 1993; Kriegler et al. 1995; Hayes et al. 1996; Galbreath et al. 2001; Dunham et al. 2002; Seehorn 2004; Kazyak et al. 2018; Pregler et al. 2018; Weathers et al. 2019). However, no formal work has attempted to relate the relative influence of habitat characteristics, land use practices, and hatchery introgression on levels of genetic diversity in Brook Trout populations. Such work would not only

promote understanding of broad-scale patterns across the landscape but would also provide contemporary genetic data that could be used to direct near-term restoration and conservation efforts.

In support of ongoing Brook Trout conservation and management in the southern Appalachian Mountains, we described patterns of genetic diversity within 108 populations of wild Brook Trout from Tennessee and modeled the impacts of land use, drainage area, and hatchery stockings on metrics of genetic diversity, effective population size, and hatchery introgression. We sought to establish relationships between genetic parameters and habitat characteristics and land use patterns to understand if populations inhabiting larger drainages with intact landscapes differed systematically from those in smaller fragmented areas. Additionally, using reference populations from native and hatchery stocks, we quantified the extent of hatchery introgression across the landscape and tested if stocking records explain observed differences in genetic diversity and introgression. Combined, our study adds to a growing body of literature that seeks to examine how landscape characteristics and land use changes affect the evolutionary trajectory of fish populations of conservation concern.

Materials and methods

Sample collection

Brook Trout distribution surveys conducted by the Tennessee Wildlife Resources Agency (TWRA) during 2010–2015 indicated 108 extant populations in Tennessee outside Great Smoky Mountains National Park and these data were used to identify stream reaches for genetic sampling. Brook Trout were collected from these areas (Fig. 1) with AC backpack electrofishing gear during May–October of 2011–2017. Sample collection was dispersed throughout most of the existing Brook Trout distribution in each stream (range of occupied stream length: 0.1–8.0 km; mean 2.0 km), including areas upstream of any waterfalls present. Map coordinates were recorded for the sample reach start and end points. Target sample size was 30 fish per population and age-0 Brook Trout (typically < 100 mm TL) were avoided, if possible, to help minimize the inclusion of siblings. We examined metrics of genetic differentiation between sample sites (see below) to ensure that our sampling did not inadvertently sample two segments of the same population.

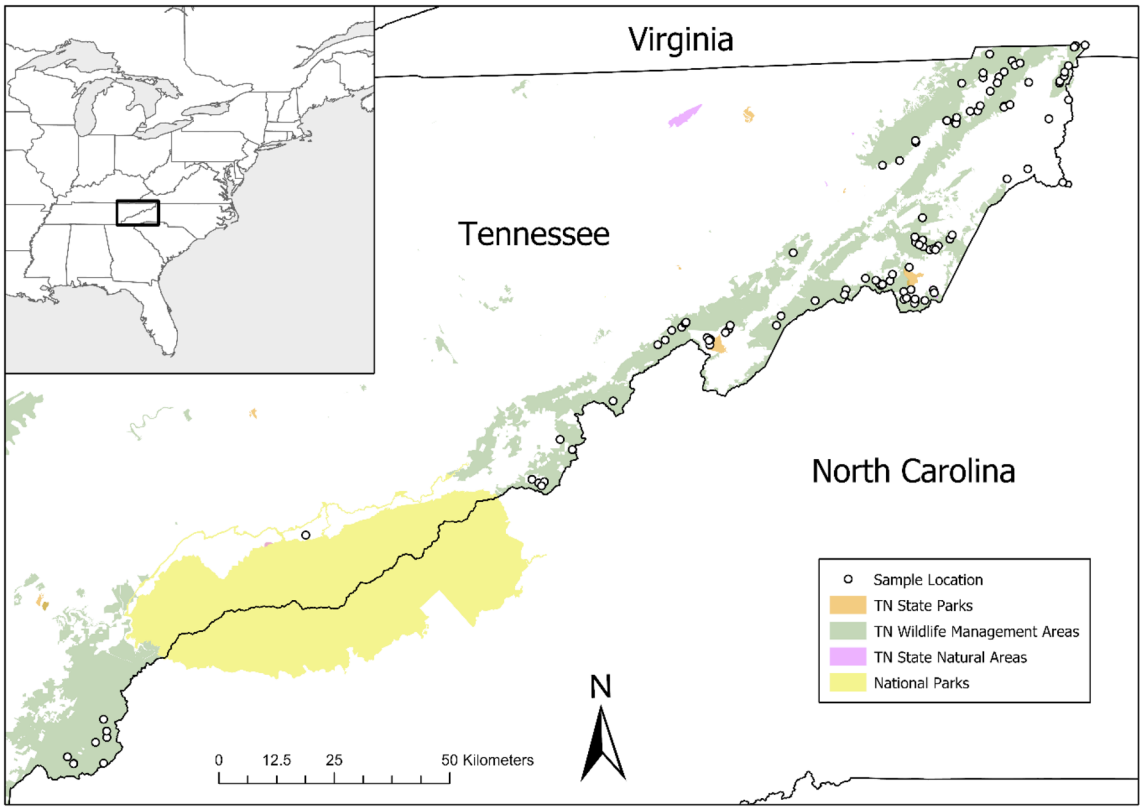

Fig. 1 Brook Trout (*Salvelinus fontinalis*) sample sites across the state of Tennessee used to survey genetic diversity and characterize landscape characteristics. Solid lines indicate state boundaries

Genetic samples (pelvic fin clips) were collected in the field, preserved with 95% ETOH in location-coded vials, and held in coolers with gel ice packs during completion of sampling. All fish were released alive following sample collection. Sample vials were then refrigerated within 12 h and held for at least one month before shipment to the analytical lab.

Laboratory protocols

Template DNA was extracted from an approximately 1 mm^2 fin clip excision using the E-Z 96 Tissue DNA plate extraction kit (Omega Bio-Tek, Norcross, GA) following manufacturers protocol. A negative control was included in each plate extraction. We quantified DNA concentrations using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA) read on a Synergy HTX Multi-Mode Microplate Reader (BioTek Instruments, Winooski, VT). All samples were normalized to a standard concentration of 20 ng/µl prior to PCR amplification.

We performed four multiplex PCR reactions to amplify 12 microsatellite loci (*Sfo*B52, *Sfo*C24, *Sfo*C28, *Sfo*C38, *Sfo*C79, *Sfo*C86, *Sfo*C88, *Sfo*C113, *Sfo*C115, *Sfo*C129, *Sfo*D75, and *Sfo*D91; see Kazyak et al. 2018 and King et al. 2012 for specific details on multiplex conditions and primer sequences). Each PCR reaction contained 1 µl of normalized DNA, 7.5 µl Qiagen Multiplex PCR Master Mix (Qiagen, Valencia, CA), 0.02–0.2 µM forward and reverse primers, and 5 µl sterile PCR H_2O, composing 15 µl reactions. The thermal cycling conditions for each reaction were: 95 °C for 15 min, followed by 35 cycles of 94 °C for 30 s, 58 °C for 90 s, and 72 °C for 90 s, completed by an extension step of 72 °C for 10 min. Each plate contained both a negative and positive control. PCR products were electrophoresed using an Applied Biosystems 3130xl genetic analyzer with Genescan 500 LIZ dye size standard (Applied Biosystems, Foster City, CA). Genemapper 4.0 software (Applied Biosystems) was used to bin, score, and output allelic data. All microsatellite scoring was automated and then checked by eye.

We screened genotypes for the presence of duplicate individuals using GenAlEx 6.503 (Peakall and Smouse 2006, 2012) and retained only unique genotypes for individuals that contained genetic data for at least 90% of the loci assayed.

Genetic metrics

Sampling stream populations of fishes may inadvertently sample family groups and inclusion of several full siblings in population genetic analyses may introduce bias (Hess et al. 2015; although see Waples and Anderson 2017). To minimize the inclusion of large family groups, we first assessed family structure among sample collections by estimating full sibship families as implemented in Colony v 2.0.5.0 (Jones and Wang 2010). Next, we generated within-population diversity statistics [observed heterozygosity (H_O), unbiased expected heterozygosity (H_E), and number of alleles (N_A)] for each collection using GenAlEx 6.503. Allelic richness was calculated using rarefaction techniques implemented in HP-Rare (Kalinowski 2005). We tested for Hardy–Weinberg equilibrium and linkage disequilibrium (LD) as calculated in Genepop v. 4.3 (Raymond and Rousset 1995) with the following parameters: dememorization = 1000, batches = 100, iterations per batch = 5000. Significance of HWE and LD tests were evaluated following a Bonferroni correction (Rice 1989). Estimates of effective population size (N_E) were generated using the LD based estimator implemented in NeEstimator v2 (Do et al. 2014). Estimates of N_E were made following a rare allele cutoff of 0.02 and confidence intervals were based on jackknife procedures. We computed pairwise estimates of genetic differentiation (F_{ST}) among sites using the algorithms implemented in Genepop with the following parameters (dememorization = 1000, batches = 100, iterations per batch = 5000). We scrutinized F_{ST} values among geographically proximate populations to screen for scenarios where the same population may have been sampled repeatedly.

Hatchery introgression

We used three complementary metrics to evaluate hatchery introgression following the methods of Kazyak et al. (2018). Briefly, we used reference collections of hatchery populations and drew comparisons with wild Brook Trout populations in Tennessee based on: (1) genetic distance, (2) discriminant analysis of principal components (DAPC; Jombart et al. 2010), and (3) genetic ancestry using Bayesian clustering analysis as implemented in the program STRUCTURE v 2.3.4 (Pritchard et al. 2000). We used three different forms of analysis as each technique differs in underlying assumptions, and combined they represent a weight-of-evidence approach to estimate hatchery introgression. We chose to use the mean of the three hatchery introgression metrics as a composite southern Appalachian Brook Trout (SABT) metric that varies on a scale from 0 to 1, with 1 representing minimal hatchery influence (highest likelihood of being a native population) and 0 representing a high degree of hatchery introgression. Details on the methods for each form of analysis and their associated correlations are described in detail below.

We compiled a reference baseline to assess hatchery introgression using genotypes from candidate hatchery strains (i.e., most likely hatchery source populations) and putatively native populations of Brook Trout from the southern Appalachian Mountains. The reference hatchery genotypes used in the current study were generated for a

previous assessment of hatchery introgression in North Carolina Brook Trout populations (Kazyak et al. 2018) and consisted of samples from 15 hatcheries (711 individuals) and represented many of the major lineages used for stocking in the eastern United States. For specifics on hatchery locations and Brook Trout strains reared at different facilities please reference Table 1 in Kazyak et al. (2018). In addition, collections of putatively native populations of Brook Trout from the Interior Basin were included for reference purposes. The set of genotypes representing native populations was also generated by Kazyak et al. (2018) and included 1595 individuals representing 61 collections from the Interior Basin in the southern Appalachian Mountains. Only a subset of the putatively native Brook Trout reference genotypes presented in Kazyak et al. (2018) were included in the current study, as several collections used by Kazyak et al. (2018) were from the current study in Tennessee. The general approach used to identify putatively natural populations involved a two-step process. First, putatively wild and known hatchery collections were compared with major phylogeographic assemblages previously identified by Stauffer and King (2014) using a neighbor-joining (NJ) tree. These comparisons were drawn using a distinct set of Brook Trout samples from the northeastern United States and the southern Appalachian Mountains. Phylogenetic analysis indicated all hatchery strains were genetically most similar to Brook Trout collections from the northern Atlantic slope (see Fig. 2 in Kazyak et al. 2018), and a major phylogenetic break was observed between hatchery collections and wild reference collections from the southern Atlantic slope and lower Interior Basin (Stauffer and King 2014; Kazyak et al. 2018). Second, wild collections from the Interior Basin were then compared against hatchery collections, and any populations that displayed appreciable frequencies of hatchery introgression were omitted.

Chord distance

We calculated the minimum genetic chord distance between each wild Brook Trout collection from Tennessee and the hatchery reference collections as a metric to evaluate the extent of hatchery introgression. Cavalli-Sforza chord distances were calculated using the *hierfstat* package (Goudet and Jombart 2015) in R (R Core Team 2020).

DAPC

We applied DAPC to the reference collections using the *adegenet* package in R (Jombart et al. 2008). Initially, the wild collections were used to fit the ordination and develop the discriminant functions (80 principal components and 2 discriminant functions were retained). Then, we used the 'predict.dapc()' function to assign each wild individual to either the hatchery or Interior Basin reference groups. The mean affinity of individuals to the Interior Basin reference collections was calculated to describe each wild collection.

STRUCTURE

First, we evaluated $k = 1$ through 10 using only the reference collections. For these runs, we used 5 iterations per k value with 300,000 burn-in steps and 300,000 recorded steps. The models included admixture but location information was not used as a prior probability. We identified an optimal k using Evanno's Δk (Evanno et al. 2005) calculated using STRUCTURE HARVESTER (Earl and vonHoldt 2012). Based on this screening, we chose to evaluate hatchery introgression based on a model that assumed 4 clusters ($k = 4$). Accordingly, we conducted additional STRUCTURE runs using the following parameters: 20 iterations of $k = 4$ with 300,000 burn-in steps followed by 300,000 recorded steps, with admixture, and no prior information based on collection location information. We used the Popflag = YES option to force STRUCTURE to focus on the differences between the reference collections and compute Q-scores for non-reference collections in that context. Model runs were aligned using CLUMPP with the LargeKGreedy algorithm and 2000 repeats (Jakobsson and Rosenberg 2007). We visually assessed the conformity of our reference collections to their expected clusters. Hatchery collections were clearly and consistently discriminated, while the three remaining clusters were present in the putatively native reference collections in varying proportions. The STRUCTURE scores we present herein reflect the sum of the Q-scores for the wild-type clusters, averaged across the individuals in each collection.

Drainage delineation

Tennessee Wildlife Resources Agency staff provided latitude and longitude for the downstream extent of each Brook Trout population. We used these point locations to delineate watersheds for each Brook Trout population using the Watershed Geoprocessing Service in ArcGIS 10.6 software (Esri Inc. 2011). The Watershed Geoprocessing Service determines the contributing area for input points using topographic layers derived from the 30-m National Elevation Dataset (U.S. Geological Survey 2020). We inspected each preliminary output watershed, adjusted the location of the downstream points where necessary (e.g., at stream confluences to ensure drainage delineation included the correct tributary), and generated final population boundaries. Previous work examining the relationship between genetic metrics and landscape characteristics for Brook Trout has utilized existing geospatial databases such as the Eastern Brook Trout Joint Venture (EBTJV) or similar patch layers (e.g., Nathan et al. 2020),

Table 1 Comparison of the performance of nested linear models for relating landscape characteristics and stocking history to Brook Trout population genetic metrics

Model	Model Structure	*df*	Hatchery Introgression			Rarefied Allelic Richness			Unbiased Expected Heterozygosity			Effective Population Size		
			R^2	AICc	ΔAIC	R^2	AICc	ΔAIC	R^2	AICc	ΔAIC	R^2	AICc	ΔAIC
1	Area * NonFrst * Stock	9	0.158	46.9	7.7	0.175	193.4	4.6	0.122	−99.1	6.6	**0.235**	**711.4307**	**0.0**
2	Area * NonFrst + Area * Stock + NonFrst * Stock	8	0.157	44.6	5.4	0.166	192.3	3.4	0.105	−99.7	6.1	0.146	717.691	6.3
3	Area * NonFrst + Area * Stock	7	0.155	42.5	3.3	0.159	190.8	2.0	0.101	−101.4	4.4	0.122	717.7398	6.3
4	Area * NonFrst + Area + Stock + NonFrst * Stock	7	0.134	45.1	5.9	0.144	192.2	3.3	0.100	−101.3	4.5	0.126	717.3942	6.0
5	Area + NonFrst + Area * Stock + NonFrst * Stock	7	0.137	44.7	5.5	0.161	190.7	1.8	0.102	−101.4	4.4	0.140	716.17	4.7
6	Area * NonFrst + Area + Stock	6	0.134	42.8	3.6	0.140	190.6	1.7	0.097	−103.0	2.8	0.097	717.8309	6.4
7	Area + NonFrst + Area * Stock	6	0.134	42.8	3.6	0.123	192.0	3.1	0.079	−101.6	4.2	0.118	716.0412	4.6
8	Area + NonFrst + Area + Stock + NonFrst * Stock	6	0.124	44.1	4.9	0.140	190.6	1.7	0.097	−103.0	2.8	0.121	715.8199	4.4
9	Area * NonFrst	5	0.028	52.8	13.6	0.021	198.3	9.4	0.013	−98.3	7.5	0.097	715.8668	4.4
10	Area * Stock	5	0.127	41.5	2.3	0.122	190.1	1.3	0.079	−103.6	2.2	0.116	714.2209	2.8
11	NonFrst * Stock	5	0.112	43.3	4.1	0.131	189.3	0.4	0.088	−104.3	1.5	0.014	722.4542	11.0
12	Area + NonFrst + Stock	5	0.121	42.1	2.9	0.111	191.0	2.2	0.078	−103.4	2.3	0.090	716.4309	5.0
13	Area + NonFrst	4	0.005	53.1	13.8	0.005	197.5	8.6	0.002	−99.5	6.2	0.089	714.5156	3.1
14	Area + Stock	4	0.118	40.3	1.1	0.110	189.1	0.2	0.077	−105.4	0.4	0.090	714.431	3.0
15	NonFrst + Stock	4	0.111	41.1	1.9	0.089	190.8	2.0	0.057	−103.8	2.0	0.007	720.9959	9.6
16	Area	3	0.001	51.3	12.1	0.001	195.8	6.9	0.002	−101.5	4.3	0.089	712.5186	1.1
17	NonFrst	3	0.004	51.0	11.8	0.004	195.5	6.7	0.000	−101.4	4.4	0.000	719.5002	8.1
18	Stock	3	**0.109**	**39.2**	**0.0**	**0.089**	**188.9**	**0.0**	**0.057**	**−105.8**	**0.0**	0.007	719.0055	7.6

Bold and underlined text indicates the model that was best supported by AICc for each genetic metric. Model terms were as follows: "Area" is the area (km^2) of the drainage associated with each collection, "NonFrst" is the proportion of each patch with non-forested land use, and "Stock" indicates if stocking was documented within the stream where each collection was obtained. Interactive effects are denoted with an asterisk, whereas simple effects are denoted with a plus symbol

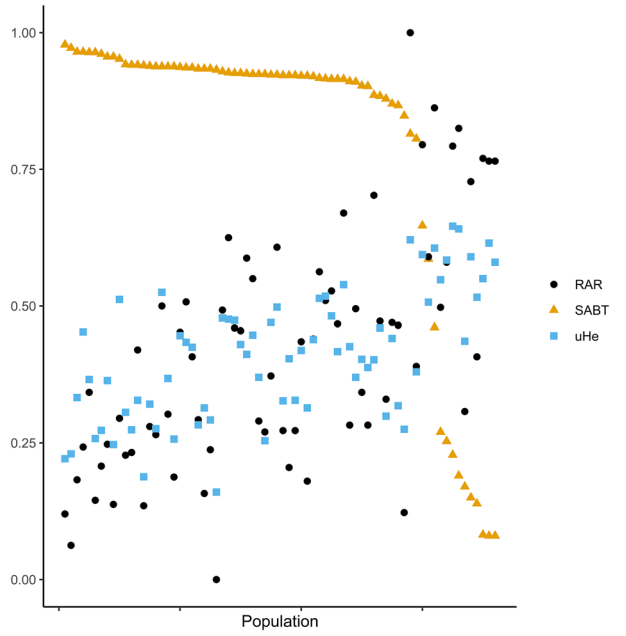

Fig. 2 Plots of unbiased levels of heterozygosity (uH_E), rarefied allelic richness (RAR), and a composite metric of extent of hatchery introgression (SABT) for collections with greater than 20 samples

and while we explored the use of the EBTJV layers for modeling efforts, we ultimately derived our own delineations (which we refer to as drainages) for several reasons. First, there were multiple Brook Trout populations in Tennessee for which EBJTV patches were either unavailable ($n = 13$) or where multiple distinct populations (based on F_{ST}) occurred within the same patch. Second, for occupied reaches below known barriers, EBTJV patches do not include the drainage area upstream of the barrier and thus may not fully characterize land use impacts in fragmented watersheds. In our study, there were several collections directly below known barriers in which the EBTJV patch layers encompassed a very small amount of habitat, and not the upstream drainage—we expected upstream landscape characteristics to influence individual populations in addition to the immediate habitat. Third, we used data provided by TWRA on the distribution of Brook Trout which was based on expert knowledge of the study system. As a result, we created a dataset which included attributes of the upstream catchment as opposed to habitat patch. We used stream length to characterize the amount of available habitat and drainage area delineation to capture watershed characteristics.

Land use patterns

Landscape metrics were calculated using ArcGIS Pro software (Esri, Inc. 2011) and Spatial Analyst tools (Esri). Each sampling point was assigned a drainage through a spatial join. The total area of forest, agriculture, urban development, and barren land within each drainage was calculated using the National Land Cover Database (NLCD) 2016 raster dataset and Tabulate Area tool (U.S. Geological Survey EROS 2020). The percentage of each land cover type was found by dividing the area of land cover type by the area of each sample's drainage. Mean, minimum and maximum elevation of drainages were calculated using Zonal Statistics tool and the 30-m resolution National Elevation Dataset (U.S. Geological Survey 2020). The length (kilometers) of roads and streams were calculated using the Summarize Within tool on the TIGER/Line roads layer and National Hydrography Dataset (NHD) flow line layer, respectively (U.S. Census Bureau 2020; U.S. Geological Survey 2020). The percentage of public lands was calculated by dividing the combined areas of Tennessee State Parks, Wildlife Management Areas (primarily the Cherokee National Forest), and State Natural Areas by the area of each drainage (Tennessee Wildlife Resources Agency).

Barriers to gene flow within each drainage (e.g., waterfalls and cascades) were identified using habitat descriptions of Bivens (1984) and TWRA field notes. However, we note that our drainages were broadly distributed across a high-gradient landscape and often had limited accessibility, and there are numerous features within many streams that may modulate connectivity to varying degrees. Many such features may reflect partial barriers to movement that may only be passable under specific flow conditions or by specific sizes of fish. Thus, while we report counts of potential barriers within each drainage, we consider these to be conservative estimates which reflect a subset of potential barriers across a complex landscape.

Principal components analysis (PCA)

Once data on land use patterns and drainage characteristics (collectively referred to here as landscape metrics) were assembled for each drainage, we performed a PCA on landscape metrics to identify highly correlated variables prior to subsequent modeling. Specifically, we were interested in identifying the maximum amount of variation explained across populations by a reduced set of variables. We applied the 'prcomp' function in R (R Core Team 2020) to: stocking (number of fish stocked), lower latitude, percent agriculture landcover, percent forested landcover, number of barriers present within a drainage, mean elevation, maximum elevation, percent of drainage publicly owned, length of roads, and drainage area. Variables were scaled to have unit variance prior to analysis, and PCA was performed on a covariance matrix.

Relative influence of hatchery introgression, habitat fragmentation, and land use patterns on genetic metrics

We used a series of linear fixed effects models in R (R Core Team 2020) to evaluate potential relationships between landscape characteristics, stocking history, and population genetic characteristics (A_R, H_O, N_E, and the composite score of hatchery introgression). For the models describing A_R, H_O, N_E, we restricted our analysis to collections with ≥ 20 samples ($n = 75$). We included all collections in the models describing the composite score of hatchery introgression, as our experience is that even small numbers of samples are typically effective for estimating this metric (the genetic differences between hatchery and endemic lineages is large and signals are typically consistent within populations; Kazyak et al. 2018). We initially defined a saturated model structure which included drainage area, the proportion of non-forested area in the drainage, stocking history both as a binary (stocked vs unstocked) and as a continuous variable (number of fish stocked), and all potential interactive effects between these variables, as PCA (above) determined that they captured the large amounts of variation across populations (Table 1). We compared the performance of the saturated model to all less complex nested models using Akaike's Information Criterion (AICc; Burnham and Anderson 2002), which balances the complexity of each model against its predictive ability. The best supported model for each genetic characteristic was identified using ΔAICc and model weights. We also report the multiple R^2 for each model to allow comparisons of the amount of variation explained. For the best supported model for each response, we present beta values to describe the direction and intensity of the relationships between the predictor variables and the response.

Results

A total of 2732 samples from 108 populations was collected and queued for DNA extraction. Of the 2732 samples collected, 2508 genotypes were retained based on acceptable levels of missing data (samples needed > 90% non-missing data) and screening for duplicates (i.e., samples could not match another sample due to either repeated sampling of the same individual in the field or laboratory error; Supplemental File 1). Of the 224 samples that did not pass quality control, 180 were duplicates and 44 were genotyped at < 90% of loci. Upon review of F_{ST} scores, we identified three collections near one another with low levels of differentiation (mean pairwise $F_{ST} = 0.06$; TN60, TN68, and TN77) relative to the overall dataset (mean pairwise $F_{ST} = 0.411$). Because there were no barriers between populations and sample sizes were quite small for two of the three collections ($n = 2$, 4), we retained one collection from this trio (TN68, which had a larger sample size) for subsequent hatchery introgression and landscape analysis (PCA, statistical modeling). All three populations were included in summaries of genetic diversity and effective population size. The final number of genotypes per population ranged from 2 to 50 (Supplemental File 2).

Genetic metrics

Sibship estimation showed that Brook Trout collections consisted primarily of unrelated individuals, but when full-sibling families were sampled, the families were small (Supplemental File 2). One population, (Rock Creek) represented a single full-sibling family and was excluded from subsequent analysis. Because full siblings were rare in our data set, we did not purge them from collections. We note that due to the limited diversity in some populations, the power to resolve family structure was somewhat limited.

Within-population diversity generally varied across all 108 collections. Observed heterozygosity was modest overall (mean $H_O = 0.40$; range 0.07–0.68), which was comparable to estimates of unbiased expected heterozygosity (mean $H_E = 0.40$; range 0.10–0.69) (Fig. 2; Supplemental File 2). The number of alleles per locus ranged from 1.25 to 5.83 (mean = 3.05), which was higher on average than estimates of rarefied allelic richness (mean 2.91; range 1.25–5.42). We observed deviations from Hardy–Weinberg equilibrium for some populations. Three populations deviated at a singular locus (*Sfo*C91), one population deviated at two loci (*Sfo*B52, *Sfo*C38), and one population deviated at three loci (*Sfo*B52, *Sfo*C88, and *Sfo*C91) (Supplemental File 2). Significant deviations from linkage equilibrium were detected in 47 populations for 27 different combinations of loci. A total of 35 populations displayed LD for a single pair of loci, nine populations deviated at two pairs, while deviations for single populations were detected at 4, 7, and 8 pairs of loci. The pair of loci that exhibited the greatest frequency of LD (39 populations) was *Sfo*C88 and *Sfo*C91. All other combinations of loci displayed LD in three or fewer populations.

Estimates of effective population size (N_E) ranged from 1.3 to 548, with a mean estimate of N_E across populations of 27.5 individuals. Eleven populations had an infinite point estimate, and 34 populations had infinity as an upper confidence bound, suggesting our ability to resolve effective population size was limited.

Hatchery introgression

The three metrics we used to characterize hatchery introgression were generally consistent within each population (average pairwise correlation coefficient 0.923, range 0.879–0.955; Fig. 3). Composite scores of hatchery introgression (SABT) ranged widely among populations (average

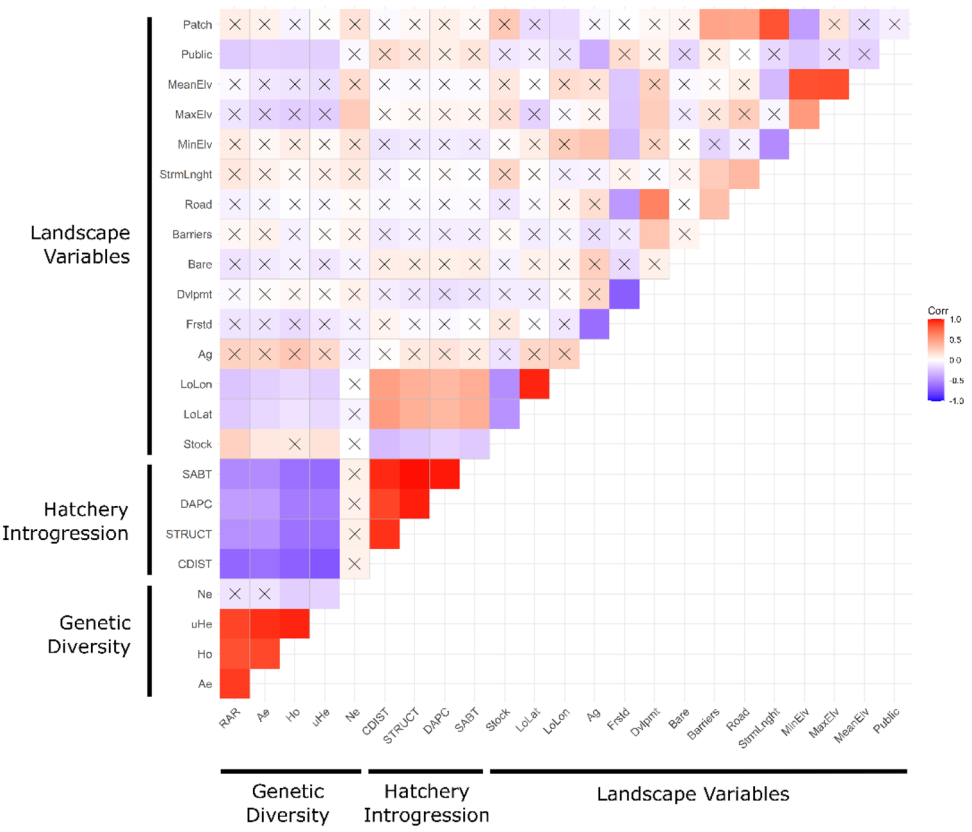

Fig. 3 A correlogram displaying the strength of correlations (denoted by colors) among and between landscape metric variables, metrics of genetic diversity, and measures of hatchery introgression for populations of Brook Trout sampled in Tennessee. Boxes with an X represent correlations that were not statistically significant ($\alpha = 0.05$ level)

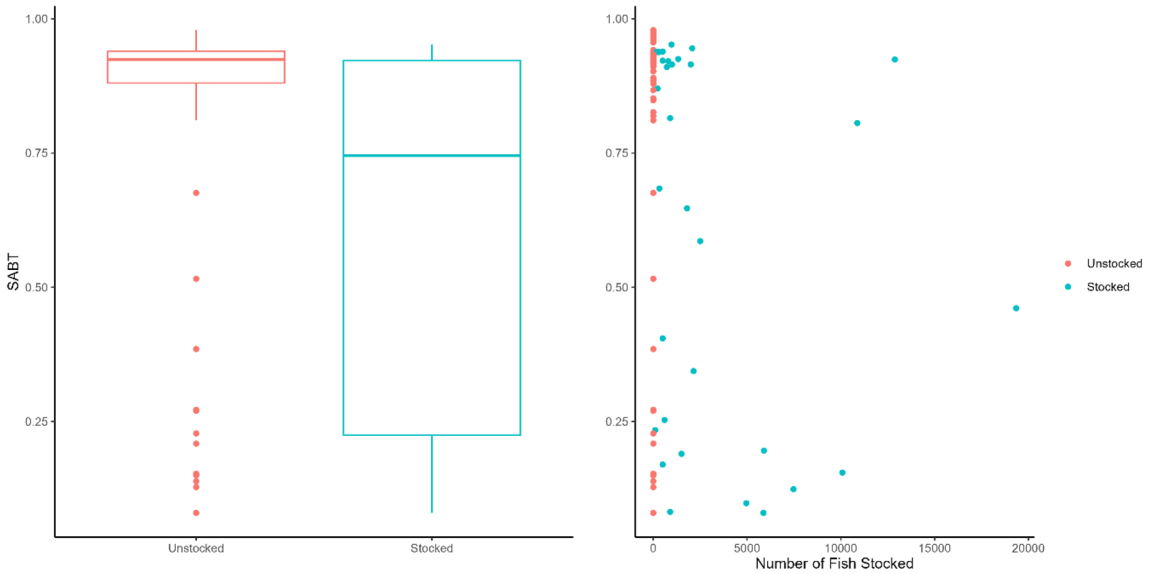

Fig. 4 The distribution of hatchery introgression among Tennessee Brook Trout populations quantified using a composite metric (SABT). The boxplot (left) displays hatchery introgression in relation to stocking as a categorical variable, and the scatterplot (right) displays introgression in relation to the numbers of known stocked Brook Trout into each of the respective populations. The horizontal bar in the boxplot represents the median value across populations and the lower and upper hinges correspond to the 25th and 75th percentiles. Upper and lower whiskers represent 1.5 times the inter-quartile range and points beyond whiskers represent outliers

0.74; range 0.08–0.98; Fig. 4), which indicated variable levels of hatchery ingression across the landscape. We observed much greater variability in SABT scores among stocked populations when compared to unstocked populations. However, many populations showed no indication of introgression and appeared to represent endemic lineages of native Brook Trout (Fig. 2).

Drainage characteristics

The drainages we assessed were typical of many occupied Brook Trout habitats in the southern Appalachian Mountains. Most drainages were small (mean 3.9 km^2; range 0.9–10.9 km^2; Supplemental File 3), drained higher elevations (maximum elevation in the drainages averaged 1338 m above sea level [ASL]; range 1052–1918 m ASL) and had limited habitat available to support wild Brook Trout populations (mean total stream length 8.5 km; range 0.7–30.2 km). Overall, the drainages showed limited anthropogenic disturbance on the contemporary landscape. Forest was the dominant land cover across all drainages considered in the study (mean 96.6%; range 71.1–100.0%). Agriculture (mean 0.7%; range 0.0–12.4%) and development were present in some drainages but were a minor component of the landscape. There were roads present in many of the drainages (mean length 3.9 km; range 0.0–41.7 km), but these were often located towards the downstream terminus of the drainage. The drainages we examined were predominantly under public ownership (mean 78.1%; range 0.0–100.0%), but there were some drainages that were predominantly under private ownership (17 drainages were ≥ 50.0% privately owned). Within the delineated drainages, the number of potential barriers identified ranged from 0 to 4 (mean = 0.7; but see comments in methods and discussion).

PCA

The first two principal components explained a combined total of 45.5% of the observed variation in landscape metrics across populations (Supplemental File 4). The most influential variable loadings for principal component 1 (PC1) were drainage area (0.46), roads (0.46), stream length (0.36), percent forest cover (−0.36), and percent agricultural cover (0.30) (Fig. 5, Supplemental File 5). Variable loadings for PC2 were highest for mean elevation (0.52), stream length (−0.43), maximum elevation (0.38), drainage area (−0.36), and percent forest cover (−0.31) (Fig. 5, Supplemental File 5). Because several landscape metrics explained similar amounts of variation across populations, we elected to retain a subset of landscape metrics for downstream analysis. Specifically, drainage area, stream length, and road all explained similar amounts of variation for PC1 and as a result we retained drainage area. Percent agriculture and percent forest were opposed in the extent of variation explained (i.e., for each PC one variable explained positive amounts of variation while the other explained negative amounts), and as

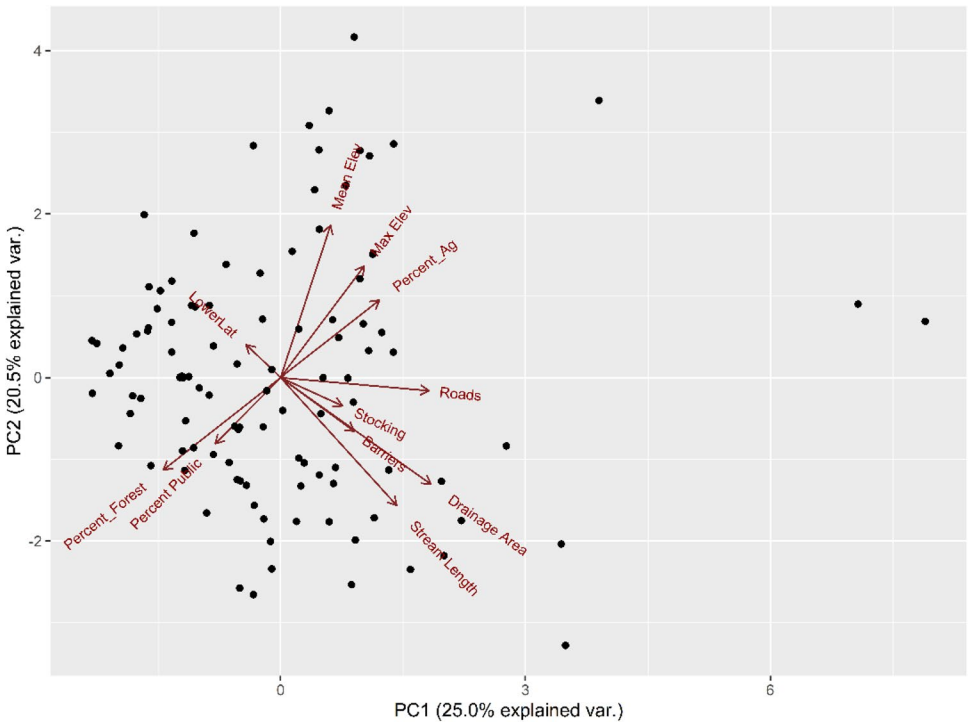

Fig. 5 Results from principal components analysis (PCA) displaying the variation in landscape variables among drainages where populations of Brook Trout were sampled in Tennessee

a result we elected to retain only percent forested coverage within a drainage.

Modeling the relationship between landscape variables and genetic characteristics

All observed correlations between landscape variables and genetic metrics were weak (e.g., $r < 0.29$), where they existed at all (Fig. 3). We chose to focus our modeling effort on three predictive landscape variables: (1) drainage area, which is assumed to reflect the size of the stream and the amount of habitat available to each population, (2) the proportion of the drainage with non-forest cover, as an index of the amount of anthropogenic disturbance, and (3) stocking history, which we assumed a priori would relate to the potential for hatchery introgression.

There were weak relationships between hatchery introgression and the three predictive variables (Table 1). The best supported model for hatchery introgression included only stocking history as a factor and explained a relatively minor amount of the overall variation (10.9%). Under this model, previously stocked drainages had higher levels of hatchery introgression (Table 2). A couple of the more complex models that included non-forested cover or drainage area received some support based on ΔAICc (< 2), but these were only able to explain a marginal amount of additional variation.

Similarly, the most supported models for rarefied allelic richness and unbiased expected heterozygosity included only stocking history as a factor (Table 1). Under the most supported model, rarefied allelic richness was generally greater in drainages with a history of stocking (8.9% of variation explained; Table 3). Similarly, unbiased expected heterozygosity was typically greater in drainages where hatchery Brook Trout had previously been stocked (5.7% of variation explained; Table 4). Many alternative model structures to explain variation in rarefied allelic richness and unbiased expected heterozygosity had substantial support based on ΔAICc (< 2; Table 1). In every instance, these alternative model structures receiving support included stocking history as a predictor, with small improvements in predictive power relative to additional model complexity.

The best model for predicting effective population size explained 23.5% of the observed variation and included a complex, three-way interaction between drainage area, non-forested land cover, and stocking history (Tables 1, 5)—making it difficult to draw straightforward inferences. A simpler model to predict effective population size using only drainage area received some support (ΔAICc = 1.1). Under this alternative model, larger drainage areas were associated with larger effective population sizes, but the relationship was weak (8.9% of variation explained).

Discussion

An understanding of land use and hatchery stocking impacts on population genetics could benefit Brook Trout management and conservation efforts in the southern Appalachian Mountains. Accordingly, we documented patterns of genetic diversity for eastern Tennessee Brook Trout populations and modelled effects of drainage area, forested cover, and hatchery stocking on effective population size (N_e) and genetic diversity. As has been reported elsewhere (e.g., Marie et al. 2010), genetic diversity (expected heterozygosity and rarefied allelic richness) was positively related to stocking effort. Interestingly the best fit models to explain genetic diversity included stocking alone (and not drainage area); however, only modest amounts of variation were explained and multiple models received substantial support (ΔAIC ≤ 2; Table 1). In contrast to genetic diversity, effective population size was best predicted by drainage area, with larger drainages being more likely to contain populations with higher N_e. Notably,

Table 2 Parameter estimates for the best supported linear model (number 18) to predict hatchery introgression within 106 Brook Trout drainages in Tennessee

Coefficient	Estimate	Standard error	*P*
Intercept	0.603	0.050	0.000
Stocking history (unstocked)	0.215	0.060	0.001

This model explained 10.9% of the observed variation in introgression scores

Table 3 Parameter estimates for the best linear model (number 18) to predict A_R within 75 Brook Trout drainages (n ≥ 20 samples) in Tennessee

Coefficient	Estimate	Standard error	*P*
Intercept	3.487	0.173	0.000
Stocking history (unstocked)	−0.555	0.208	0.009

This model explained 8.9% of the observed variation in allelic richness

Table 4 Parameter estimates for the best linear model (number 18) to predict uH_e within 75 Brook Trout drainages (n ≥ 20 samples) in Tennessee

Coefficient	Estimate	Standard error	*P*
Intercept	0.458	0.024	0.000
Stocking history (unstocked)	−0.061	0.029	0.039

This model explained 5.7% of the observed variation in expected heterozygosity

Table 5 Parameter estimates for the best linear model (number 1) to predict N_e within 75 Brook Trout drainages (n ≥ 20 samples) in Tennessee

Coefficient	Estimate	Standard error	*P*
Intercept	8.201	17.613	0.643
Area	4.373	3.125	0.166
Non-forested Cover	383.369	570.237	0.503
Stocking history (unstocked)	8.168	21.688	0.708
Area:Non-forested Cover	− 74.424	70.338	0.294
Area:Stocking history (unstocked)	− 4.163	4.480	0.356
Non-forested Cover:Stocking history (unstocked)	− 1037.166	656.729	0.119
Area:Non-forested Cover:Stocking history (unstocked)	338.345	121.140	0.007

This model explained 23.5% of the observed variation in effective population size

stronger support was observed for a complex model which included a three-way interaction term to explain effective population size, emphasizing a complex, potentially context-specific relationship among variables. Stocking had only a weak, positive correlation with hatchery introgression, and the presence or absence of stocking better explained levels of genetic diversity than total numbers of fish stocked. Hatchery introgression varied widely and was generally higher in stocked populations, but strong signals of introgression were detected in a fraction of populations with no known stocking history. Combined, this evidence suggests that establishing relationships between anthropogenic impacts and measures of genetic diversity may require finer-scale landscape metrics than drainage area or forested cover and that legacy impacts not captured in our modeling may continue to shape contemporary genetic patterns in Tennessee Brook Trout populations and those elsewhere in the southern Appalachian Mountains.

Stream fishes occupying small drainages with limited or low-quality habitat are predicted to experience an elevated risk of extirpation via demographic or environmental variation (Boyce 1992). In Tennessee, we expected to see a positive relationship between drainage area and genetic diversity, as increased area is expected to contain greater amounts of habitat, support larger populations, and facilitate gene flow among populations (Hilderbrand and Kershner 2000). Indeed, numerous studies have identified a positive relationship between patch or drainage area and genetic diversity among salmonids (Neville et al. 2006; Whiteley et al. 2010; Kovach et al. 2015; Buonaccorsi et al. 2017), and for Brook Trout in particular (Whiteley et al. 2013, 2014; Nathan et al. 2020). In northern Virginia, the area of patches containing Brook Trout populations was positively related to both allelic richness (R^2 range 0.57–0.81) and mean expected heterozygosity (R^2 range 0.55–0.98) (Whiteley et al. 2013). Similarly, patch area and catchment area were identified as the best predictors of genetic diversity ($R^2_{adj} = 0.25$) in Brook Trout populations in Connecticut and measures of area and genetic diversity were positively correlated (Nathan et al. 2020). In our study, drainage area was only included in the best model explaining variation in effective population size (Table 1). Although many of the less-supported models for genetic diversity and hatchery introgression included drainage area, the improvements gained through the addition of this predictor were always minimal, and simple univariate models based on drainage area only explained 0.1–0.2% of the observed variation (Table 1). We note that our approach relied upon drainage-level characteristics as opposed to patch-level, which may influence predictive abilities; however, 62% of our drainages were identical to the EBTJV patches, and we would generally expect to see similar relationships between area and levels of genetic diversity.

Overall, we observed very weak correlations between landscape characteristics and measures of genetic diversity (Fig. 3), and this may be explained by several potential mechanisms. First, the area of drainages containing Brook Trout in Tennessee were small (mean 3.9 km^2, median 3.4 km^2, range 0.9–10.9 km^2) relative to previous studies (Whiteley et al. 2013; average = 37.3 km^2, range 9.9–108 km^2) and to the southeast United States in general (median 8.6 km^2; Whiteley et al. 2013). Smaller populations are subject to greater levels of demographic and genetic stochasticity (Lande 1993; Wood et al. 2014), suggesting that evolutionary forces such as genetic drift may more strongly affect levels of genetic diversity in Tennessee Brook Trout populations than landscape factors such as drainage area. Second, the presence of historical stocking in the best models explaining genetic diversity highlights its role in shaping contemporary levels of genetic diversity in Tennessee Brook Trout populations. While stocking can increase levels of genetic diversity, the introduced diversity would have been derived from non-native sources (Habera and Strange 1993), may lead to genetic homogenization across populations (Marie et al. 2010), and may negatively affect fitness, resiliency, and adaptive potential (Hallerman 2003; Naish et al. 2007). Lastly, legacy effects (see below) not captured in our current analysis may also be an important factor to consider in future modeling efforts.

Brook Trout are particularly susceptible to physical and chemical changes to the environment, and land use practices

such as logging and agriculture have been implicated in localized declines (MacCrimmon and Campbell 1969; Hudy et al. 2005; Stranko et al. 2008; DeWeber and Wagner 2015; Kanno et al. 2015). We expected Brook Trout in intact landscapes with higher levels of forested cover to display higher levels of genetic diversity and have larger effective population sizes. Instead, we did not observe a meaningful impact of land cover on either heterozygosity or allelic richness (Table 1; Fig. 3), and the observed relationship with effective population size was weak and complex. Numerous factors have been proposed to explain a lack of clear relationship between land use and biological response (Allan 2004), and we argue that legacy effects may be largely responsible. Changes to land cover (e.g., natural forest converted to agriculture) can alter soil nutrient dynamics and biodiversity levels, and these impacts remain detectable decades and even centuries later (Fraterrigo et al. 2005; Compton and Boone 2000). Stream surveys in North Carolina identified that aquatic invertebrate and fish biodiversity were best explained by historic (40 years prior) and not contemporary whole watershed land use (Harding et al. 1998). The French Broad and Little Tennessee River watersheds examined by Harding et al. (1998), drain from North Carolina into Tennessee and contain ~ 15% of our study sites. Thus, there is direct evidence at the drainage basin scale that legacy effects have impacted stream community diversity and abundance in our study system—these alterations may also be responsible for shaping contemporary patterns of Brook Trout genetic diversity. Future investigations into drivers of genetic structure may benefit from quantifying historical anthropogenic impacts in addition to contemporary changes.

In contrast to models that best explained metrics of genetic diversity, N_e was best explained by a model that included interactions between drainage area, forested cover, and stocking, but this relationship was weak and explained only a minor amount of the observed variation ($R^2 = 0.24$). A second model which received substantial support included only drainage area, but this model had lower explanatory power ($R^2 = 0.09$). In general, effective population size estimates were small (mean = 27.5; median = 12.9) and varied little across populations, which may be an artifact of the low carrying capacities of headwater streams in Tennessee. Habera et al. (2001) noted that Brook Trout occurred at low abundances across Tennessee, and attributed the low abundance to limited food supplies stemming from soft, infertile waters typical of the southern Appalachian Mountains. As smaller populations may be more impacted by stochastic processes than large ones, random processes such as genetic drift could be obscuring relationships between effective population size and landscape characteristics that would otherwise be observed. Another factor that may explain our limited ability to explain variation in N_e was the resolution of our genetic data set. While the molecular markers we used are highly polymorphic and have been used to characterize Brook Trout throughout their range (e.g., King et al. 2012; Kazyak et al. 2015, 2018), the populations we examined were small and displayed lower levels of diversity relative to range-wide estimates. Both of those factors may decrease our ability to explain variation in N_e estimates.

Additional factors may also explain the weak relationships observed between genetic diversity, N_e, and landscape characteristics. Our modeling efforts utilized landscape characteristics described at the drainage-level, which differed slightly from previous population genetic studies of Brook Trout (e.g., Peterson et al. 2008; Whiteley et al. 2010, 2013). As mentioned previously, we attempted to utilize a framework common to prior studies (e.g., EBTJV patch layers), but approximately 12% of our sites were not represented by available patches, and additional populations were genetically distinct but combined into common patches. As a result, we generated a novel landscape character dataset which included attributes of the upstream catchment, and this scale of description may be too coarse to establish tight landscape-genetic relationships. However, 62% of our drainages were identical to EBTJV layers, suggesting that our landscape approach alone is unlikely to explain the observed weak relationships. Other factors, such as in-stream features common in many high gradient stream landscapes can inhibit gene flow to varying degrees (Weathers et al. 2019). We identified some potential barriers (e.g., waterfalls and cascades) within our drainages and consider these to be a conservative estimate of such features. Although it is intractable to identify all such features in landscape-scale studies, these barriers undoubtedly impact the genetic characteristics of wild populations and obscure our ability to make generalities. We note that ongoing local and regional efforts to collect and organize fish passage survey information (e.g., Southeast Aquatic Connectivity Program) may help fill this data gap. A failure to establish strong relationships between landscape attributes and genetic diversity and N_e suggests that there was either a lack of substantive variation in characteristics across sites, that the level of detail was too coarse to capture relevant patterns and processes, or that genetic metrics are not influenced by landscape variables.

The genetic impacts of stocking on native populations of Brook Trout have been studied throughout their range (e.g., Marie et al. 2010; White et al. 2018; Beer et al. 2019; Kazyak et al. 2018), and overall, the extent of change appears to be context specific. Specifically, several studies have identified minimal signs of hatchery introgression in Brook Trout populations despite historical or ongoing stocking efforts (Annett et al. 2012; White et al. 2018; Beer et al. 2019; Kazyak et al. 2018), while others have noted widespread impacts (Marie et al. 2010). We observed significant variation among Tennessee populations with respect to hatchery influence and showed that hatchery introgression

was only weakly related to stocking intensity (Fig. 3). The failure for stocking to have a strong impact on hatchery influence could be explained by several factors including poor survival of hatchery-origin Brook Trout in the wild (Danzmann and Ihseen 1995; Flowers et al. 2019), harvest by recreational anglers (Annett et al. 2012), or emigration from the site of stocking (Flowers et al. 2019). Because of difficulties in rearing Brook Trout from the southern Appalachian Mountains in the hatchery environment, non-native strains (namely from the northern Appalachian stocks; Habera and Strange 1993) have been used as brood stock for hatchery supplementation efforts. Reduced fitness of hatchery-origin Brook Trout in Tennessee streams may be driven by a mismatch between native and non-native environments. In addition, hatchery-origin fish may maintain reproductive isolation from wild conspecifics via temporal differences in spawn timing or utilization of distinct spawning habitats (Quinn et al. 2000; Fleming and Petersson 2001).

The presence of hatchery influence in populations without known stocking events represents an important finding of management relevance. There are several possible mechanisms that may explain the presence of hatchery alleles beyond the point of introduction including dispersal from nearby stocked sites and undocumented introductions. Data on movement and dispersal have indicated that stocked Brook Trout typically exhibit low levels of dispersal, with most fish moving < 2 km from the stocking site (Flowers et al. 2019). Likewise, movement levels among native populations of Brook Trout occupying smaller tributary systems in West Virginia were also low (< 175 m total dispersal; Petty et al. 2012), suggesting that dispersal among our research sites would be an unlikely explanation for the presence of hatchery alleles into unstocked populations. Alternatively, hatchery introgression in unstocked sites may be the product of undocumented Brook Trout introductions or translocations by the public or by fishery managers. The presence of non-sanctioned, angler-mediated translocations is a widespread problem globally (Ellender and Weyl 2014; Hargrove et al. 2015) and has occurred in Tennessee. For example, non-native Alabama Bass (*Micropterus henshalli*) were illegally introduced into Parksville Reservoir and may negatively affect native Black Bass (*Micropterus* spp.) through hybridization and competition (Moyer et al. 2014). Additionally, Brook Trout restoration efforts in Great Smoky National Park were stymied by the illegal introduction of non-native Rainbow Trout (*Oncorhynchus mykiss*) following their initial eradication (Moore and Kulp 2019). It is also possible that some stocking events prior to 1952 were not documented, after which TWRA began keeping detailed stocking records. The U.S. Forest Service also stocked hatchery Brook Trout fingerlings and translocated wild fish in the Cherokee National Forest during the 1970s and 1980s—some of which may have been undocumented. Taken together, unauthorized or undocumented fish introductions, along with movement of Brook Trout among streams may explain hatchery introgression in the unstocked sites we studied. Moving forward, known levels of hatchery introgression will be informative for conservation planning, as populations with elevated levels of non-native alleles may be excluded as a source for translocations efforts and may become the focus of future genetic restoration work.

Conservation implications

Our statewide genetic assessment of Tennessee Brook Trout populations provides valuable insights for future conservation efforts. Although we failed to identify strong relationships between land use, drainage area, and hatchery stockings on metrics of genetic diversity, N_e, and hatchery introgression, this finding is noteworthy for several reasons. First, the ability to predict the genetic characteristics of populations in unsampled space in Tennessee and other portions of the southern range of eastern Brook Trout is likely to be limited. Second, given that many Brook Trout populations in Tennessee are small and located on public lands, processes such as genetic drift may pose a larger threat to future persistence than land use practices such as agriculture or development. In addition to relating genetic diversity to landscape characteristics, we identified signals of hatchery introgression across the landscape which raises two key points. The detection of hatchery introgression in unstocked sites suggests legacy effects from either illegal stocking by anglers or undocumented stocking by state or federal agencies. If hatchery introgression occurred because of illegal, angler-mediated introductions, then outreach and education programs may help prevent future introductions outside of official conservation efforts (Moore and Kulp 2019). Additionally, translocations may be used as part of ongoing Brook Trout restoration efforts in Tennessee, and knowledge of introgression levels could help identify appropriate donor populations to prevent the spread of non-native alleles across the landscape. Lastly, populations with hatchery introgression had higher levels of genetic diversity, and testing for the presence of hatchery influence within native populations could benefit our understanding of sources of diversity (e.g., native vs. introduced) in order to prioritize populations for conservation. Moving forward, we have provided a detailed description of the genetic characteristics of Brook Trout in Tennessee and these results could play an important role in future conservation and management efforts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10592-021-01404-8.

Acknowledgements We thank the Biology Department at Tennessee Technological University for providing molecular laboratory access. We appreciate the efforts of J. Emmert, Z. Harbison, C. Chapman,

D.R. Lindbom, M. Thurman, T. Scott, and W. Collier who helped with Brook Trout sampling across the rugged landscape of eastern Tennessee. Funding for this study was provided by TWRA, Tennessee Trout Unlimited, the U.S. Forest Service, and Federal Aid in Sport Fish Restoration funds. Use of trade, product, or firm names does not imply endorsement by the U.S. Government.

Author contributions DCK and JSH designed the study, performed statistical analyses, and drafted the manuscript. OB and BAL worked together to complete all laboratory procedures. KR and KAF completed the GIS analyses. JWH and JH were instrumental in the development of the study and were responsible for field collections. All authors participated in substantive conversations during the development of the manuscript and contributed to its preparation.

Funding Funding for this study was provided by TWRA, Tennessee Trout Unlimited, and Federal Aid in Sport Fish Restoration funds.

Data availability Multilocus genotypes are provided in Supplemental File 1 in GenAlEx format. Summary data for each collection is provided in Supplemental File 2. Drainage delineations are provided in Supplemental File 3 as an ArcGIS shapefile.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest or competing interests.

References

Allan JD (2004) Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu Rev Ecol Evol Syst 35:257–284

Annett B, Gerlach G, King TL, Whiteley AR (2012) Conservation genetics of remnant coastal brook trout populations at the southern limit of their distribution: population structure and effects of stocking. Trans Am Fish Soc 141:1399–1410

Beer SD, Cornett S, Austerman P, Trometer B, Hoffman T, Bartron ML (2019) Genetic diversity, admixture, and hatchery influence in Brook Trout (*Salvelinus fontinalis*) throughout western New York State. Ecol Evol 9:7455–7479

Bivens RD (1984) History and distribution of Brook Trout in the Appalachian Region of Tennessee. Master's thesis. University of Tennessee, Knoxville. https://trace.tennessee.edu/utk_gradthes/2526/

Blum MJ, Bagley MJ, Walters DM, Jackson SA, Daniel FB, Chaloud DJ, Cade BS (2012) Genetic diversity and species diversity of stream fishes covary across a land-use gradient. Oecologia 168:83–95

Bolstad GH, Hindar K, Robertsen G, Jonsson B, Sægrov H, Diserud OH, Fiske P, Jensen AJ, Urdal K, Næsje TF, Barlaup BT (2017) Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. Nat Ecol Evol 1:1–5. https://doi.org/10.1038/s41559-017-0124

Boyce MS (1992) Population viability analysis. Annu Rev Ecol Syst 23:481–506

Buonaccorsi VP, Malloy J, Peterson M, Brubaker K, Grant CJ (2017) Population genomic analysis of brook trout in Pennsylvania's Appalachian Region. Trans Am Fish Soc 146:485–494

Burnham K, Anderson D (2002) Model selection and multimodel inference: a practical information theoretic approach, 2nd edn. Springer, New York, p 488

Compton JE, Boone RD (2000) Long-term impacts of agriculture on soil carbon and nitrogen in New England forests. Ecology 81:2314–2330

Currens KP, Hemmingsen AR, French RA, Buchanan DV, Schreck CB, Li HW (1997) Introgression and susceptibility to disease in a wild population of rainbow trout. N Am J Fish Manag 17:1065–1078. https://doi.org/10.1577/1548-8675(1997)017%3c1065:IASTDI%3e2.3.CO;2

Danzmann RG, Ihssen PE (1995) A phylogeographic survey of brook charr (*Salvelinus fontinalis*) in Algonquin Park, Ontario based upon mitochondrial DNA variation. Mol Ecol 4:681–698

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14:209–214

Dunham JB, Rieman BE, Peterson JT (2002) Patch-based models to predict species occurrence: lessons from salmonid fishes in streams. In: Scott JM, Morrison ML, Heglund PJ (eds) Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, DC, pp 327–334

DeWeber JT, Wagner T (2015) Predicting brook trout occurrence in stream reaches throughout their native range in the eastern United States. Trans Am Fish Soc 144:11–24

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Eastern Brook Trout Joint Venture (EBTJV) (2011) Conserving the Eastern Brook Trout: Action Strategies. https://easternbrooktrout.org/reports/ebtjv-conservation-strategy-2011#:~:text=The%20vision%20of%20the%20EBTJV,threats%20and%20disturbances%3B%20(2). Accessed 2 Dec 2020

Ellender BR, Weyl OL (2014) A review of current knowledge, risk and ecological impacts associated with non-native freshwater fish introductions in South Africa. Aquat Invas 9:117–132

Esri, Inc. (2011) ArcGIS desktop: release 10.6. Environmental Systems Research Institute, Redlands

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

Fausch KD, Torgersen CE, Baxter CV, Li HW (2002) Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes: a continuous view of the river is needed to understand how processes interacting among scales set the context for stream fishes and their habitat. Bioscience 52:483–498

Ferchaud A-L, Laporte M, Perrier C, Bernatchez L (2018) Impact of supplementation on deleterious mutation distribution in an exploited salmonid. Evol Appl 11:1053–1065

Flebbe PA, Roghair LD, Bruggink JL (2006) Spatial modeling to project southern Appalachian trout distribution in a warmer climate. Trans Am Fish Soc 135:1371–1382

Fleming IA, Petersson E (2001) The ability of released, hatchery salmonids to breed and contribute to the natural productivity of wild populations. Nord J Freshw Res 5:71–98

Flowers HJ, Kwak TJ, Fischer JR, Cope WG, Rash JM, Besler DA (2019) Behavior and survival of stocked trout in southern Appalachian Mountain streams. Trans Am Fish Soc 148:3–20

Fraterrigo JM, Turner MG, Pearson SM, Dixon P (2005) Effects of past land use on spatial heterogeneity of soil nutrients in southern Appalachian forests. Ecol Monogr 75:215–230

Galbreath PF, Adams ND, Guffey SZ, Moore CJ, West JL (2001) Persistence of native southern Appalachian brook trout populations in the Pigeon River system, North Carolina. N Am J Fish Manag 21:927–934

Gossieaux P, Lavoie E, Sirois P, Thibault I, Bernatchez L, Garant D (2020) Effects of genetic origin on phenotypic divergence in

Brook Trout populations stocked with domestic fish. Freshw Ecol 11:e03119
Goudet J, Jombart T (2015) hierfstat: estimation and tests of hierarchical F-statistics. R package version 0.04-22. http://CRAN.Rproject.org/package=hierfstat
Gozlan RE, Britton JR, Cowx I, Copp GH (2010) Current knowledge on non-native freshwater fish introductions. J Fish Biol 76:751–786
Gresswell RE, Vondracek B (2010) Chapter 18 coldwater streams. In: Hubert W, Quist M (eds) Inland fisheries management in North America, 3rd edn. American Fisheries Society, Bethesda, pp 587–618
Guffey SZ (1998) A population genetic study of southern Appalachian Brook Trout. PhD dissertation. University of Tennessee
Habera J, Moore S (2005) Managing southern Appalachian brook trout: a position statement. Fisheries 30:10–20
Habera JW, Strange RJ (1993) Wild trout resources and management in the southern Appalachian Mountains. Fisheries 18:6–13
Habera JW, Strange RJ, Bivens RD (2001) A revised outlook for Tennessee's brook trout. J Tenn Acad Sci 76:68–73
Hallerman EM (2003) Coadaptation and outbreeding depression. In: Hallerman EM (ed) Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, pp 239–259
Harbicht AB, Alshamlih M, Wilson CW, Fraser DJ (2014) Anthropogenic and habitat correlates of hybridization between hatchery and wild brook trout. Can J Fish Aquat Sci 71:688–697
Harding JS, Benfield EF, Bolstad PV, Helfman GS, Jones EB (1998) Stream biodiversity: the ghost of land use past. Proc Natl Acad Sci 95:14843–14847
Hargrove JS, Weyl OL, Allen MS, Deacon NR (2015) Using tournament angler data to rapidly assess the invasion status of alien sport fishes (*Micropterus* spp.) in Southern Africa. PLoS ONE 10:e0130056
Hargrove JS, Weyl OL, Austin JD (2017) Reconstructing the introduction history of an invasive fish predator in South Africa. Biol Invas 19:2261–2276
Hayes JP, Guffey SZ, Kriegler FJ, McCracken GF, Parker CR (1996) The genetic diversity of native, stocked, and hybrid populations of brook trout in the southern Appalachians. Conserv Biol 10:1403–1412
Heath DD, Heath JW, Bryden CA, Johnson RM, Fox CW (2003) Rapid evolution of egg size in captive salmon. Science 299:1738–1740
Hess JE, Campbell NR, Docker MF, Baker C, Jackson A, Lampman R, McIlraith B, Moser ML, Statler DP, Young WP, Wildbill AJ (2015) Use of genotyping by sequencing data to develop a high-throughput and multifunctional SNP panel for conservation applications in Pacific lamprey. Mol Ecol Resour 15:187–202
Hilderbrand RH, Kershner JL (2000) Conserving inland cutthroat trout in small streams: how much stream is enough? N Am J Fish Manag 30:513–520
Horak D (1995) Native and nonnative fish species used in state fisheries management programs in the United States. In: Schramm HL, Piper RG (eds) Uses and effects of cultured fishes in aquatic ecosystems, Symposium 15. American Fisheries Society, Bethesda, pp 61–67
Hudy M, Thieling TM, Gillespie N, Smith EP (2005) Distribution, status, and perturbations to brook trout within the eastern United States. Final report to the steering committee of the Eastern Brook Trout Joint Venture
Hudy M, Thieling TM, Gillespie N, Smith EP (2008) Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the eastern United States. N Am J Fish Manag 28:1069–1085
Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806
Jombart T et al (2008) Package 'adegenet.' Bioinform Appl Note 24:1403–1405
Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:1–15
Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Resour 10:551–555
Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189
Kanno Y, Letcher BH, Rosner AL, O'Neil KP, Nislow KH (2015) Environmental factors affecting brook trout occurrence in headwater stream segments. Trans Am Fish Soc 144:373–382
Kazyak DC, Hilderbrand RH, Keller SR, Colaw MC, Holloway AE, Morgan RP, King TL (2015) Spatial structure of morphological and neutral genetic variation in Brook Trout. Trans Am Fish Soc 144:480–490
Kazyak DC, Rash J, Lubinski BA, King TL (2018) Assessing the impact of stocking northern-origin hatchery Brook Trout on the genetics of wild populations in North Carolina. Conserv Genet 19:207–219. https://doi.org/10.1007/s10592-017-1037-4
King TL, Lubinski BA, Burnham-Curtis MK, Stott W, Morgan RP II (2012) Tools for the management and conservation of genetic diversity in brook trout (*Salvelinus fontinalis*): tri and tetranucleotide microsatellite markers for the assessment of genetic diversity, phylogeography, and historical demographics. Conserv Genet Resour 4:539–543
Kovach RP, Muhlfeld CC, Wade AA, Hand BK, Whited DC, DeHaan PW, Al-Chokhachy R, Luikart G (2015) Genetic diversity is related to climatic variation and vulnerability in threatened bull trout. Glob Change Biol 21:2510–2524
Kriegler FJ, McCracken GF, Habera JW, Strange RJ (1995) Genetic characterization of Tennessee brook trout populations and associated management implications. N Am J Fish Manag 15:804–813
Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am Nat 142:911–927
Le Luyer J, Laporte M, Beacham TD, Kaukinen KH, Withler RE, Leong JS, Rondeau EB, Koop BF, Bernatchez L (2017) Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proc Natl Acad Sci 114:12964–12969
Lennon RE (1967) Brook trout of Great Smoky Mountains National Park. U.S. Bureau of Sport Fisheries and Wildlife Technical Paper 15
Lunn T, Munks S, Carver S (2017) The impacts of timber harvesting on stream biota: an expanding field of heterogeneity. Biol Conserv 213:154–166
MacCrimmon HR, Campbell JS (1969) World distribution of brook trout, *Salvelinus fontinalis*. J Fish Board Can 26:1699–1725
Malone EW, Perkin JS, Leckie BM, Kulp MA, Hurt CR, Walker DM (2018) Which species, how many, and from where: integrating habitat suitability, population genomics, and abundance estimates into species reintroduction planning. Glob Change Biol 24:3729–3748
Marie AD, Bernatchez L, Garant D (2010) Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). Mol Ecol 19:2025–2037
McCracken GF, Parker CR, Guffey SZ (1993) Genetic differentiation and hybridization between stocked hatchery and native brook trout in Great Smoky Mountains National Park. Trans Am Fish Soc 122:533–542
Moore SE, Kulp MA (2019) Overcoming challenges to native fish restoration in a National Park. In: Kreuger CC, Taylor WW, Youn S (eds) From catastrophe to recovery: stories of fishery management success. American Fisheries Society, Bethesda, pp 45–64

Moyer GR, Williams AS, Jolley ML, Ragland BJ, Churchill TN (2014) Genetic confirmation and assessment of an unauthorized fish introduction in Parksville Reservoir, Tennessee. J Southeast Assoc Fish Wildl Agencies 1:64–69

Naish KA, Taylor JE, Levin PS, Quinn TP, Winton JR, Huppert D, Hilborn R (2007) An evaluation of the effects of conservation and fishery enhancement hatcheries on wild populations of salmon. Adv Mar Biol 53:61–194. https://doi.org/10.1016/S0065-2881(07)53002-6

Nathan LR, Kanno Y, Letcher BH, Welsh AB, Whiteley AR, Vokoun JC (2020) Evaluation of genetic structuring within GIS-derived brook trout management units. Trans Am Fish Soc 149:681–694

Nathan LR, Welsh AB, Vokoun JC (2019) Watershed-level brook trout genetic structuring: evaluation and application of riverscape genetics models. Freshw Biol 64:405–420

Neville HM, Dunham JB, Peacock MM (2006) Landscape attributes and life history variability shape genetic structure of trout populations in a stream network. Landsc Ecol 21:901–916

Ozerov MY, Gross R, Bruneaux M, Vähä JP, Burimski O, Pukk L, Vasemägi A (2016) Genomewide introgressive hybridization patterns in wild Atlantic salmon influenced by inadvertent gene flow from hatchery releases. Mol Ecol 25:1275–1293

Pavlova A, Beheregaray LB, Coleman R, Gilligan D, Harrisson KA, Ingram BA, Kearns J, Lamb AM, Lintermans M, Lyon J, Nguyen TT (2017) Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: a call for assisted gene flow. Evol Appl 10:531–550

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539

Peterson DP, Rieman BE, Dunham JB, Fausch KD, Young MK (2008) Analysis of trade-offs between threats of invasion by nonnative brook trout (*Salvelinus fontinalis*) and intentional isolation for native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*). Can J Fish Aquat Sci 65:557–573

Peterson DP, Rieman BE, Horan DL, Young MK (2014) Patch size but not short-term isolation influences occurrence of westslope cutthroat trout above human-made barriers. Ecol Freshw Fish 23:556–571

Petty JT, Hansbarger JL, Huntsman BM, Mazik PM (2012) Brook trout movement in response to temperature, flow, and thermal refugia within a complex Appalachian riverscape. Trans Am Fish Soc 141:1060–1073

Pregler KC, Kanno Y, Rankin D, Coombs JA, Whiteley AR (2018) Characterizing genetic integrity of rear-edge trout populations in the southern Appalachians. Conserv Genet 19:1487–1503

Pringle CM, Naiman RJ, Bretschko G, Karr JR, Oswood MW, Webster JR, Welcomme RL, Winterbourn MJ (1988) Patch dynamics in lotic systems: the stream as a mosaic. J N Am Benthol Soc 7:503–524

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Quinn TP, Unwin MJ, Kinnison MT (2000) Evolution of temporal isolation in the wild: genetic divergence in timing of migration and breeding by introduced chinook salmon populations. Evolution 54:1372–1385

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Sanz N, Cortey M, Pla C, García-Marín JL (2006) Hatchery introgression blurs ancient hybridization between brown trout (*Salmo trutta*) lineages as indicated by complementary allozymes and mtDNA markers. Biol Conserv 130:278–289

Seehorn T (2004) Brook trout in the Chattooga drainage: a stream suitability index. Clemson University

Stauffer JR, King TL (2014) Designation of a neotype for brook trout, Salvelinus fontinalis. Proc Biol Soc Washington 127:557–567

Stranko SA, Hurd MK, Klauda RJ (2005) Applying a large, statewide database to the assessment, stressor diagnosis, and restoration of stream fish communities. Environ Monit Assess S108:99–121

Stranko SA, Hilderbrand RH, Morgan RP, Staley MW, Becker AJ, Roseberry-Lincoln A, Perry ES, Jacobson PT (2008) Brook trout declines with land cover and temperature changes in Maryland. N Am J Fish Manag 28:1223–1232

Sundström LF, Petersson E, Höjesjö J, Johnsson JI, Järvi T (2004) Hatchery selection promotes boldness in newly hatched brown trout (*Salmo trutta*): implications for dominance. Behav Ecol 15:192–198

Trushenski J, Flagg TH, Kohler CH (2010) Use of hatchery fish for conservation, restoration, and enhancement of fisheries. In: Kohler C, Hubert WA (eds) Inland fisheries management in North America, 3rd edn. American Fisheries Society, Bethesda, pp 261–293

U.S. Census Bureau (2020) TIGER Data Products Guide. https://www.census.gov/programs-surveys/geography/guidance/tiger-data-products-guide.html

U.S. Geological Survey (2020) National Elevation Dataset. https://viewer.nationalmap.gov/basic/

U.S. Geological Survey EROS (2020) National Land Cover Dataset 2016. https://earthexplorer.usgs.gov/

Wade AA, Hand BK, Kovach RP, Luikart G, Whited DC, Muhlfeld CC (2017) Accounting for adaptive capacity and uncertainty in assessments of species' climate-change vulnerability. Conserv Biol 31:136–149

Waples RS, Anderson EC (2017) Purging putative siblings from population genetic data sets: a cautionary view. Mol Ecol 26:1211–1224

Weathers TC, Kazyak DC, Stauffer Jr JR, Kulp MA, Moore SE, King TL, Carlson JE (2019) Neutral genetic and phenotypic variation within and among isolated headwater populations of brook trout. Trans Am Fish Socs 148:58–72

White SL, Miller WL, Dowell SA, Bartron ML, Wagner T (2018) Limited hatchery introgression into wild brook trout (*Salvelinus fontinalis*) populations despite reoccurring stocking. Evol Appl 11:1567–1581

Whiteley AR, Hastings K, Wenburg JK, Frissell CA, Martin JC, Allendorf FW (2010) Genetic variation and effective population size in isolated populations of coastal cutthroat trout. Conserv Genet 11:1929–1943

Whiteley AR, Coombs JA, Hudy M, Robinson Z, Colton AR, Nislow KH, Letcher BH (2013) Fragmentation and patch size shape genetic structure of brook trout populations. Can J Fish Aquat Sci 70:678–688

Whiteley AR, Hudy M, Robinson ZL, Coombs JA, Nislow KH (2014) Patch-based metrics: a cost effective method for short and long-term monitoring of EBTJV wild Brook Trout populations. In: Wild Trout XI: looking back and moving forward. Wild Trout Symposium, West Yellowstone

Wood JL, Belmar-Lucero S, Hutchings JA, Fraser DJ (2014) Relationship of habitat variability to population size in a stream fish. Ecol Appl 24:1085–1100