Microsatellite analysis of brook trout, *Salvelinus fontinalis*, in western Pennsylvania

Julia M. Horiates

A thesis submitted to the Department of Biology, East Carolina University, in partial fulfillment of the requirements for Biology Honors Thesis

Advisor: Dr. Christopher Balakrishnan, Ph.D
East Carolina University
Department of Biology

April 26th, 2017

Biology Honors Thesis

Microsatellite analysis of brook trout, *Salvelinus fontinalis*, in western Pennsylvania

Julia M. Horiates

APPROVED BY:

Director of Thesis: _____
Christopher N. Balakrishnan, Ph.D

Assessor: _____
Susan B. McRae, Ph.D

Director of Undergraduate Studies: _____
Jean-Luc Scemama, Ph.D

On this day, 4/27/17

I hereby declare I am the sole author of this thesis. It is the result of my own work and is not the outcome of work done in collaboration, except for the fieldwork focused in obtaining the fin clip samples, which were completed by Dr. David Argent and Dr. William Kimmel. It has not been submitted elsewhere as coursework for this or another degree.

Signed: Julia M Horiates Date: 4/27/17

Julia M. Horiates

Microsatellite analysis of brook trout, *Salvelinus fontinalis*, in western Pennsylvania

Julia M. Horiates

Department of Biology, East Carolina University, Greenville N.C. 27858-4353, U.S.A.

Abstract

In western Pennsylvania, the brook trout, *Salvelinus fontinalis*, has dramatically declined in population due to the environmental stressors of mining, logging, water withdrawal, and acid deposition. Brook trout are game fish and are valued economically in this region which makes their declining populations concerning. To better understand brook trout population dynamics in western Pennsylvania, we conducted a genetic analysis to determine patterns of population connectivity and genetic diversity. Microsatellites are short repeat sequences within the genome, and are highly susceptible to mutation making them suitable as neutral markers for analyzing patterns of regional population variation. Therefore, we conducted a microsatellite analysis of 16 individuals from 4 streams, for a total of 64 samples. Seven tri-and tetranucleotide microsatellite DNA markers were analyzed to investigate allelic diversity of brook trout. Amplified microsatellite loci were analyzed using Geneious 6.0 software that estimated PCR product sizes and converted sizes into estimates of the number of microsatellite repeats at each locus. Genepop 4.2 program was used to obtain the fixation index (F_{ST}) to measure levels of population differentiation and connectivity. The results showed moderate divergence (average pairwise divergence F_{ST}= 0.24; range of 0.06-0.30) with highly significant differentiation ($P < 0.001$). My study suggests that even populations that are spatially close to one another show evidence of divergence. This finding

suggests limited dispersal between streams and will inform species management strategies for this species.

Acknowledgments

I sincerely thank Dr. Christopher Balakrishnan for serving as my mentor and for his support and guidance. Additionally, thank you to collaborators Dr. David Argent and Dr. William Kimmel at California University of Pennsylvania for providing the fin clip samples. Furthermore, thank you to Dr. Sue McRae for serving as my assessor and both the Balalab and McRae lab for support and encouragement. Also, thank you to Dr. Denise Mayer for assistance with genotyping. To my family and friends, thank you for encouragement and love. Lastly, I would like to thank the ECU Honors College 2014 Scholarly Activity Awards for Students (SAAS) Grant for helping fund this project and the ECU Biology Department for providing their facilities.

Table of Contents

Introduction 10

Study goals 15

Methods 16

Fin Clip Sampling and DNA Extraction 16

PCR and Genotyping 18

Data Analysis 18

Results 19

Discussion 25

Literature cited 27

List of Tables

Table 1. F_{ST} results for population comparisons 24

Table 2. Characteristics of 7 microsatellite DNA loci 24

List of Figures

Figure 1. Regional map of brook trout population distribution and current conservation status along Appalachian Mountains 12

Figure 2. Map of sampling locations in southwestern Pennsylvania 17

Figure 3. Allelic frequencies for microsatellite *Sfo*C28 20

Figure 4. Allelic frequencies for microsatellite *Sfo*C24 20

Figure 5. Allelic frequencies for microsatellite *Sfo*C38 21

Figure 6. Allelic frequencies for microsatellite *Sfo*B52 21

Figure 7. Allelic frequencies for microsatellite *Sfo*D75 22

Figure 8. Allelic frequencies for microsatellite *Sfo*C86 22

Figure 9. Allelic frequencies for microsatellite *Sfo*C113 23

Introduction

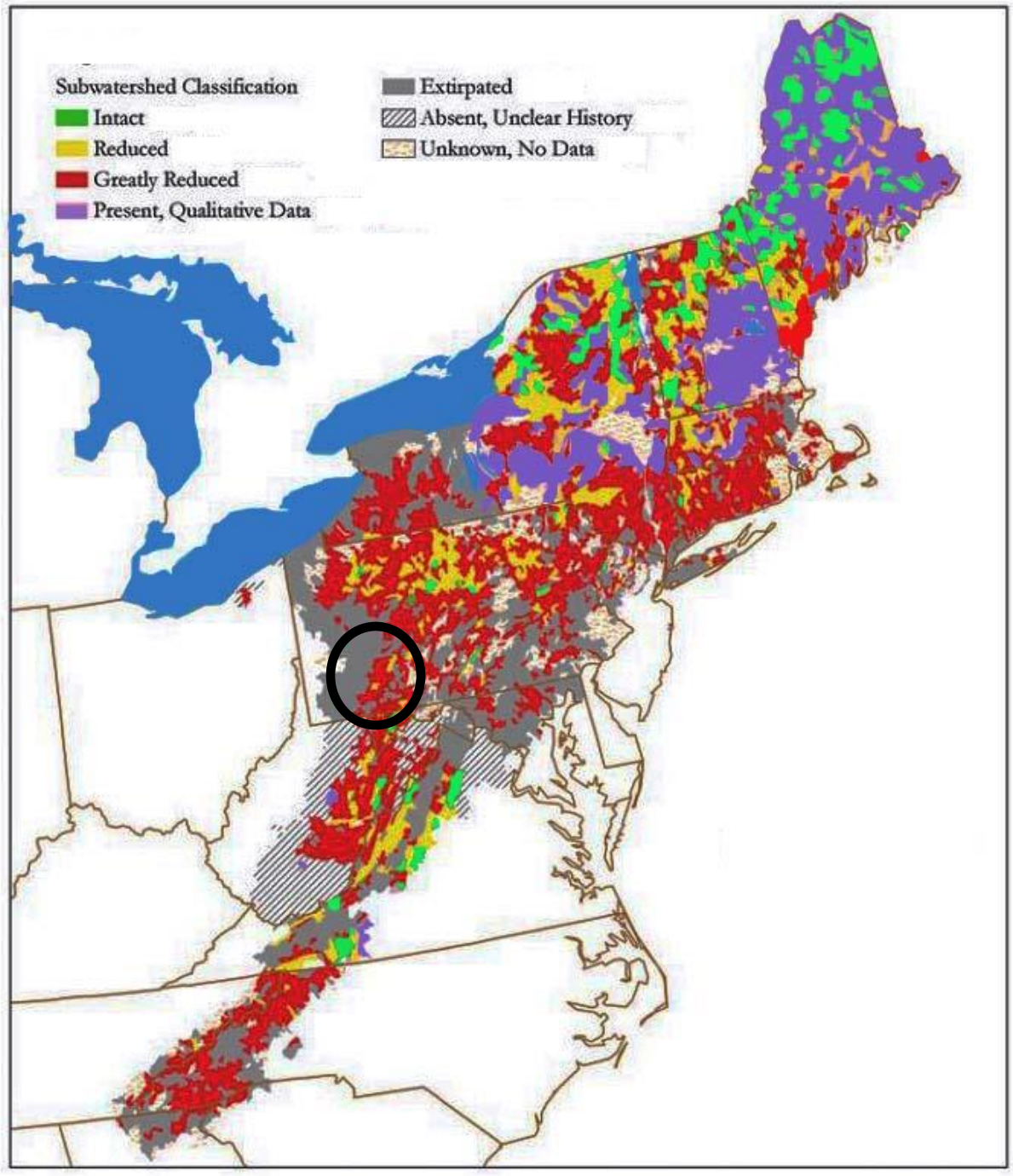
The family *Salmonidae*, or salmonids, contain salmons, trout, chars, freshwater whitefishes, and grayling species (Nelson 1994). Brook trout, *Salvelinus fontinalis*, were once misclassified as trout and the common name trout has remained since. Brook trout are now considered a species of char because of the light colored pattern on a dark background on their scales, lack of teeth on their vomer, and small sized scales (Nelson 1994). True trout species tend to have vomer teeth, dark spots on a light base coloration, and large scales (Nelson 1994). Despite the earlier mistaken identity, brook trout populations are of great interest due to their environmental indicator species status.

As Appalachian Mountain natives, brook trout populations are found historically from Georgia to southern Canada (Sevon et al., 1999). The Appalachian Mountains' cold, clear waters make it an ideal home for brook trout (King et al., 2012; Whiteley et al. 2013). Brook trout are relatively tolerant to changes in water pH and have been documented in a range of 5.5 to 8 pH (a side effect of acid deposition), but migrate when water increases beyond 18°C (Power 1980). Populations of brook trout have been introduced to other locations such as the U.S. Pacific coast states, Alaska, and Hawaii for sport fishing purposes (McPhail and Strouder 1997). With an expanded habitat, other native salmonid species, such as graylings, are declining in those areas due to competition with brook trout (McPhail and Strouder 1997). Brook trout is considered an invasive species in the Pacific states. Nevertheless, there is a decline in population numbers in parts of its native range including the Appalachian Mountains. Other salmonids introduced to the Appalachian Mountains, such as rainbow and brown trout, provide competition to brook trout. These species combined

with the environmental stressors from man-made landscape changes are known contributors to reducing habitat and stream connectivity for brook trout (Whiteley et al., 2013).

The Eastern Brook Trout Joint Venture (EBTJV) has studied historic populations along the U.S. portion of the Appalachian Trail and reported that populations in only 9% of the original watersheds are thriving, most of which reside in Maine (Trout Unlimited 2005). To restore populations, different strategies are being explored. West Virginia is currently researching reducing invasive salmonid populations to lessen competition (Conserving Strategy 2011). Abandoned mine drainage, which increases acid deposition, is also a recurring problem that harms brook trout populations in West Virginia as well (Conserving Strategy 2011). Tennessee, North Carolina, and Virginia aim to improve water quality and acquire more land to conserve current populations (Conserving Strategy 2011). Virginia is also increasing vegetation along watersheds and sub-watersheds to improve nutrient supply and help with the buffering supply (Conserving Strategy 2011). Without these conservation efforts, brook trout populations in the Appalachian Mountains, are expected to suffer a 77% habitat decline by the end of the 21st century (Wegner et al., 2011). The current status of brook trout populations is displayed in Figure 1.

Figure 1: Regional map of brook trout population distribution and current status in Appalachian Mountain Range. The Pennsylvania area of interest for this study is circled in black. Adapted from Trout Unlimited, 2005.



Brook trout were added Pennsylvania's State Wildlife Action Plan in 2005 and continue to be monitored (Argent and Kimmel 2013). In southeastern Pennsylvania, population numbers have declined dramatically due to a variety of environmental stressors such as mining, logging, water withdrawal, and acid deposition (Power 1980). Disturbances to the surrounding environment and closed canopies tend to affect the quality and temperature of the water. Researchers are concerned with the current environmental state in the Appalachian Mountains due to the recent decline of brook trout as they are environmental indicators. Out of all the native brook trout habitats, Pennsylvania has been declared as the state with the most reduced or extant watersheds, which makes it an important location to study. Populations of brook trout in the Laurel Highlands of Pennsylvania are of concern after a profiles of air and stream temperature were collected and indicated a warming climate (Argent and Kimmel 2013). It is possible that the progression towards warmer mean temperatures in brook trout habitats can decrease gene flow and ultimately lead to genetic isolation (Argent and Kimmel 2013).

A key goal of conservation biology is to maintain genetic diversity (Pilgrim et al. 2012). Genetic diversity is as a critical measure of population health (lack of inbreeding). Brook trout are also of economic, cultural, and recreational value as they are the state fish of Pennsylvania, North Carolina, Maine and other states. These fish serve as a major income for fisheries and hatcheries in Pennsylvania, with a 35% share of the total season catch commercially (Argent and Kimmel 2013). It is integral to understand the genetic variation of brook trout to better preserve their diversity and thus conserve the species (Pilgrim et al. 2012).

Genetic analyses of brook trout are needed to develop conservation practices for this species. To study gene flow and the effects of temperature shifts, microsatellites markers are useful. In this study, I use analyses of microsatellite markers to describe patterns of genetic variation in brook trout. Microsatellites are short tandem nucleotide repeats within the genome (Mburu and Hanotte 2005). They are highly susceptible to mutation and thus suitable for analyzing patterns of regional population variation (Pilgrim et al. 2012). Microsatellite alleles are characterized by the number of repeats, with mutations increasing or decreasing the number of repeated elements (Mburu and Hanotte 2005). By using microsatellites and analyzing their differentiation and F_{st} values, I can observe if there are signs of gene flow or isolation. Maintaining gene flow reinforces the need for more conservation management of this species due to the decline in populations. In order to sustain the brook trout populations, connectivity will need to be determined genetically as morphological differences are not always accurate.

Study Goals

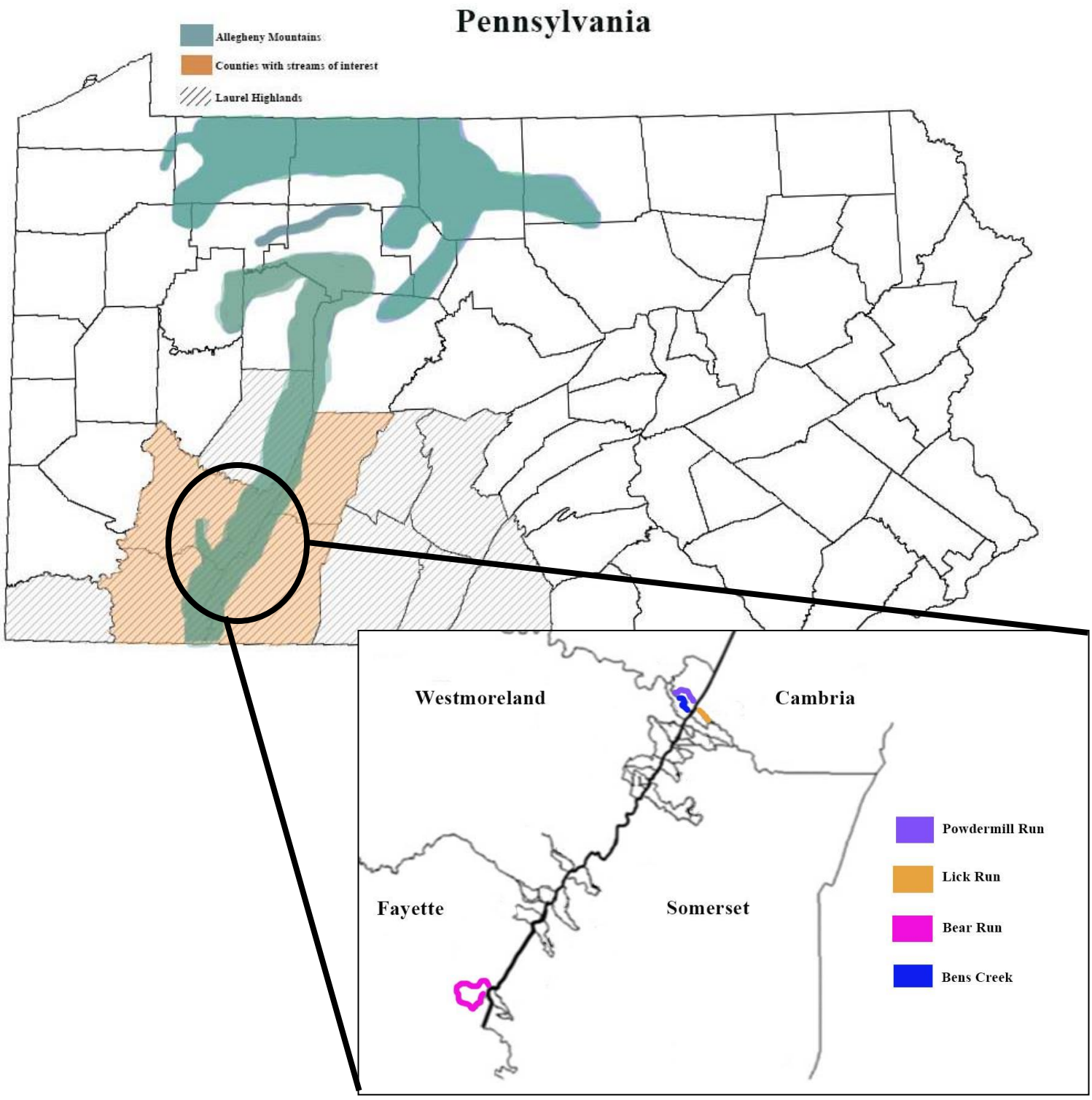
This project aims to conduct a microsatellite analysis of brook trout to determine patterns of genetic diversity and to infer patterns of population connectivity in western Pennsylvania. These data will aid in restoration efforts and allow humans to intervene in repopulating the species in certain watersheds if stream depletion is eventually reached. In addition, diversity sustains a population's ability to evolve in an ever-changing environment. Therefore, the questions we are trying to answer are as follows: 1) How much genetic variation remains in each population? 2) Which of four stream populations sampled are closely related? 3) Are populations sampled from different streams genetically differentiated? And if so, 4) which populations are most similar and/or divergent? The findings of this work will potentially contribute to continued conservation efforts of brook trout populations.

Methods

Fin Clip Sampling and Extraction

This work was conducted in collaboration with Dr. David Argent and Dr. William Kimmel, based at California University of Pennsylvania. Specimens of brook trout were collected by these collaborators from four streams in western Pennsylvania: Baldwin Run (40.35, -79.05), Powdermill Run (40.36, -79.03), the North Branch of Ben's Creek (40.23, -79.04) and Bear Run (39.90, -79.43). The first three of these streams eventually drain into the Conemaugh River and the last drains to the Youghiogheny River (Figure 2). Fin clip samples from 20 individuals per population were collected and used as a source for genomic DNA. DNA was isolated from the fin clips using a DNeasy Blood and Tissue Kit (Qiagen), and the instructions were followed as stated by the manufacturer.

Figure 2: Map of sampling locations in southwestern Pennsylvania. Stream locations are within 4 counties in the Laurel Highlands. The estimated location of the Allegheny Mountain Range, part of the Appalachian Mountain Range, is included for reference. The 4 streams are Powdermill Run, Bens Run, Lick Run, and Bear Run.



PCR and Genotyping

PCR and genotyping were completed using a multiplex approach where 2-3 primer pairs labeled with different fluorescent colors were used per PCR reaction, and the products were run on an Applied Biosystems (ABI) 3710 sequencer. Primers were based off those previously described (King et al 2012). The following primers were multiplexed together: *Sfo*C24 and *Sfo*C28; *Sfo*B52, *Sfo*C38, and *Sfo*D75; *Sfo*C86, *Sfo*C115, and *Sfo*C113. Thermocycler steps were as follows: 1. 3 minutes at 95 ° C; 2. 30 seconds at 95 ° C; 3. 45 seconds at 56 ° C; 4. 2 minutes at 72 ° C; 5. Repeat Step 2, 34 times; 6. 10 minutes at 72 ° C; 7. Hold at 72 ° C. Geneious 6.0.6 software was used to estimate PCR product sizes which were converted into estimates of the number of microsatellite repeats at each locus.

Data Analysis

Genetic diversity statistics (Pi, Theta) were estimated using GENEPOP (version 4.2). The same software was used to test for genetic differentiation among populations using F_{ST} statistics with P<000. To monitor the loci pairs, a log likelihood ratio test for linkage disequilibrium was also completed using the same GENEPOP software. Diversity at each locus was quantified by calculating expected heterozygosity (H_E) and number of alleles (N_A) in GENEPOP as well.

Results

I examined eight microsatellite loci in 64 individual brook trout among 4 streams in western Pennsylvania, 16 per stream. Out of the 8 microsatellites, only 7 were successfully amplified (all but S*fo*C115 with repeat motif $CTCA_{21}$). All of the loci that were genotyped were polymorphic, with between 4 and 17 alleles among them (Table 1). Likelihood ratio tests indicated that two of the seven loci were found to be in significant linkage disequilibrium, *Sfo*B52 and *Sfo*C113.

Allelic frequencies for each microsatellite surveyed are shown in Figures 3-9. Moderate allelic differentiation is present in microsatellites *Sfo*C28, *Sfo*C38, *Sfo*D75, and *Sfo*C113. Slight allelic differentiation is observed in microsatellites *Sfo*C24 and *Sfo*C86. Great allelic differentiation is observed in microsatellite *Sfo*B52. Based on these allele frequencies, the Fixation index, F_{ST}, estimates are presented in Table 1. Population comparisons are as follows: F_{ST} between Bens Creek and Powdermill run is 0.17; between Lick Run and Powdermill Run is 0.14; between Lick Run and Bens Creek is 0.04; between Bear Run and Powdermill Run is 0.29; between Bear Run and Bens Creek is 0.29; and between Bear Run and Lick Run is 0.24. A Fisher's exact test revealed all pairwise F_{ST} values to be statistically significant, $P < 0.001$ indicating significant population substructure.

.

Figure 3: Allelic frequencies for microsatellite *SfoC*28.

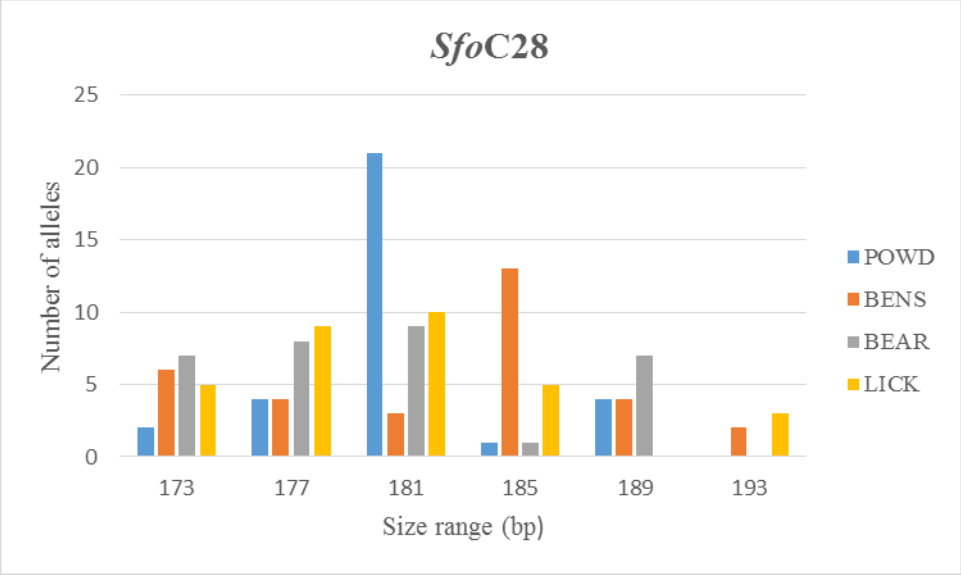


Figure 4: Allelic frequencies for microsatellite *Sfo*C24.

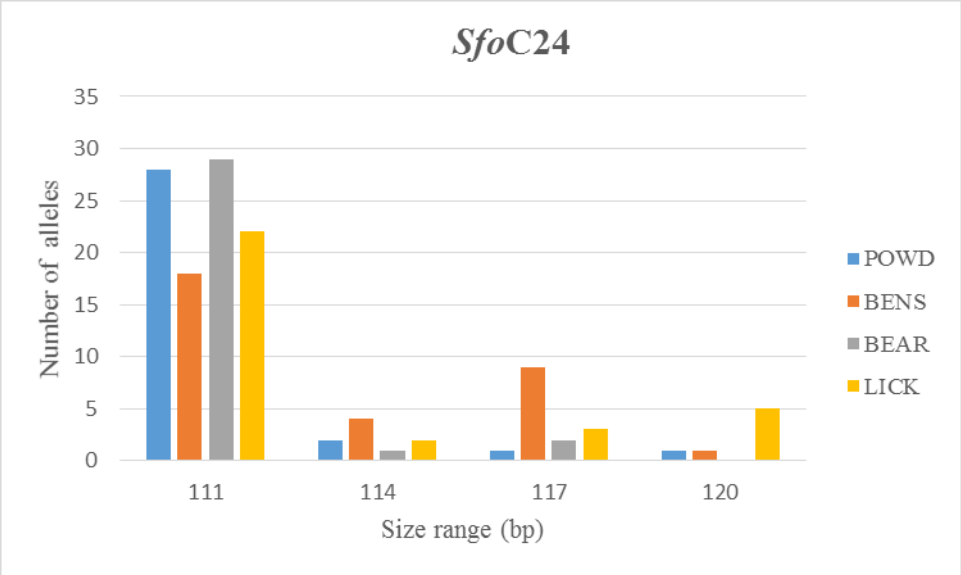


Figure 5: Allelic frequencies for microsatellite *Sfo*C38.

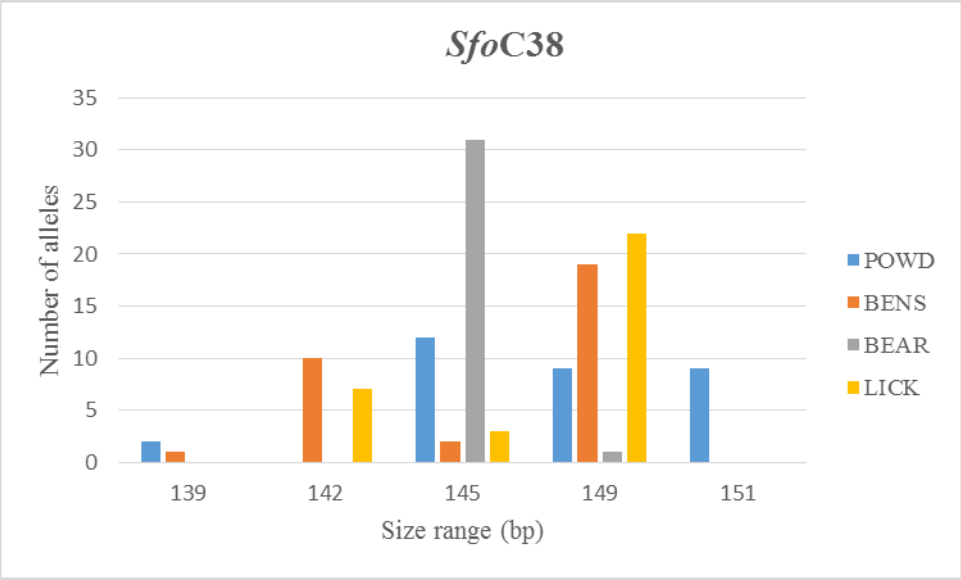


Figure 6: Allelic frequencies for microsatellite *Sfo*B52.

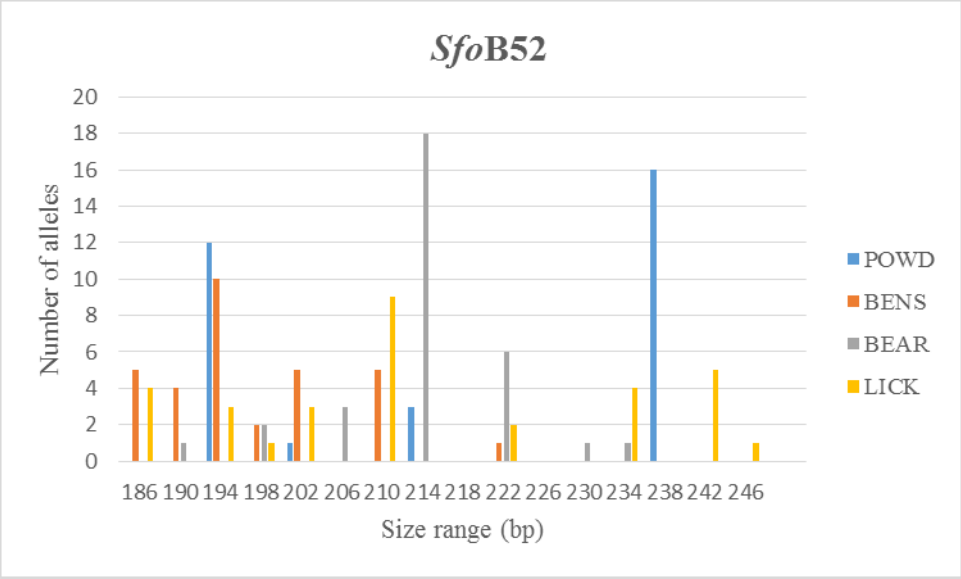


Figure 7: Allelic frequencies for microsatellite *Sfo*D75.

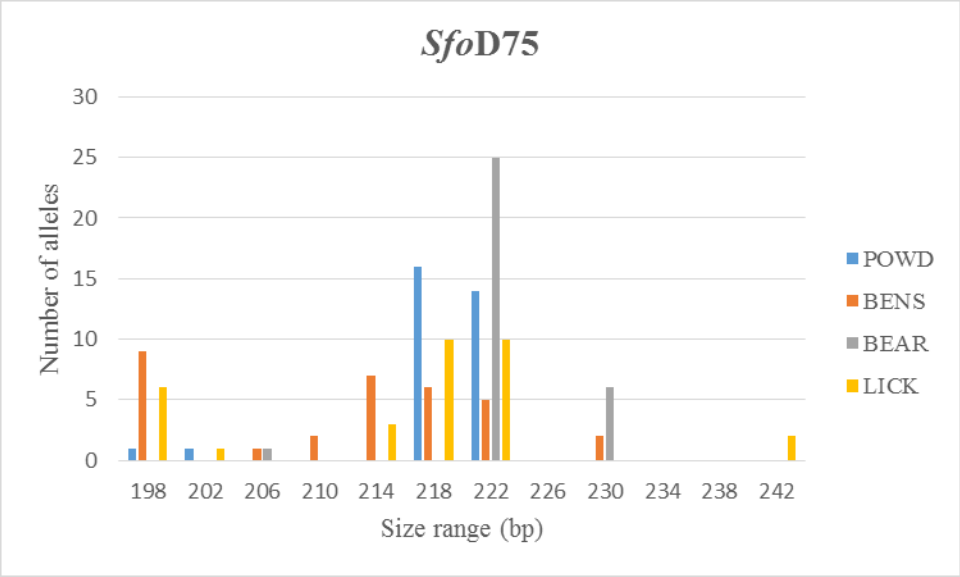


Figure 8: Allelic frequencies for *Sfo*C86.

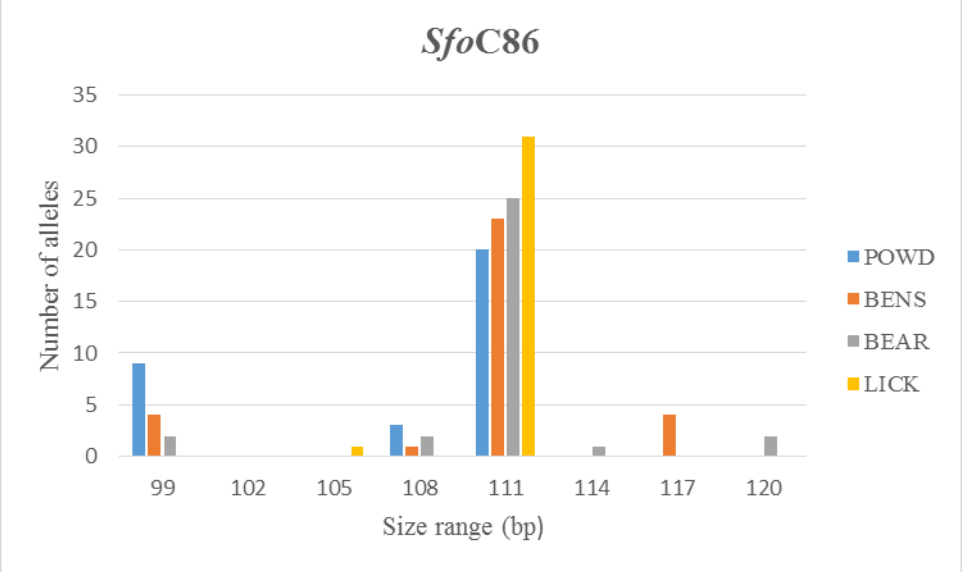


Figure 9: Allelic frequencies for SfoC113.

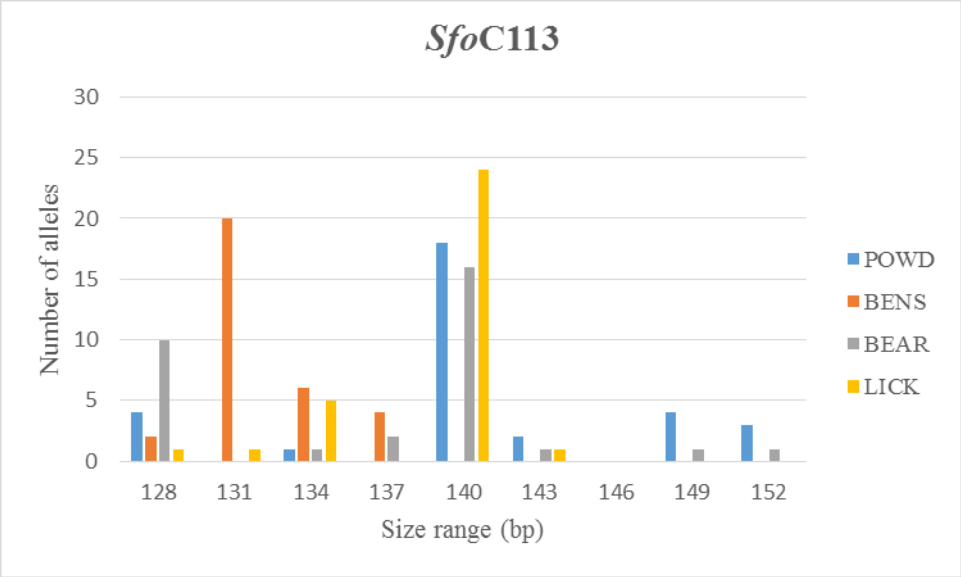


Table 1: F_{ST} results for population comparisons. Significant differentiation is observed between all population pairs as determined by Fisher's exact test, $P < 0.001$.

	POWD	BENS	LICK	BEAR
POWD	~			
BENS	0.17	~		
LICK	0.14	0.04	~	
BEAR	0.29	0.29	0.27	~

Table 2: Characteristics of 7 microsatellite DNA loci. H_O is observed heterozygosity and H_E is expected heterozygosity. N_A is the number of alleles per loci. Adapted from King et al. 2013.

Locus	Repeat Motif	Size Range	Ho	HE	NA
Sfo C28	(GCGT)10	173-193	0.46	0.59	5
Sfo C24	(GAT)10	111-120	0.4	0.42	4
Sfo C38	(GAT)9	139-151	0.75	0.4	17
Sfo B52	(GCGT)12	186-246	0.69	0.58	7
Sfo C86	(GAT)8	99-120	0.6	0.47	9
Sfo D75	(TAGA)17	198-242	0.58	0.68	8
Sfo C113	(GAT)12	128-152	0.36	0.61	7

Discussion

Genetic diversity is important in conservation as it can represent the health of its surrounding environment (Pilgrim et al. 2012). Through this survey of 8 microsatellites, we revealed 7 microsatellites had varying levels of allelic differentiation. Great allelic differentiation is detected in 1 locus, 4 loci show moderate differentiation, and the remaining 2 show slight differentiation. All loci are polymorphic with 4-17 alleles per locus, but the populations are genetically differentiated, which suggests that they are isolated from each other, and may have distinctive features.

As microsatellites are generally selectively neutral, this allowed us examine the history of brook trout populations from the four selected streams. Microsatellites can reveal the history of gene flow and genetic drift (Mburu and Hanotte 2005), which can occur in brook trout populations through migration and isolation, respectively. The relatively high F_{ST} values in our analysis reveals that brook trout populations in western Pennsylvania show some evidence of isolation. The fixation index results displayed moderate divergence (average pairwise divergence F_{ST}= 0.24; range of 0.06-0.30) with highly significant differentiation ($P < 0.001$). Specifically Bear Run, as the most southern stream, shows the highest amount of differentiation when compared to all other stream populations. This suggests, as might be expected, that gene flow is limited between Bear Run to its northern counterparts. Lick Run and Bens Creek have a relatively low pairwise F_{ST} value of 0.04 (across all loci), indicating some gene flow through migration patterns. Future studies should consider sampling the remaining streams in the Pennsylvania Laurel Highlands to further analyze the brook trout isolation patterns between subwatersheds.

Linkage disequilibrium tests presented one Loci 2 and 7 as significant, suggesting that these loci may be close together on a chromosome. All other loci pairs were found not to be significantly linked ($P < 0.01$). Therefore, most of the loci in this study can be considered as independent. However, previous studies have not indicated a similar results between *Sfo*B52 and *Sfo*C113. This result may suggest inbreeding that affected this two markers within this study. I also observed differences between observed heterozygosity (H_O) and expected heterozygosity (H_E). These differences, however, indicate possible deviations from Hardy-Weinberg expectations. Such deviations are expected, however, if populations are not mating randomly as we have observed here.

Genetic diversity within the brook trout populations is integral to the future of their populations as they adapt to the changing environment. Also, noting morphological characteristic in addition to genetic variation may yield some interesting data. Stream characteristics such as gradient, water temperature, and pH coinciding with a genetic analysis would be valuable to hypothesize about migratory patterns, possible barriers, and watershed or sub-watershed quality between sampling locations. It is important to note that microsatellites are noncoding regions in the genome so an experiment aiming towards coding genes may show divergence as well. Divergence in coding genes may provide insights into functional adaptation as well. This would aid in understanding population structure of brook trout and how it is influenced by landscape changes, informing conservation efforts.

Literature cited

Argent, D. and Kimmel, W. (2013). Potential impacts of climate change on brook trout (*Salvenlinus fontinalis*) populations in streams draining the Laurel Hill in Pennsylvania. *Journal of Freshwater Ecology* 28:4, 489-502.

Conservation Strategy/Habitat Work Group. (2011). Conserving the Eastern Brook Trout: Action Strategies. *Eastern Brook Trout Joint Venture*.

King, T.L., Lubinski B.A., Burnham-Curtis, M.K., Scott, W., and Morgan II, R.P. (2012). Tools for the management and conservation in brook trout (*Salvelinus fontinalis*): tri- and tetranucleotide microsatellite markers for the assessment of genetic diversity, phylogeography, and historical demographics. *Conservation Genetic Resources* 4, 539-543.

Mburu, D., and Hanotte, O. (2005). A practical approach to microsatellite genotyping with special reference to livestock population genetics. *IRLI Biodiversity Project*.

McPhail, J.D., and Strouder, D.J. (1997). Pacific Salmon and Their Ecosystems: Status and Future Options. *The Origin and Speciation of* Oncorhynchus. New York, New York.

Nelson, J.S. (1994) Fishes of the world. Third edition. John Wiley & Sons, Inc., New York. 600.

Pilgrim, B.L., Perry, R.C., Keefe, D. G., Perry, E. A., and Marshall, H. D. (2012). Microsatellite variation and genetic structure of brook trout (*Salvelinus fontinalis*) populations in Labrador and neighboring Atlantic Canada; evidence for ongoing gene flow and dual routes of post-Wisconsin colonization. *Ecology and Evolution* 5, 885-898

Power, G. (1980). "The brook charr". *Charrs: salmonid fishes of the genus Salvelinus*. 141–203.

Schwartz, M. K., Luikart, G. and Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution* 22, 25–33.

Sevon, W.D., Fleeger, G.M., and Shepps, V.C. (1999). Pennsylvania and the Ice Age. *Pennsylvania Geological Survey* 2, 6-30.

Trout Unlimited. (2005). Eastern Brook Trout: Status and Threats. *Eastern Brook Trout Joint Venture Report on Eastern Brook Trout: Status and Threats*. <http://easternbrooktrout.org/reports/eastern-brook-trout-status-and-threats/view> Accessed 2017 April 14.

Wenger S.J., Isaak D.J., Luce C.H., Neville, H.M., Fasuch, K.D., Dunham, J.B., Daulwalter, D.C., Young, M.K., Elsner, M.M., Rieman, B.E., Hamlet, A.F., and Williams, J.E. (2011). Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14175-14180.

Whiteley, A.R., Coombs, J.A., Hudy, M., Robinson, Z., Colton, A.R., Nislow, K.H., and Letcher, B.H. (2013). Fragmentation and patch size shape genetic structure of brook trout populations. *Canadian Journal of Fisheries and Aquatic Sciences* 70, 678-688.