

## **Patch-based metrics: A cost effective method for short- and long-term monitoring of EBTJV wild Brook Trout populations?**

Andrew R. Whiteley<sup>1,2</sup>, Mark Hudy<sup>3</sup>, Zachary Robinson<sup>1</sup>, Jason A. Coombs<sup>1,2</sup>, and Keith Nislow<sup>2</sup>

<sup>1</sup>Department of Environmental Conservation, University of Massachusetts, Amherst, MA

<sup>2</sup>U.S. Forest Service, Northern Research Station, University of Massachusetts, Amherst, MA

<sup>3</sup>U.S. Geological Survey, Ecosystems Mission Area, Reston, VA

### **Abstract**

The wild Brook Trout *Salvelinus fontinalis* resource throughout the range of the Eastern Brook Trout Joint Venture (EBTJV) has been significantly reduced over the last 150 years and faces ongoing and future threats from climate change, land use changes, invasive species, and loss of genetic integrity. Monitoring both short- and long-term trends on individual Brook Trout populations and the resource as a whole are important needs of managers. Currently, standard population estimates using mark-recapture and depletion removal estimates are not viable for large scale monitoring because of expense, inability to detect trend (i.e. large coefficient in variation), and problems expanding the sample to the entire population. However, extensive fine-scale occupancy data (at the catchment level) exist for many states. We used this fine-scale catchment data to identify unique “patches” of Brook Trout. We define a “patch” as a group of contiguous catchments occupied by wild Brook Trout. Patches are not connected physically (separated by a dam, unoccupied warm water habitat, downstream invasive species, etc.) and are generally assumed to be genetically isolated. The median patch size from Pennsylvania to the southern range distribution edge is 850 ha and 85.3% of patches were less than 3,000 ha in area. With preliminary patch-level genetic data from Virginia, we found a strong positive relationship between patch size and effective number of breeders ( $N_b$ ), with notable outliers associated with patches that contain reclaimed habitat (positive residuals) and the presence of invasive Rainbow Trout *Oncorhynchus mykiss* (negative residuals). We also found that subsamples from large patches yield similar estimates of genetic metrics, which suggests that our patch-based approach should be applicable even to potentially problematic large patches. We recommend the use of patches for large-scale monitoring of eastern Brook Trout. Recommended patch metrics include: number of patches with allopatric populations (Brook Trout only), number of patches with sympatric nonnative trout populations (Brook Trout with Rainbow Trout or Brown Trout *Salmo trutta*), average size of patches, number of patches increasing in size (connectivity), number of patches decreasing in size, number of patches with decreasing or stable genetic diversity, and number of patches with increasing, decreasing or stable number of effective breeders ( $N_b$ , an indicator of reproductive output and success). A monitoring design combining fixed annual “sentinel” patches and a rotating panel design for other patches has the potential to be a cost effective tool for managers to detect trends in wild Brook Trout populations.

### **Introduction**

The Brook Trout *Salvelinus fontinalis* is a sentinel species that serves as an excellent indicator of headwater ecosystem health in its native range in eastern North America. Multiple anthropogenic stressors have eliminated or severely reduced Brook Trout populations over the last 200 years (Hudy et al. 2008). Monitoring efforts are needed to assess Brook Trout population status. These monitoring efforts should be scale-appropriate and attempt to monitor demographic and genetic contributions to population resilience.

We suggest that the ‘habitat patch’ concept would be highly useful for eastern Brook Trout conservation. A habitat patch for a headwater salmonid is generally defined as a continuous network of thermally suitable habitat (Isaak et al. 2010). Brook Trout in their native range tend to occur in discrete patches of habitat, especially in the southern portion of their native range. Southern Brook Trout populations (approximately from Pennsylvania south) have been more anthropogenically influenced than northern populations (Hudy et al. 2008), and tend to occur in small habitat patches, often isolated by dams (Hudy et al. 2008; Bain and Wine 2010). Even suitable Brook Trout habitat below anthropogenic (dams) or natural (waterfall) barriers may show effects of isolation due to warmwater habitat (Meisner 1990) or downstream invasive species (Moore et al. 1983; Moore et al. 1986; Strange and Habera 1998). Patch size may be a readily obtainable metric with high conservation utility for eastern Brook Trout populations if it is closely related to population persistence and resilience. Further, the patch concept may provide the optimal scale to collect a wide variety of demographic and genetic metrics related to Brook Trout population status.

Genetic metrics collected at the patch scale offer an opportunity to understand historical effects and current demography and may offer an integrative assessment of population status. Two genetic metrics, in particular, can be used across a wide array of taxa and allow inference of two critical components of population resilience. The first indicator, the effective number of breeders ( $N_b$ ) provides information about reproductive output and success (Waples and Do 2010; Hare et al. 2011; Whiteley et al. 2012). This metric combines information from the number of families produced by the parents of a given cohort, the variance in reproductive success among those parents, and early family-dependent survival of the offspring produced (Waples and Do 2010; Christie et al. 2012). Estimates of  $N_b$  can be used to rank population risk and can serve as the foundation for monitoring efforts. The second indicator, genetic diversity (allelic diversity and heterozygosity), provides information about past events such as population bottlenecks that render a population less able to adapt to future conditions (Gienapp et al. 2008; Allendorf et al. 2013). Low genetic diversity is associated with reduced resistance to disease, increased levels of inbreeding, and lower efficacy of natural selection associated with directional and episodic environmental change .

Here, we use extensive fine-scale Brook Trout occupancy data at the catchment level to identify and provide summary statistics for Brook Trout habitat patches for in the southern portion of the species’ native range. We define a patch as a group of contiguous catchments occupied by wild Brook Trout. Patches are not connected physically (separated by a dam, unoccupied warm water habitat, downstream invasive species, etc.) and are generally assumed to be genetically divergent from one another. We then test for a relationship between patch size and both genetic diversity and  $N_b$  in a set of 19 patches from Virginia. Finally, we present a case study of a large patch to demonstrate the application of our patch concept to potentially problematic larger patches.

## **Methods**

### *Habitat patches and sampling*

We defined a patch for Brook Trout as a group of occupied contiguous catchment polygons from the U.S. Geological Survey (USGS) National Hydrology Dataset (NHD) Plus catchment GIS layer (7<sup>th</sup> level, 14-digit hydrologic unit codes (HUCs)). We began a patch in the catchment polygon where a Brook Trout population had been documented based on previous occurrence data and then expanded that patch (dissolving catchment polygon boundaries) to include all catchment polygons upstream until a barrier to fish passage, such as a dam or lake, was encountered or the stream ran dry. Patches above barriers began in the catchment polygon above the reservoir and continued until the stream ran dry or another barrier was reached. Patch

area ) was calculated by summing the area of all catchments contained within that patch. Patches were delineated in the following states: Pennsylvania, New Jersey, Maryland, West Virginia, Virginia, Tennessee, North Carolina, South Carolina, and Georgia.

We developed a sampling protocol designed to obtain samples from patches that yield unbiased estimates of various genetic metrics, with particular focus on  $N_b$  (effective number of breeders). The key consideration is that samples must be close to random with respect to family membership of the individuals collected (Whiteley et al. 2012). Our goal was to spread out sampling locations within a patch for the purpose of reducing full-sibling overrepresentation while increasing family representation to minimize bias in  $N_b$  estimates. To accomplish this, we calculated the length of the main stem within a patch using a geographic information system GIS and divided it into three equal reaches. In larger or irregular patches, supplemental reaches were added, or main stem reaches were replaced by large tributary reaches to enable better spatial sampling of the patch.

To examine genetic variation and the effective number of breeders ( $N_b$ ) within a patch and relate these metrics to patch size, we attempted to sample 25 young of year (YOY) trout from each reach to achieve a target sample size of 75 YOY. Reaches were sampled in a downstream to upstream order. Fin clips were collected and additional data such as presence or absence of invasive trout and the presence of age-1 and older Brook Trout were collected. If YOY were extremely abundant at the start of a reach, the sample was spread over approximately 100 m to prevent family over-representation. If YOY were not found in a reach after sampling 100 m, additional sampling sites within that reach were attempted by dividing reach length by two and spacing the sites approximately 1 to 2 km apart. If fish were not obtained in downstream reaches, upstream reaches were supplemented, if possible, to achieve a goal of at least 50 YOY per patch. A total of 19 patches were sampled, ranging in area from 590 to 10,880 ha.

We exhaustively sampled YOY Brook Trout from a large patch in the Dry River Basin of Rockingham County, Virginia to determine how within-patch differentiation may influence patch-level genetic metrics. We obtained nine samples from five sites within this patch (Table 1). Sampling methods followed Whiteley et al. (2013).

### *Genetic methods*

We used eight microsatellite loci following the procedures of Whiteley et al. (2013). We also followed the procedures of Whiteley et al. (2013) to estimate all population genetic summary statistics, including observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity per locus and population, mean number of alleles ( $A$ ), mean allelic richness per population ( $AR$ ; mean number of alleles scaled to the smallest sample size), and  $N_b$ . Details regarding testing for Hardy-Weinberg proportions and linkage disequilibrium can also be found in Whiteley et al. (2013). We used a linear model to examine the effect of patch area on estimates of genetic variation within patches and  $N_b$ . Response variables in separate models included mean observed number of alleles ( $A_O$ ), mean observed heterozygosity, or point estimates of  $N_b$ . We used a logit transformation for heterozygosity (Warton and Hui 2011). All analyses were performed with the *stats* package in R version 2.15.0 (R Development Core Team 2006).

We followed the procedures of Whiteley et al. (2013) to estimate  $F$ -statistics and test for genic differentiation. We used Meirmans and Hedrick's unbiased estimator  $G''_{ST}$  (Meirmans and Hedrick 2011) for estimates of overall and pairwise  $F'_{ST}$ . We used Nei's unbiased estimator of  $G_{ST}$  (Nei 1987) for estimates of overall and pairwise  $F_{ST}$ . We combined locus-specific exact tests for allele frequency (genic) differentiation with Fisher's method (Ryman 2006). We used the B-Y FDR correction method to control the type I error rate for results from this combined test (Benjamini and Yekutieli 2001; Narum 2006). To minimize any biases associated with family-level structure, we first reconstructed full-sibling families within each sample with COLONY

version 1.2 (Wang 2004) and then randomly sampled one full-sib per family to create a separate data set for tests of genetic differentiation. We also followed the procedures of Whiteley et al. (2013) to test for population-level genetic structure with STRUCTURE ver. 2.3.1 (Pritchard et al. 2000). We performed five runs for each of  $K = 1$  to 4 with and without location as a prior.

## Results

### *Habitat patches*

Brook Trout patch area from Pennsylvania south through Georgia ranged from 21 to 47,766 ha (Figure 1). Mean patch area was 1,854 ha and median patch area was 850 ha. The vast majority of patches were small, for example, 85.3% (2,310 out of 2,708 patches) were less than 3,000 ha in area. If we remove the 202 patches less than 200 ha, which may cause downward bias in measures of central tendency, the mean patch area becomes 1,993 ha (+139 ha), the median becomes 935 ha (+85 ha), and 84% (2,108 of 2,506) of the patches are less than 3,000 ha.

### *Patch-based genetic monitoring*

We observed a strong positive relationship between  $\hat{N}_b$  and patch area for 19 Virginia Brook Trout habitat patches (Figure 2a). Patch area explained 49% of the variation in  $N_b$  ( $t_{17} = 4.0$ ,  $P = 0.0009$ ). While this is a substantial portion of the variation in  $N_b$ , these data indicate that factors other than patch area influence  $N_b$ . Two pH-remediated patches (Hudy et al. 2000) that have low hydrologic variability had high  $N_b$  for their patch area (Figure 2a, triangles). Three patches with invasive Rainbow Trout had low  $N_b$  for their patch area (Figure 2a, squares). We also observed a positive but weaker relationship between patch area and within-population mean heterozygosity (Figure 2b) and mean number of alleles (data not shown). Patch area explained 20% of the variation in heterozygosity ( $t_{17} = 2.1$ ,  $P = 0.05$ ; Figure 2b) and only 6% of the variation in mean number of alleles ( $t_{17} = 1.0$ ,  $P = 0.32$ ).

### *Large patch case study*

Location within the Dry River Basin relative to a dam had a strong influence on within-sample genetic diversity and  $N_b$  (Table 1). The above-dam Dry Run patch (DN-a) had the least genetic variation and the smallest  $N_b$  (Table 1). The above-dam Dry River patch (DV-a) had the largest estimates for genetic diversity ( $H_S$ ,  $A_O$ , and  $AR$ ) but smaller  $N_b$  estimates than below-dam sites. All below dam sites, including the tributary sample, had similar values of genetic diversity and  $N_b$  (Table 1). Two  $N_b$  estimates had upper CI limits that included infinity (Table 1). These are likely due to smaller sample sizes relative to a larger true  $N_b$  and are not likely to be reliable estimates.

We examined genetic divergence among sites by taking a random subsample of one full-sib per family for all analyses. Overall  $F'_{ST}$  in the entire Dry River Basin was 0.331 (0.258-0.408) and overall  $F_{ST}$  was 0.096 (0.068 – 0.132). Pairwise  $F'_{ST}$  ranged from 0.001 to 0.258. Pairwise  $F_{ST}$  ranged from 0.005 to 0.629 (Table 2). Only one of the 28 exact tests for genic differentiation was not significant based on Fisher's method and controlling the FDR with the B-Y correction method ( $\alpha = 0.05$ ; Table 2).

STRUCTURE analyses were consistent with two or three genetic groups in the Dry River patch as a whole. All models clustered the above-dam Dry Run patch (DN-a) separately as one group. All models also clustered the below dam samples (DV-b-main stem a & b, DV-b-main stem b, and DV-b-tributary) together in a single group. There was some discrepancy among models related to the above-dam Dry River (DV-a) samples, which is consistent with the moderate genetic differentiation between DV-a and the below-dam sites (Table 2). With a location prior,  $K = 3$  clearly outperformed other models and split DV-a into a third genetic group

(Fig. 3a). With no location prior, log-likelihoods increased from  $K = 1$  to  $K = 4$  then decreased.  $K = 2$  grouped DN-a separately and the other group contained the remainder of the sites (fig. 3c).  $K = 3$  had an additional group that occurred primarily in DV-a (Fig. 3b).  $K = 4$  was biologically implausible; a fourth group was scattered throughout each of the DV-a and DV-b sites (data not shown).

## Discussion

### *Habitat patches*

The patches we defined appear to be the most appropriate spatial scale for eastern Brook Trout management. Our approach provides a workable method to define the scale at which demographic and genetic monitoring could be performed for this species. Because we focus on both Brook Trout presence and population discontinuities, our approach defines what are likely to be demographically independent units. It is clear that even if there is downward bias due to artificially small patches in our analysis, our results reveal a striking tendency towards small and fragmented populations in the southern portion of their native eastern Brook Trout range. It is highly likely that patches of approximately 3,000 ha or less contain single populations, therefore roughly 85% of Brook Trout patches are likely to contain single populations that may suffer from the variety of well-described small population effects (Gienapp et al. 2008). Another possible weakness of our approach occurs with very large patches, which may contain metapopulations and multiple demographically independent subpopulations (see below). Our patch-based approach appears to be highly effective for the vast majority of the southern portion of the Brook Trout eastern range, application to more intact and potentially continuous northern habitat may present a challenge that we are currently working to address.

### *Patch-based genetic monitoring*

The strong positive relationship between  $N_b$  and patch area indicates that patch area alone serves as an important driver of reproductive success and output in this set of patches. Our results also provide some indication regarding additional factors that influence among-patch variation in  $N_b$ . Two pH-remediated patches (Hudy et al. 2000) that have low hydrologic variability had relatively large  $\hat{N}_b$  for their patch area. A likely cause for this is the high proportion of available quality habitat compared to patches with lower productivity or increased flow variability. On the other hand, the three patches that contained invasive Rainbow Trout and had relatively low  $\hat{N}_b$  for their patch area are likely to have a low proportion of available quality habitat due to displacement.

The weaker relationship between within-patch genetic diversity and patch area indicates that factors other than patch area influence the maintenance of genetic diversity and evolutionary potential within patches. There was a cluster of patches with low heterozygosity for their patch area and another cluster of patches with high heterozygosity for their patch area. Patches with low genetic diversity for their size may have undergone bottlenecks in the past. Aspects of hydrological variability, fragmentation effects, and presence of invasive species could have caused bottlenecks. Patches with high genetic diversity for their size are likely to have maintained population sizes above levels that lead to precipitous loss of genetic variation. These patches are likely to be relatively un-fragmented, have high habitat quality throughout a greater proportion of the patch, and be less influenced by invasive species.

These preliminary results indicate that factors beyond patch size must be considered if we are to comprehensively model Brook Trout genetic indicators of resilience. In addition to patch size, important drivers of variation in genetic indicators of resilience are likely to be hydrologic variability, stream temperature, fragmentation, and presence of invasive species. We are

currently working to expand our analysis to include a greater number of patches that vary more widely in ecological characteristics. A critical next step is to understand how environmental and landscape factors influence genetic metrics and to integrate them with ecological models to improve forecasts of future population status.

### *Large patch case study*

Larger patches may present a difficulty for our patch-based genetic monitoring approach. Large patches are more difficult to effectively sample and smaller subsamples may yield biased estimates of genetic metrics. The rarity of large patches reveals that these potential problems only apply to a very limited portion of eastern Brook Trout populations. Furthermore, our large patch case study provides promising results that suggest that subsamples of large patches should be representative of the patch as a whole. In the large below-dam patch, we sampled two main stem sites and a tributary site closer to the headwater limits of the patch. We observed statistically significant allele frequency divergence among the tributary site and the main stem sites. However, all STRUCTURE models grouped the below-dam sites together. These results suggest that metapopulation structure occurs below the dam and that populations may be demographically independent. However, the important point relative to our patch sampling approach is that values of genetic diversity and  $N_b$  were similar within all below-dam sites. Therefore, if we had sampled only a small portion of this large patch, we would have obtained representative values of the focal genetic metrics.

The likely evolutionary mechanism responsible for the similarity of values of genetic diversity for the within-patch subpopulations is gene flow. Gene flow among multiple genetically differentiated populations in larger patches would maintain genetic diversity within each of the subpopulations (Jorde and Ryman 1996). The mechanistic explanation for the similarity in  $N_b$  estimates among sites within the large below-dam patch is less clear. The DV-b-tributary, a headwater tributary site that is connected to the below-dam metapopulation, had similar  $\hat{N}_b$  values to the other below-dam sites and much greater  $\hat{N}_b$  than the adjacent but above-dam DN-a. Small amounts of gene flow are less likely to influence the LD signal (relative to genetic diversity) within each subpopulation (Jorde and Ryman 1996; Palm et al. 2003). Therefore, gene flow may not provide a mechanistic explanation of the similarity of  $N_b$  estimates in the below-dam patch estimates of  $N_b$ . We previously suggested that consistently lower  $\hat{N}_b$  in above-dam relative to below-dam patches in a larger series of Brook Trout patches was due to limited spawning habitat in above-dam patches (Whiteley et al. 2013). We predict that greater spawning site availability in the connected stream (DV-b-tributary) relative to the two isolated patches (DN-a and DV-a) is the primary cause of these differences.

We recommend a patch-based monitoring program for eastern Brook Trout population status. The program could center on patches defined as we describe here. We recommend the following patch metrics: number of patches with allopatric populations (Brook Trout only), number of patches with sympatric nonnative trout populations (Brook Trout with Rainbow Trout or Brown Trout), average size of patches, number of patches increasing in size (connectivity), number of patches decreasing in size, number of patches with decreasing or stable genetic diversity, and number of patches with increasing, decreasing or stable  $N_b$  (as an indicator of reproductive output and success). The program could focus on a set of sites that are visited every year (sentinel sites) and other sets of sites that are visited in a rotating manner such that each set would be visited once every 5 years. It might be possible to use future estimates of  $N_b$  from these sites, obtained with the appropriate sampling strategy (Whiteley et al. 2012), to monitor population trend (Tallmon et al. 2010).

## Acknowledgements

M. Burak and M. Page collected genetic data. C. Anderson, J. Courtwright, A. Fitzgerald, C. Kyger, H. Mohn, B. Stanley, and B. Trumbo helped with field collections. The following organizations provided financial assistance or volunteer support: James Madison University, George Washington and Jefferson National Forest; Virginia Department of Game and Inland Fisheries; U.S. Forest Service, Northern Research Station; University of Massachusetts Amherst; and U.S. Geological Survey Conte Anadromous Fish Research Laboratory.

## References Cited

- Allendorf, F. W., G. Luikart, and S. N. Aitken. 2013. Conservation and the Genetics of Populations. John Wiley & Sons, West Sussex, UK.
- Bain, M. B., and M. L. Wine. 2010. Testing predictions of stream landscape theory for fish assemblages in highly fragmented watersheds. *Folia Zoologica* 59:231-239.
- Christie, M. R., M. L. Marine, R. A. French, R. S. Waples, and M. S. Blouin. 2012. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity* 109:254-260.
- Gienapp, P., C. Teplitsky, J. S. Alho, J. A. Mills, and J. Merila. 2008. Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology* 17:167-178.
- Hare, M. P., L. Nunney, M. K. Schwartz, D. E. Ruzzante, M. Burford, R. S. Waples, K. Ruegg, and F. Palstra. 2011. Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology* 25:438-449.
- Hudy, M., D. M. Downey, and D. W. Bowman. 2000. Successful restoration of an acidified native brook trout stream through mitigation with limestone sand. *North American Journal of Fisheries Management* 20:453-466.
- Hudy, M., T. M. Theiling, N. Gillespie, and E. P. Smith. 2008. Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the eastern United States. *North American Journal of Fisheries Management* 28:1069-1085.
- Isaak, D. J., C. H. Luce, B. E. Rieman, D. E. Nagel, E. E. Peterson, D. L. Horan, S. Parkes, and G. L. Chandler. 2010. Effects of climate change and wildfire on stream temperatures and salmonid thermal habitat in a mountain river network. *Ecological Applications* 20:1350-1371.
- Jorde, P. E., and N. Ryman. 1996. Demographic genetics of brown trout (*Salmo trutta*) and estimation of effective population size from temporal change of allele frequencies. *Genetics* 143:1369-1381.
- Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: F-ST and related measures. *Molecular Ecology Resources* 11:5-18.
- Meisner, J. D. 1990. Effect of climatic warming on the southern margins of the native range of brook trout, *Salvelinus fontinalis*. *Canadian Journal of Fisheries and Aquatic Sciences* 47:1065-1070.
- Moore, S. E., G. L. Larson, and B. Ridley. 1986. Population control of exotic rainbow trout in streams of a natural area park. *Environmental Management* 10:215-219.
- Moore, S. E., B. Ridley, and G. L. Larson. 1983. Standing crops of brook trout concurrent with removal of rainbow trout from selected streams in Great Smoky Mountains National Park. *North American Journal of Fisheries Management* 3:72-80.
- Palm, S., L. Laikre, P. E. Jorde, and N. Ryman. 2003. Effective population size and temporal genetic change in stream resident brown trout (*Salmo trutta*, L.). *Conservation Genetics* 4:249-264.

- Pritchard, K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- R Development Core Team. 2006. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Strange, R. J., and J. W. Habera. 1998. No net loss of brook trout distribution in areas of sympatry with rainbow trout in Tennessee streams. *Transactions of the American Fisheries Society* 127:434-440.
- Tallmon, D. A., D. Gregovich, R. S. Waples, C. S. Baker, J. Jackson, B. L. Taylor, E. Archer, K. K. Martien, F. W. Allendorf, and M. K. Schwartz. 2010. When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Molecular Ecology Resources* 10:684-692.
- Waples, R. S., and C. Do. 2010. Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* 3:244-262.
- Warton, D. I., and F. K. C. Hui. 2011. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* 92:3-10.
- Whiteley, A. R., J. A. Coombs, M. Hudy, Z. Robinson, A. R. Colton, K. H. Nislow, and B. H. Letcher. 2013. Fragmentation and patch size shape genetic structure of brook trout populations. *Canadian Journal Fisheries and Aquatic Sciences* online early:
- Whiteley, A. R., J. A. Coombs, M. Hudy, Z. Robinson, K. H. Nislow, and B. H. Letcher. 2012. Sampling strategies for estimating brook trout effective population size. *Conservation Genetics* 13:625-637.

## Figure Legends

Figure 1. Histogram of patch area (in hectares ) for eastern Brook Trout patches from Pennsylvania to Georgia. Patches were defined with extensive Brook Trout occupancy data and the approach described in the text.

Fig. 2. Effective number of breeders ( $N_b$ ; panel A) and genetic diversity (mean within-population heterozygosity; panel B) regressed against patch area (ha) for 19 habitat patches of Brook Trout in Virginia. Habitat patches that have been pH-remediated and have stable flow patterns relative to all others are shown as triangles. Habitat patches that contain invasive Rainbow Trout are shown as squares. All other sites lack these characteristics.  $R^2 = 0.49$  for the regression in panel A.  $R^2 = 0.20$  for the regression in panel B.

Fig. 3. STRUCTURE analysis of large Dry River case study Brook Trout patch from Virginia. This entire patch is 15,904 ha but was divided according to our methodology into a large downstream patch (DV-b, 10,880 ha), a small above-dam patch (DN-a; 1,217 ha), and a large above-dam patch (DV-a, 3,807 ha). Sample locations are shown as black dots. STRUCTURE plots show the proportion of the genome ( $Q$ ) of each individual assigned to each population sample. One full-sibling was randomly chosen from each family for all analyses. Shown in (a) is the best-supported STRUCTURE admixture model with a location prior. Shown in (b)  $K = 3$ , and in (c)  $K = 2$ ; both with an admixture model and no location prior. Each row corresponds to an individual and sample sites are separated by horizontal bars. Each of the clusters was given a color that corresponds to the colors in the map. The below-barrier shade of gray was used for the combined DV-a and DV-b cluster in (c).

Table 1. Genetic summary statistics for young-of-the-year (YOY) Brook Trout (captured in five sites in the Dry River Basin, Virginia. In the site names, the initial ‘a’ corresponds to above-dam and the initial ‘b’ corresponds to below-dam. The year of sampling is shown for each sample. There were two main stem Dry River below-dam sites, labeled ‘1’ and ‘2’. Summary statistics are based on all individuals sampled ( $N$ ). Measures are as follows: mean expected heterozygosity ( $H_S$ ), mean number of observed alleles per population ( $A_O$ ), allelic richness standardized to  $N = 27$  ( $AR$ ), and LDNe-based single-sample estimates of the effective number of breeders that gave rise to the cohort examined ( $N_b$ ).

Site name	Patch area (ha)	$N$	$H_S$	$A_O$	$AR$	$N_b$
DV-b-mainstem-1-2010	10,880	99	0.777	8.8	7.8	191.2 (140.3-279.8)
DV-b-mainstem-1-2011	10,880	67	0.770	9.3	7.9	146.4 (100.9-238.8)
DV-b-mainstem-2-2010	10,880	44	0.774	9.4	8.5	375.5 (174.2-INF)
DV-b-tributary-2012	10,880	86	0.761	10.3	8.2	125.2 (90.1-186.3)
DV-a-2010	3,807	379	0.780	10.9	8.2	66.6 (57.8-76.5)
DV-a-2011	3,807	510	0.797	11.6	8.8	75.0 (60.9-91.4)
DN-a-2010	1,217	46	0.565	3.4	3.4	4.9 (3.8-8.7)
DN-a-2011	1,217	27	0.392	2.8	2.8	40.2 (12.6-INF)

Table 2. Genetic differentiation among eight samples of YOY Brook Trout from five sites in the Dry River Basin, Virginia.  $F_{ST}$  is above the diagonal,  $F'_{ST}$  is below the diagonal. Bold values were not significant following Fisher’s method for combining  $P$ -values across eight exact tests per sample (B-Y FDR correction, nominal  $P = 0.013$ ).

Site	DVb- mainstem- 1-2010	DVb- mainstem- 1-2011	DVb- mainstem- 2-2010	DVb- tributary- 2012	DVa-2010	DVa-2011	DNa-2010	DNa-2011
DV-b-mainstem-a-2010	--	<b>0.001</b>	0.005	0.031	0.020	0.018	0.150	0.230
DV-b-mainstem-a-2011	<b>0.005</b>	--	0.003	0.020	0.014	0.014	0.141	0.230
DV-b-mainstem-b-2010	0.023	0.014	--	0.021	0.017	0.012	0.147	0.240
DV-b-tributary-2012	0.135	0.090	0.091	--	0.029	0.026	0.158	0.258
DV-a-2010	0.095	0.065	0.079	0.130	--	0.002	0.153	0.245
DV-a-2011	0.084	0.066	0.055	0.119	0.011	--	0.142	0.244
DN-a-2010	0.462	0.438	0.454	0.477	0.484	0.449	--	0.155
DN-a-2011	0.570	0.570	0.595	0.629	0.620	0.618	0.304	--

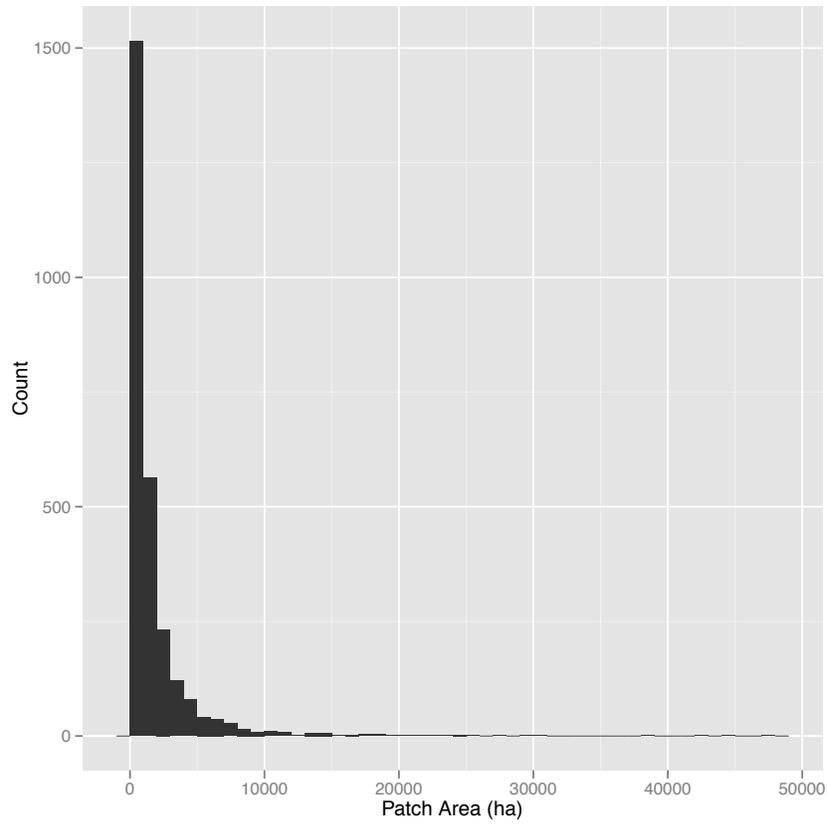


Fig. 1

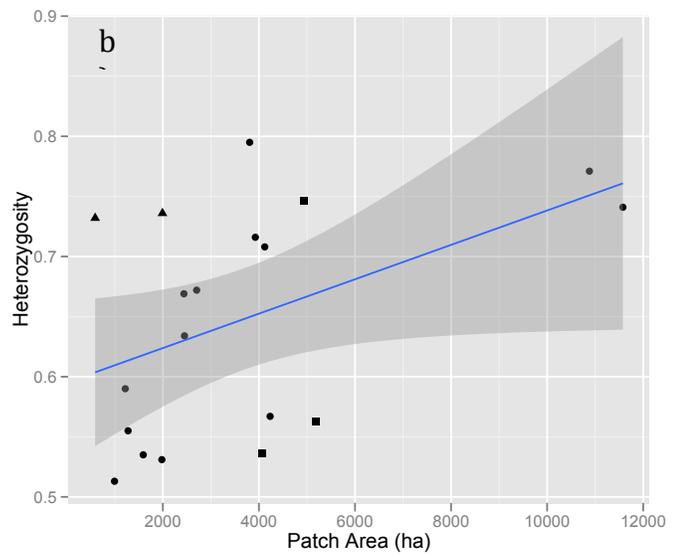
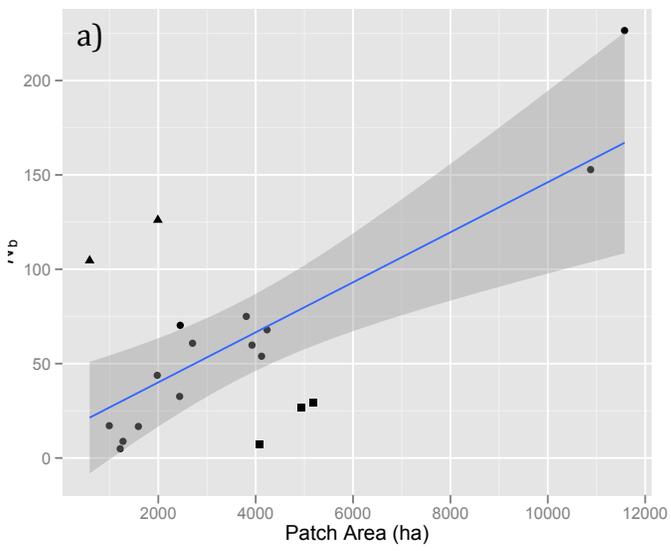


Fig. 2

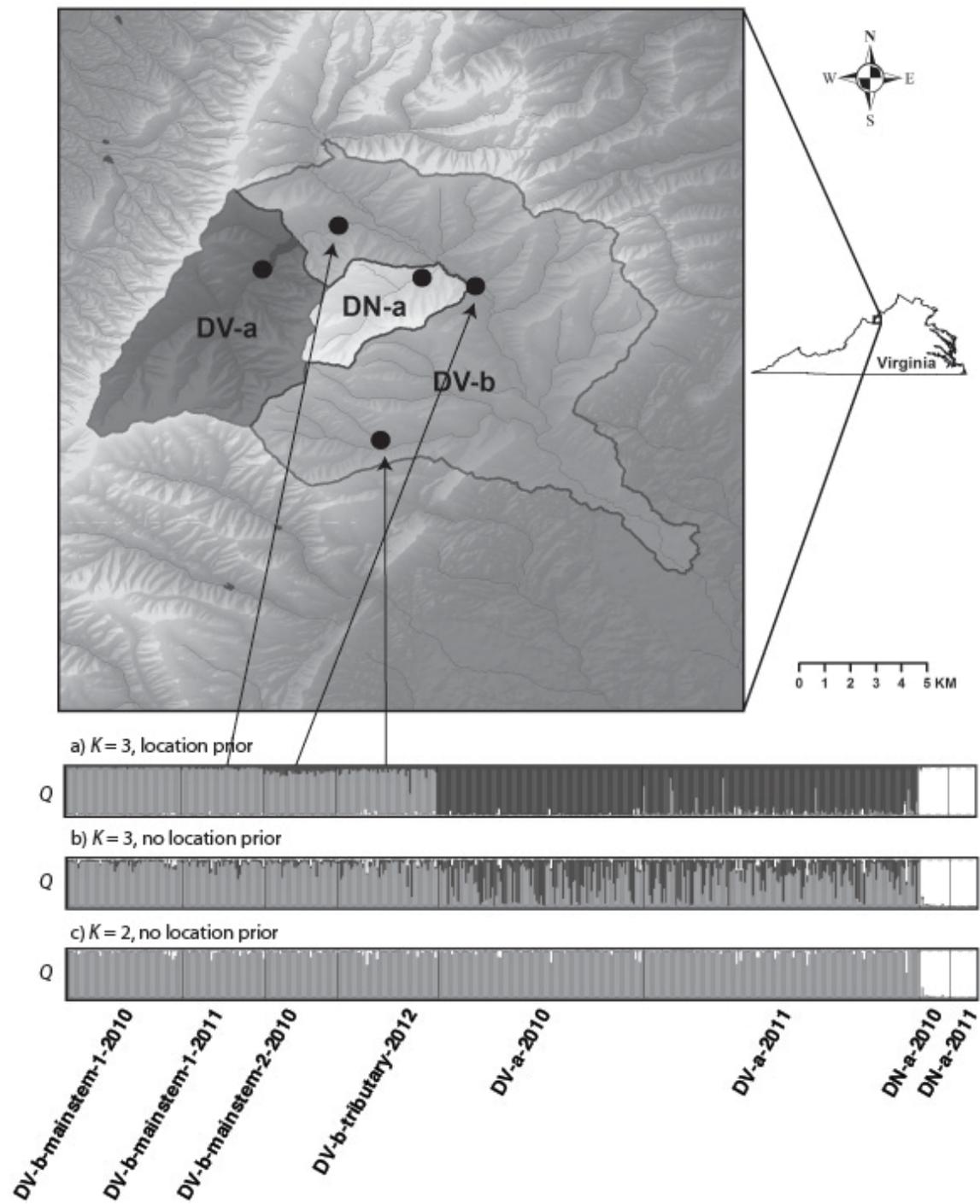


Fig. 3