

MS. SHANNON WHITE (Orcid ID : 0000-0003-4687-6596)

Article type : Original Article

Limited hatchery introgression into wild brook trout (*Salvelinus fontinalis*) populations despite reoccurring stocking

Shannon L. White<sup>a,b,\*</sup>, William L. Miller<sup>a,b</sup>, Stephanie A. Dowell<sup>c</sup>, Meredith L. Bartron<sup>c</sup>, and Tyler Wagner<sup>d</sup>

<sup>a</sup> Pennsylvania Cooperative Fish and Wildlife Research Unit, Pennsylvania State University, University Park, USA

<sup>b</sup> Department of Ecosystem Science and Management, Pennsylvania State University, University Park, USA

<sup>c</sup> U.S. Fish and Wildlife Service, Northeast Fishery Center, 227 Washington Avenue, Lamar, Pennsylvania 16848, USA

<sup>d</sup> U.S. Geological Survey, Pennsylvania Cooperative Fish and Wildlife Research Unit, Pennsylvania State University, University Park, USA

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/eva.12646

This article is protected by copyright. All rights reserved.

\*Corresponding author: slw361@psu.edu, 413 Forest Resources Building, University Park, Pennsylvania 16802, USA, Tel: +1-804-387-3498, Fax: +1-814-863-4710

Due to increased anthropogenic pressures on many fish populations, supplementing wild populations with captive-raised individuals has become an increasingly common management practice. Stocking programs can be controversial due to uncertainty about the long-term fitness effects of genetic introgression on wild populations. In particular, introgression between hatchery and wild individuals can cause declines in wild population fitness, resiliency, and adaptive potential, and contribute to local population extirpation. However, low survival and fitness of captive-raised individuals can minimize the long-term genetic consequences of stocking in wild populations, and to date the prevalence of introgression in actively stocked ecosystems has not been rigorously evaluated. We quantified the extent of introgression in 30 populations of wild brook trout (*Salvelinus fontinalis*) in a Pennsylvania watershed, and examined the correlation between introgression and 11 environmental covariates. Genetic assignment tests were used to determine the origin (wild vs. captive-raised) for 1742 wild-caught and 300 hatchery brook trout. To avoid assignment biases, individuals were assigned to two simulated populations that represented the average allele frequencies in wild and hatchery groups. Fish with intermediate probabilities of wild ancestry were classified as introgressed, with threshold values determined through simulation. Even with reoccurring stocking at most sites, over 93% of wild-caught individuals probabilistically assigned to wild origin, and only 5.6% of wild-caught fish assigned to introgressed. Models examining environmental drivers of introgression explained less than 3% of the among-population variability, and all estimated effects were highly uncertain. This was not surprising given overall low introgression observed in this study. Our results suggest that introgression of hatchery-derived genotypes can occur at low rates, even in actively stocked ecosystems and across a range of

habitats. However, a cautious approach to stocking may still be warranted, as the potential effects of stocking on wild population fitness and the mechanisms limiting introgression are not known.

Keywords: Introgression, captive stocking, brook trout, *Salvelinus fontinalis*

## Introduction

Supplementation of wild populations with captive-raised individuals is an increasingly important management strategy for species of social, commercial and recreational value (Araki et al. 2007; Naish et al. 2007; Stowell et al. 2015). With continued population declines in many fish species due to habitat loss, climate change, and historic overharvest, stocking programs have expanded to meet growing recreational demands and conservation goals (Araki & Schmid, 2010). Recreational stocking can provide an immediate increase in the size of the harvestable population and help to maintain populations that were formally sustained through natural reproduction. The maintenance of recreational opportunities for many species now largely relies on stocking enhancement programs (Askey et al. 2013), with some naturally reproducing populations comprised almost exclusively of individuals with captive ancestry (Evans & Willox, 1991; Ford et al. 2006).

Once released, captive fish, which tend to be larger and more aggressive than their wild counterparts (Huntingford, 2004), can prey upon juveniles from wild populations, compete for food, and usurp high-quality habitats (Naish et al. 2007). However, captive fish often have low reproductive success, with relative fitness often less than half that of wild counterparts (Araki et al. 2008; Christie et al. 2014), and may be quickly removed from the environment through dispersal,

harvest, or natural mortality (Baer et al. 2007). As such, recreational fish stocking can sometimes have small or no detectable effect on native fish communities (Weaver & Kwak 2013).

Nonetheless, even with the potential for low survival and reproduction, recreational stocking still remains controversial due to the long-term negative effects that captive individuals can have on wild populations (Utter, 2003; Weber & Fausch, 2003; Wollebæk et al. 2010). In particular, successful interbreeding between hatchery and wild individuals, defined here as introgression, has the potential to threaten long-term wild population viability by altering patterns of genetic diversity (Reisenbichler & McIntyre 1977; Waples, 1991; Utter, 2003). Inbreeding, genetic drift, and unintentional and/or artificial selection in captive populations can increase the prevalence of genotypes with commercially valuable traits and alter allelic frequencies relative to wild populations. The phenotypes of captive individuals can confer a reduction in fitness relative to wild populations by altering the behavior, morphology, physiology, and timing of life history events (Ford et al. 2006; Naish et al. 2013). Consequently, captive individuals, and their offspring, may not be successful in natural environments (Saikkonen et al. 2011; Stringwell et al. 2014; but see Allendorf et al. 2004 for a counterexample and discussion of heterosis).

Introgression can rapidly (*i.e.*, in as little as one generation; Muhlfeld et al. 2009) modify wild population genetic diversity (Ryman & Laikre 1991; Bowman et al. 2017), disrupt locally adapted gene complexes (Hallerman, 2003; Naish et al. 2007), homogenize genetic structure (Hindar et al. 1991; Marie et al. 2010), introduce maladaptive phenotypes into a population (Bolstad et al. 2017), disrupt expression of biologically relevant genes (Lamaze et al. 2013), and increase disease susceptibility (Currrens et al. 1997). These genetic consequences of introgression, combined with reduced survival, reproduction, and competitive ability of introgressed offspring, can compromise

population resiliency and future adaptive potential by reducing population sizes and eroding among-population genetic variability (Tufto, 2017). Small, isolated populations may be particularly threatened by hatchery introgression due to limited gene flow (Lynch & O’Hely, 2001; Ford, 2002), which can increase the long-term prevalence of domestic genotypes in wild populations.

Negative genetic consequences of stocking on wild populations have been documented in many fish species, particularly in commonly stocked species such as salmonids (Wollebæk et al. 2010). However, most studies to date that quantify hatchery introgression into wild populations have focused primarily on accidental releases from captive facilities and legacy effects of terminated stocking programs (see Araki & Schmid, 2010 for a review, and Kazyak et al. 2018 for a large-scale study on brook trout, *Salvelinus fontinalis*). Few studies have provided empirical estimates for introgression in populations managed with relatively long-term (>10 years, in many cases) and recurring stocking events focused on population supplementation (but see Valiquette et al. (2014) and Létourneau et al. (2017) for examples of lacustrine lake (*Salvelinus namaycush*) and brook trout populations). With repeated exposure to captive individuals, this common management scenario is predicted to induce high rates of introgression (Fleming & Petersson, 2001), particularly as the intensity and duration of stocking increases (Araguas et al. 2004). However, failure of stocked individuals to survive and reproduce, which may vary as a function of local and regional habitat features, could minimize rates of captive introgression. In this case, the genetic effects of stocking could vary considerably across the landscape. Accordingly, a better empirical understanding of the relative rates of introgression in multiple stocked populations may help assess the potential long-term effects of recreational stocking programs.

We quantified the degree of captive-stock introgression in 30 wild brook trout populations and developed models to explore the correlation between introgression and local and regional habitat characteristics. Brook trout is a species of recreational value and high conservation concern throughout its native range in eastern North America. Despite ongoing efforts by state and federal agencies to restore populations, brook trout population size and geographic range have declined significantly in recent decades. Declines are largely due to instream habitat loss, deforestation, nonnative species introductions, and climate change (Hudy et al. 2008).

Because brook trout is a popular sport fishery in Pennsylvania, approximately 800,000 adult brook trout are stocked annually by the Pennsylvania Fish and Boat Commission (PFBC) and PFBC cooperative nurseries to increase recreational angling opportunities. Stocking is sometimes implemented on streams with existing wild brook trout populations and on streams adjacent to wild populations. As such, we predicted that introgression may occur in streams that are directly stocked and, because brook trout are capable of moving in excess of 10 km (Davis et al. 2015), we also predicted introgression could occur in tributaries to stocked streams. Quantifying potential indirect effects of stocking is important because considerations for how hatchery stocking can influence wild populations are typically restricted to the stream of direct stocking efforts. However, movement of stocked individuals into nearby wild populations could result in unintended and unforeseen interactions between hatchery and wild trout that extend beyond the spatial scale of direct stocking efforts.

## Methods

### *Study Area*

Loyalsock Creek is a 1,284 km<sup>2</sup> watershed located primarily in Lycoming and Sullivan counties in northcentral Pennsylvania (Figure 1). This watershed is predominantly forested, and wild brook trout inhabit most tributaries to Loyalsock Creek. The mainstem of Loyalsock Creek is a large, warm/coolwater stream that is not thermally suitable for brook trout during summer, but is used seasonally for residence and migration.

The PFBC has maintained a recreational trout stocking program in Loyalsock Creek and select tributaries in the watershed for several decades. Because we sampled fish in 2015 (see *Sample Collection and Genotyping*), we examined stocking records from 2006 to 2015. Over this time period, 21 of our 30 study sites were either directly stocked or within 2 km of a stocking location. In streams that were stocked, stocking occurred once or twice per year from March to May. The average number of adult brook trout stocked per year at each site ranged from 151 to 3491 at an average density of 125 fish/km (PFBC and cooperative nurseries combined, see Table S1 for more information). Average size of stocked fish was approximately 254 to 305 mm. In Pennsylvania, average wild adult brook trout density is 5 fish/100 m<sup>2</sup> (Wagner et al. 2014) and average size of adult brook trout in Loyalsock Creek sampled during this study was approximately 120 mm. As such, brook trout density and biomass can increase by several orders of magnitude following a stocking event.

Only PFBC-operated state fish hatcheries are used in PFBC stocking efforts, and the majority of fish stocked in the Loyalsock Creek watershed originate from the PFBC's Tylersville State Fish Hatchery, with limited stocking from Benner Spring State Fish Hatchery. While captive fish are occasionally moved among hatcheries, wild fish have not been used to supplement hatchery

broodstocks in Pennsylvania for at least 50 years (Brian Wisner, PFBC, personal communication).

Sport fishing organizations also stock streams in the Loyalsock Creek watershed with fish from a

PFBC cooperative nursery. The PFBC provides fish to its cooperative nurseries, and so genetic

samples collected from PFBC hatcheries are representative of fish from cooperative nurseries.

Private landowners sometimes stock fish from privately operated, non-PFBC sponsored hatcheries.

Records are not maintained for private stocking events, but they represent a minimal source of

stocked fish in the watershed.

### *Sample Collection and Genotyping*

Using backpack electrofishing, we collected between 2 and 197 (average = 58) tissue samples from adult brook trout (*i.e.*, >100 mm; Whiteley et al. 2012) from each of 30 sites distributed throughout the Loyalsock Creek watershed from June 2015 to April 2017. In January 2017, tissue samples were obtained from 50 to 100 individuals from five brook trout strains at PFBC state fish hatcheries, including the Oswayo (OSW), Tylersville (TYL) and Bellefonte (BELL) strains and two representative groups from the Benner Springs strain (BNSPB, BNSPLP). For all fish, we excised a 5-mm<sup>2</sup> portion of the upper caudal fin and preserved the tissue in 95% non-denatured ethanol.

Genomic DNA was extracted for all samples following the manufacturers' protocols for the MagBind® Tissue DNA Kit KF (Omega Biotek, Norcross, GA). All DNA extractions were carried out using the Kingfisher® Flex (ThermoScientific, Waltham, MA) automated purification system. Samples were genotyped at 12 microsatellite loci developed for use in brook trout, including *SfoC*-113, *SfoD*-75, *SfoC*-88, *SfoD*-100, *SfoC*-129, *SfoC*-24, *SfoB*-52, *SfoC*-28, *SfoC*-38, *SfoC*-79, *SfoC*-86, and *SfoD*-91 (King et al. 2012).



Loci were combined into three multiplexes for polymerase chain reaction (PCR) amplification. Each 15  $\mu$ l-PCR reaction consisted of 1.5  $\mu$ l of genomic DNA extract along with 1.5X PCR buffer, 3.75 mM MgCl<sub>2</sub>, 0.3175 mM dNTPs, 0.08-0.18  $\mu$ M of each forward and reverse primer, 0.08 units/ $\mu$ L of GoTaq Flexi DNA polymerase, and deionized water. The amplification cycle for all multiplex mixes consisted of an initial denaturation at 94° C for 2 min, followed by 35 cycles of: 94° C denaturation for 45 sec, 56° C annealing for 45 sec, and 72° C extension for 2 min. A final extension was conducted at 72° C for 10 min.

An Applied Biosystems 3130xl genetic analyzer (ThermoFisher Scientific) was used for fragment analysis. Alleles were individually scored using GeneMapper® version 4.1 software (ThermoFisher Scientific) by two independent readers. Genotypes were obtained for >99% of loci across all individuals. For purposes of quality assurance, we re-extracted and genotyped 10% of all samples and compared the results to those of original samples. The genotyping error rate was less than 1% per locus.

Because some sites were sampled more than once, we used the program Cervus 3.0.7 (Kalinowski et al. 2007) to identify duplicate genotypes. The probability of two unique individuals sharing the same multilocus genotype was < 0.0001%, and so duplicate genotypes were considered repeated samples and removed from the analysis. In total, we sampled 1742 unique wild and 300 hatchery brook trout (Table 1). For sites with repeat sampling, we combined all samples into the same population. As we quantified assignment probabilities independently for each fish without respect to population origin, combining samples does not affect the outcome of our analysis.

### *Simulating Population Centroids*

Differences in sample size and allelic richness, which were both highly variable across the populations sampled (Table 1), have been shown to bias individual estimates of assignment probability using traditional algorithms found in the programs STRUCTURE and GeneClass (Halbisen & Wilson, 2009; Wang, 2017). In particular, there is a tendency for algorithms to bias assignment towards populations of larger sample size and higher genetic diversity. To minimize these effects, we quantified each individual's probability of wild origin,  $P(\text{wild})$ , following the methods of Karlsson et al. (2014). The method described by Karlsson et al. (2014) circumvents the aforementioned challenges by simulating centroids for the wild and hatchery populations that represent average allele frequencies across all populations, and uses these centroids, rather than putative sample populations, as the basis for individual fish assignment.

Provided that there is significant differentiation in wild and captive populations (an assumption we tested, as described below), restricting assignment to either hatchery or wild origin is a more powerful analysis that results in less ambiguous assignment probabilities. This is particularly advantageous when there is at least intermittent gene flow among populations or between groups (*i.e.*, hatchery and wild), as traditional approaches may fail to reach assignment consensus when probabilities are divided among several populations. Further, because centroids represent the average allele frequencies of populations within the same group, this dichotomous classification scheme minimizes spurious assignments related to inexhaustive sampling of potential source populations. This was particularly important for this analysis, as it was not possible to sample hatcheries used for private stocking or all streams in the Loyalsock Creek watershed.

Wild and hatchery population centroids were constructed from 22 randomly selected fish from each wild population and 50 randomly selected fish from each hatchery population. Sample size was based on the smallest reasonable population sample size within each group. Only two individuals were captured from mainstem Loyalsock Creek, and so fish from that site were not used to generate population centroids. Sampling was performed without replacement, and the randomly sampled wild and hatchery populations consisted of 638 and 250 unique fish, respectively. Using the program HybridLab (Nielsen et al. 2006), we simulated mating events within the randomly selected wild and hatchery populations to obtain 500 simulated wild and 500 simulated hatchery individuals. These simulated individuals comprised the wild and hatchery centroids.

To evaluate the precision of the centroid generation method, we constructed ten unique wild and hatchery centroids, each time using a different set of randomly selected individuals. Average within-group  $F_{ST}$  for both wild and hatchery centroids was  $<0.001$  (variance  $<0.00001$ ) and average  $F_{ST}$  between hatchery and wild centroids was 0.07. This demonstrates that assignment probabilities and classification are unlikely to be affected by the centroid generation process. Additionally, HybridLab conditions random mating on observed allele frequencies, thus ensuring that centroids are centered within the observed wild and hatchery populations and representative of average within-group allelic variance. However, we checked this assumption with a principal coordinate analysis (PCoA) of pairwise  $F_{ST}$  and tested for conformance to Hardy-Weinberg expectations in GenAEx 6.5.

Using non-wild genotypes to generate the wild centroid could mischaracterize the wild centroid and bias individual assignments. Ultimately, because the majority of individuals in our study assigned to wild origin (see *Results: Population Assignments of Sampled Fish*), the relative

abundance of wild trout relative to hatchery or introgressed supplants the influence of non-wild genotypes in the centroid. However, we evaluated this assumption by generating ten centroids using only individuals from populations where all fish assigned to wild origin. Average  $F_{ST}$  between these centroids and the ten centroids generated from the full dataset was 0.003, which is an order of magnitude lower than the average  $F_{ST}$  among wild populations of brook trout in this study (0.07) and most wild brook trout populations reported in the literature (*e.g.*, Kelson et al. 2015). Because using more individuals to generate the wild centroid helps minimize assignment bias by increasing the amount of explained genetic variation in our samples, we elected to use the centroid generated from the full dataset for individual assignments. However, individual assignments with the more restrictive centroids would be nearly identical to the centroids generated from the full dataset.

#### *Estimating Individual Wild Probability*

We estimated  $P(\text{wild})$  for each sampled wild and hatchery individual using the program STRUCTURE v 2.3.4 (Pritchard et al. 2000) executed through the ParallelStructure package in R (Besnier & Glover, 2013). To minimize assignment biases due to unequal sample sizes (Kalinowski 2011), a separate STRUCTURE analysis was completed for each observed individual. Each analysis included the 500 simulated wild fish and 500 simulated hatchery fish (*i.e.*, the population centroids), as well as the genotype for one sampled individual. For each STRUCTURE run, we applied 50,000 repetitions as burn-in and 100,000 repetitions after burn-in with no *a priori* information on sample location and assumed two populations ( $K = 2$ ), which corresponded to the dichotomy between wild and hatchery populations. The coefficient of ancestry ( $q$ ) to the wild cluster was interpreted as the  $P(\text{wild})$  for each observed individual.

### *Assignment of Individuals to Wild, Hatchery, or Introgressed Origin*

To identify introgressed individuals, we needed to determine upper and lower thresholds for  $P(\text{wild})$  commensurate with interbreeding between a wild and hatchery fish. To accomplish this, we used HybridLab to simulate 500 wild-hatchery matings using the 638 wild and 250 hatchery fish randomly selected above. We then completed 500 independent runs of STRUCTURE, this time including the 500 simulated wild fish, 500 simulated hatchery fish, and one simulated wild-hatchery cross.

We developed an expected distribution for  $P(\text{wild})$  for introgressed individuals using the 500  $q$ -values from the STRUCTURE analysis of the simulated wild-hatchery crosses. Values of  $P(\text{wild})$  that fell between the 2.5 and 97.5 percentile were considered characteristic of introgressed origin, and observed wild and hatchery individuals with a  $P(\text{wild})$  that fell within this range were classified as introgressed. Fish with a  $P(\text{wild})$  below the 2.5 percentile were classified as pure hatchery origin, and fish with a  $P(\text{wild})$  above the 97.5 percentile were classified as pure wild origin. Others have recommended different methods for identifying introgressed individuals, including the use of 5<sup>th</sup> and 95<sup>th</sup> percentiles (Karlsson et al. 2014) and raw STRUCTURE  $q$ -values between 0.10 and 0.90 (Harbicht et al. 2014; Vähä & Primmer, 2006). However, we elected to use a larger interval to characterize introgressed  $P(\text{wild})$  because it offers a more powerful estimate of introgression and is more likely to assign individuals as introgressed rather than originating from a hatchery or wild source.

Because the true ancestry of wild-caught individuals is unknown, the accuracy of our assignment method could not be evaluated using the empirical data. Therefore, we repeated the analysis on a simulated dataset with a known number of introgressed individuals. This analysis also enabled us to evaluate the accuracy of individual assignments when using different assignment thresholds for classifying introgression (*i.e.*, using a narrower range than the 2.5 and 97.5 percentiles). Our method detected introgressed individuals with zero error when using the 2.5 and

97.5 percentiles. However, error rate increased to 30% when using the 5 and 95 percentiles as assignment thresholds (see Supplementary Information for more details).

Using a larger range of values of  $P(\text{wild})$  to characterize introgression increases the likelihood of assigning individuals to admixed origin, which also likely increases the ability to detect post-F1 introgression. To evaluate whether our methods were capable of detecting post-F1 introgression, we conducted a parallel analysis using the program NewHybrids (Anderson and Thompson, 2002) to determine the probability that a sampled fish was from wild, hatchery, first-generation, or post-first-generation origin (representing the cumulative probability of F2 or backcross). Assignment probabilities were determined by running 100,000 sweeps of four chains after 25,000 burn-in sweeps with a Jeffrey's prior for  $\theta$  and  $\pi$ .

#### *Environmental Correlates to Introgression*

We used a hierarchical logistic regression to model the probability of a wild-caught fish assigning as introgressed as a function of several environmental covariates. Wild-caught fish that assigned to a putative hatchery origin were removed prior to analysis. Environmental covariates included site-level predictors that described local habitat and water quality (temperature, dissolved oxygen, conductivity, hardness, total alkalinity, pH, stream width, and adult brook trout density calculated with the Zippin [1958] removal method [J.M. Niles, unpublished data, Susquehanna University Freshwater Research Initiative]). We also included watershed-level predictors that were derived from Geographic Information Systems (watershed area [ $\text{km}^2$ ], distance to closest stocking reach, and land use). To account for environmental stochasticity, which could alter the interannual probability of introgression, we averaged site-level data from 2013 to 2015 (see Table S2 for the average and range of values for each covariate). To avoid issues associated with multicollinearity, highly correlated (absolute value of Pearson's  $r > 0.7$ ) variables were not included in the same model. We also excluded land use because the majority of sites occurred in a state forest, and difference in

land use among watersheds was not biologically meaningful (*e.g.*, all but one site had a watershed with >90% forest cover). All covariates were transformed into z-scores prior to analysis.

We fitted four candidate models to evaluate hypotheses about potential environmental correlates to introgression. The multi-scale model included all site- and watershed- level predictors, and represented the hypothesis that introgression is mediated by habitat at multiple scales. The second model included only site-level covariates, and represented the hypothesis that site-level factors are the dominant environmental predictors of introgression. In contrast, the third model included only watershed-level covariates and was used to evaluate the hypothesis that large-scale habitat features are more important than local properties in predicting introgression. The fourth model, the literature-supported model, was similar to the multi-scaled model, but only included covariates that have been previously identified in the literature as being important predictors of introgression in salmonids (Marie et al. 2012; Splendiani et al. 2013; Harbicht et al. 2014). All models included sample site as a random effect to account for the fact that there were multiple individuals collected from the same site, and we would expect genotypes from individuals collected within a site to be more similar than genotypes from individuals among sites (see Table S3 for model descriptions).

Models were executed using the lme4 package in R (Bates et al. 2014) and models were compared using an information theoretic approach to compare competing hypotheses (Burnham & Anderson 2002). Specifically, we calculated  $\Delta AIC_c$  and Akaike weights ( $w_i$ ) for each model.

## Results

### *Population Centroids*

Wild populations were genetically more diverse than hatchery populations with an average of 6.62 alleles per locus (SE=0.16) and expected heterozygosity of 0.68 (SE= 0.01) compared to 4.63 alleles per locus (SE=0.28) and an expected heterozygosity of 0.60 (SE=0.02) in hatchery populations (Table 1). There was significant genetic differentiation between observed wild and hatchery populations. Average  $F_{ST}$  within hatchery and wild populations was 0.06 (range: 0.01 to 0.08) and 0.07 (range: 0.01 to 0.16), respectively, and average  $F_{ST}$  among hatchery and wild populations was 0.13 (range 0.06 to 0.22). All pairwise  $F_{ST}$  values were significant at  $p < 0.001$ .

Samples from wild and hatchery populations were separated most significantly by axis 1 on the PCoA, which explained 28.6% of the total variance in the dataset. Axis 2 did not result in separation between the groups, and only explained 9.2% of the total variance. Simulated wild and hatchery population centroids were located in the middle of their respective populations on the PCoA (Figure 2) and were distinguishable in STRUCTURE (Figure 3) with no individuals incorrectly assigning to the competing cluster. Both centroid populations met the expectations of Hardy-Weinberg equilibrium ( $p > 0.015$  at all loci, with Bonferroni-corrected  $\alpha = 0.002$ ).

### *Population Assignments for Sampled Fish*

The expected distribution of  $P(\text{wild})$  for an introgressed fish had a lower 2.5 percentile of 0.06 and an upper 97.5 percentile of 0.94 (Figure 4). Therefore, when determining individual assignments, a fish that had a  $P(\text{wild})$  between 0.06 and 0.94 was classified as introgressed. Fish with a  $P(\text{wild})$  that fell below or above this range were classified as hatchery and wild origin, respectively.



Of the 1742 wild-caught fish sampled, 16 (<1%) assigned to pure hatchery origin and 97 (5.6%) to introgressed origin (Figure 5A). Average  $P(\text{wild})$  for all wild-caught fish was 0.97; however, there was considerable variation in average  $P(\text{wild})$  across populations (Figure 6). There were 23 populations with average  $P(\text{wild})$  greater than 0.97, and only two populations had an average  $P(\text{wild})$  less than 0.90, including SWAM (0.89) and DSLB (0.77). For both sites, low average  $P(\text{wild})$  was due to the presence of multiple fish assigning to hatchery origin. Of note, DSLB is the site with highest stocking densities and, while only two fish assigned as introgressed, we did detect eight fish of pure hatchery origin. Sites that had the largest proportion of introgressed individuals included CONK (28%) and FLAG (18%). However, there were 20 populations comprised of <5% introgressed individuals, with 10 populations having no fish assigning as introgressed.

Of the 300 hatchery fish sampled, 252 (84%) classified as pure hatchery origin and 48 (16%) as admixed origin (Figure 5B). No individual sampled from a hatchery assigned to pure wild origin. All hatchery strains stocked in the Loyalsock Creek watershed (BNSPLP, BNSPB, and TYL) had an average  $P(\text{wild}) < 0.005$ , and all fish from those strains assigned as pure hatchery origin.

All hatchery fish that assigned as introgressed were from the OSW (average  $P(\text{wild})$  0.04) and BELL (average  $P(\text{wild})$  0.11) strains. Wild fish have not been used in Pennsylvania hatcheries for at least 50 years, and so we know that these fish are not the product of wild-hatchery introgression. A more likely explanation is that hatchery fish that assigned as introgressed are the offspring from two different hatchery strains, at least one of which more genetically similar to wild populations than the hatchery strains used in our analysis. Unlike TYL and BNSP, hatcheries containing the OSW and BELL strains receive excess hatchery fish from across Pennsylvania at the end of each stocking season, and subsequently use those fish as broodstock (Brian Wisner, PFBC, personal communication). Our study only characterized fish from four of 13 state fish hatcheries (and each state fish hatchery can maintain multiple unique strains). Thus, although ambiguous genotypes could have resulted in the

incorrect assignments, a more likely scenario is that introgressed hatchery fish are the result of unexplained genetic variation from unsampled hatchery strains. Ultimately, fish from the OSW and BELL strains are not stocked at, or near, our sample sites by PFBC, so these assignment errors do not influence the accuracy of our results. However, the OSW and BELL strains were retained in our analysis to help account for uncertain sources of stocked fish used in undocumented, private stocking events.

There was consensus between NewHybrids and STRUCTURE for 1933 of 2042 individuals (95%). NewHybrids assigned 52 individuals as introgressed, all to post-F1 origin. STRUCTURE assigned 41 of these individuals (79%) as introgressed, and 11 to hatchery origin. The 11 fish that assigned as introgressed by NewHybrids but to hatchery origin by STRUCTURE were from one of the hatchery strains (thus representing an assignment error by NewHybrids).

Of the remaining 98 discrepancies, 36 individuals (37%) had inconclusive ancestry in NewHybrids (*i.e.*, <80% assignment probability to any one category, Sloss et al. 2008), and 61 individuals (62%) assigned as wild by NewHybrids but introgressed in STRUCTURE. Only one individual assigned to hatchery in NewHybrids and introgressed in STRUCTURE. This suggests that, while the two methods provided similar results, the STRUCTURE simulation method offers a more conservative and, in the case of hatchery individuals, more accurate, estimate of introgression. We elected to make inferences about individual population origin based on the STRUCTURE simulation method, as STRUCTURE is a more powerful algorithm when using fewer loci (Vähä & Primmer, 2006) and because identification of cohort structure was not necessary to accomplish the objectives of this analysis. However, it is worth noting that NewHybrids assigned all individuals to post-F1 origin, suggesting that our STRUCTURE method was likely capable of detecting post-F1 introgression.

## *Environmental Correlates to Introgression*

The two top-ranked models, the watershed-level model and the site-level model, were nearly indistinguishable based on  $\Delta AIC_c$  ( $\Delta AIC_c = 0.35$ ) with weights of 0.48 and 0.40, respectively. There was very little support for either the multi-scaled ( $w_i = 0.07$ ) or literature-supported ( $w_i = 0.05$ ) models. All models explained very little among-site variation in introgression, with  $R^2 < 0.03$  for all models (Table S3).

Given that the watershed-level and site-level models had similar support, we averaged these two models to generate a consensus model. The 95% confidence intervals for all parameters in the consensus model overlapped zero (see Table S4 for parameter estimates and standard errors).

### **Discussion**

Supplemental stocking with captively propagated individuals is frequently practiced to increase recreational opportunities (Naish et al. 2007) and mitigate declines of threatened and endangered populations (Fraser 2008). Frequent stocking is predicted to increase the incidence of captive introgression into wild populations (Fleming and Petersson, 2001); however, few studies have evaluated the prevalence of introgression in actively stocked populations. After sampling 30 wild brook trout populations, many of which at or near recent stocking locations, we found low incidences of introgression. Less than 6% of all wild-caught fish probabilistically assigned to admixed origin, and, on average, 94% of all fish within each sampled population assigned to pure wild origin. Our results suggest that the risk of domestic introgression in ecosystems with a history of repeated stocking may be minimal for some species and populations.

Introgression can alter native population genetic diversity, disrupt the frequency of locally adapted gene complexes, and decrease population fitness (Naish et al. 2007). However, our understanding of the probability of these events occurring is based largely on studies of unintentional releases, legacy effects of terminated stocking practices, and effects of large-scale fishery augmentation projects (Fleming and Petersson 2001; Ford et al. 2006; Naish et al. 2007). While these studies provide valuable insights, they may not be adequate surrogates for direct study of present-day sport fish stocking programs. Many documented instances of accidental releases occur in relatively large habitat patches, such as estuaries and bays, where rapid dispersal of escaped individuals away from the source can increase survival and reproduction (Piccolo and Orlikowska, 2012). Conversely, stream salmonids are not as vagile, and are primarily constrained to low-order watersheds (Kanno et al. 2011) where increased angling pressure may lead to higher harvest rates, particularly for domestic fish which are often more susceptible to angling (García-Marín et al. 1999). Headwater streams are also spatially and temporally complex systems that are prone to rapid changes in flow, temperature, and food availability. This stochasticity may exacerbate the phenotypic mismatch of hatchery fish with wild environments, thereby further decreasing fitness and survival of hatchery individuals in the wild.

Other studies of domestic introgression focus on populations that were last stocked over 20 years ago, when broodstocks had been held in captivity for fewer generations. Although significant genetic change can occur over short timespans (Christie et al. 2014), minimal genetic differences between historic hatchery and wild stocks are not expected to significantly influence patterns of genetic diversity in present-day wild populations (Ford, 2002; Naish et al. 2007, but see Stowell et al. 2015 for a counterexample in cutthroat trout (*Oncorhynchus clarkii*) populations).

A significant limitation with studies of historic and accidental releases is that power to detect introgression decreases with generation time (Vähä & Primmer, 2006). Assuming wild and hatchery individuals are genetically distinct (an assumption that is not always met when studying historic populations, as discussed above), first-generation hybrids can be readily detected using relatively simple genetic assignment tests. However, post-F1 offspring can be difficult to identify without multiple diagnostic loci or comparisons to reference samples collected before captive release (Mадiera et al. 2005; Vähä & Primmer, 2006). Similarly, backcrosses with wild or captive individuals and variability in allelic inheritance can confound the expected distribution of ancestry coefficients, which makes assignment to any specific cohort difficult without a large number (> 24) of informative loci (Vähä & Primmer, 2006). These factors can inflate variance estimates and limit the ability to detect introgression, even at the first-generation level (Karlsson et al. 2014).

We circumvent these challenges by studying populations that are currently stocked, which should have a large proportion of first-generation introgressed offspring relative to other cohort classes if interbreeding between stocked and wild fish is occurring. While we base assignment thresholds from only first-generation simulated crosses, the use of conservative thresholds for individual assignments provided high power to detect introgression beyond the F1 cohort (as determined via independent analysis in NewHybrids). Thus, our assignment methods present an improvement over other methods (Vähä & Primmer, 2006), and suggests that single crosses, rather than more extensive simulation of additional cohorts, are sufficient to identify contemporary introgression. This method is particularly advantageous for detecting introgression given that simulations beyond the first generation become intractable, and the power to assign individuals to specific cohorts declines as the distribution of assignment values begin to overlap among groups. Attempts to simulate beyond the first generation are further complicated for species like brook trout, where overlapping generations will increase the variance of expected distributions and decrease assignment power.

The power of the analysis draws from the generation of an expected distribution for introgressed  $P(\text{wild})$ , which is favored in a logistical sense because it does not require historic reference samples. Using this distribution as the basis for population assignment removes uncertainty associated with interpretation of  $q$ -values from STRUCTURE output (van Wyk et al. 2016). Further, it allows researcher to adjust assignment thresholds to balance the tradeoff between conservatism and bias when estimating introgression in their study system. For example, a larger range of  $q$ -values used to classify introgression may be more conservative, but could introduce bias by misclassifying wild or hatchery individuals as introgressed. The significance of this tradeoff will likely depend on the research and conservation objectives.

The analysis further gains power by restricting assignment to two population centroids, rather than putative sample populations, which decreases assignment errors by avoiding assumptions of assignment algorithms and minimizing bias associated with sampling design (Karlsson et al. 2014; Wang, 2017). In particular, this approach negates the need to exhaustively sample wild and hatchery populations, provided that samples are representative of total within-population variation (Karlsson et al. 2014). Application of this method to simulated data here (Supplementary Information) and by Karlsson et al. (2014) using populations of comparable genetic differentiation as our observed wild and hatchery populations suggests that this method is a reliable alternative for individual assignment tests.

As one of the most commercially and recreationally valuable fish taxa, captive releases of salmonid species have been occurring for nearly 150 years. As such, salmonids have been the focal species in the majority of studies of introgression in fish and wildlife populations (Araki & Schmid, 2010). Many of these studies have produced equivocal results, even when limiting comparisons within a single species. For example, farmed Atlantic salmon (*Salmo salar*) unintentionally released into large rivers have shown a high propensity to interbreed with native populations and disrupt wild

population genetics, while other populations invaded by captive escapees showed limited introgression (Hansen et al. 2009; Glover et al. 2013; Ozerov et al. 2016). Likewise, studies of fluvial brown trout (*Salmo trutta*) and lacustrine lake trout and brook trout suggest that rates of introgression vary across a landscape, even when populations are separated at small spatial scales (Madiera et al. 2005; Almodóvar et al. 2006; Halbisen & Wilson, 2009; Marie et al. 2012).

Our results corroborate those by Humston et al. (2012) and suggest variable, but overall lower degrees of introgression than other studies of stream salmonids (see Fleming & Petersson, 2001 for comparison). Of note, all sites comprised of >10% introgressed individuals are not directly stocked by PFBC or cooperative nurseries, including one site (CONK) which has no record of stocking for over 50 years and is isolated by a downstream impoundment. While the possibility of private stocking cannot be dismissed, this result suggests that there may be legacy effects of genetic introgression that continue to accumulate in populations long after the termination of management activities. Additionally, while natural selection may be able to purge deleterious functional genes within several generations, (Harbicht et al. 2014), introgression may remain detectable in noncoding regions of the genome (Tufto, 2017).

Why the degree of introgression varies significantly across species and among populations remains uncertain. Some have suggested a negative correlation between the prevalence of introgression and population size (Harbicht et al. 2014). This is presumably because competition for limited resources is higher in populations that approach carrying capacity, which decreases the probability of domestic fish surviving and reproducing in the wild. This hypothesis has support from studies that have found higher rates of introgression in populations inhabiting high-elevation and/or low pH environments, both of which are known to limit brook trout productivity (Marie et al. 2012; Splendiani et al. 2013; Harbicht et al. 2014).

Results of our models correlating introgression to environmental predictors explained little variation in introgression among populations. This was not unexpected given that introgressed individuals constituted <6% of our sample, and so there was minimal among-site variability to model. However, even though the estimated effects for both site- and watershed-level predictors had large uncertainty, the direction of the estimated effects generally agreed with published literature. For example, effect estimates were negative for pH and adult density and positive for temperature. This corroborates findings from others, who have found empirical evidence that introgression was more common in lakes with low pH, high temperature, and smaller adult densities (Marie et al. 2012; Harbicht et al. 2014; Létourneau et al. 2017). Likewise, the predicted direction for parameter estimates for stream width and watershed area were both negative, indicating the potential for decreased introgression in larger stream reaches. As smaller streams are characterized by stochasticity at finer temporal and spatial scales, this result could support findings from Splendiani et al. (2013) that suggest increased introgression at sites with more unstable stream flows. Given the overall uncertainty of our parameter estimates, future work would benefit from more direct study of introgression in stream ecosystems to determine the effect of habitat on stocked fish survival and reproduction.

Variable fitness of stocked individuals may also explain why introgression differs across studies. High mortality, low-quality gametes, and low offspring survival are typical of captive individuals released into wild populations (Araki & Schmid, 2010), particularly as generation time in captivity increases. This can be due to reductions in genetic diversity and selection for genotypes and phenotypes adapted to the hatchery environment (Naish et al. 2007), which may not be optimal for persistence in the wild. While we did not explicitly examine individual survival and reproduction, few hatchery brook trout are collected during summer surveys by PFBC staff (Jason Detar, PFBC, personal communication), suggesting limited long-term survival of brook trout stocked in the Loyalsock Creek watershed.



While there was minimal introgression at any single site, 21 sites had at least one wild-caught individual assign to either introgressed or hatchery origin. In some cases, hatchery influence was detected in populations located several kilometers from the nearest stocking location and with no history of previous stocking. Domestic trout can readily disperse after release (Bettinger & Bettoli, 2002), particularly in spring when trout can use mid-reach rivers (which often exceed trout thermal tolerance in summer, Aunins et al. 2015) as movement corridors. While mortality is often higher for more mobile, stocked fish (Aarestrup et al. 2005), our results suggest that stocked trout may move into, and survive in, wild populations and/or genetic invasion may occur through a stepping-stone mechanism (Hitt et al. 2003). This result highlights the need to potentially consider the influence of stocking at larger spatial scales that extend beyond the area of direct stocking effort.

Although we found strong evidence of minimal introgression, results from this study should be viewed circumspectly when assessing the potential risks of introgression from stocking. We did not evaluate possible declines in wild population fitness as a result of direct competition and/or failed mating attempts with stocked fish (Weber & Fausch, 2001; McGinnity et al. 2003). Likewise, we only evaluated neutral microsatellite markers and did not test for genetic introgression at loci linked to fitness (Harbicht et al. 2014) or at other areas of the genome where introgression may occur more readily (Ozerov et al. 2016). Evidence from other studies of stream salmonids suggest that rates of introgression can be high, (Fleming & Petersson, 2001; Muhlfeld et al. 2009), and that it can lead to nonrandom changes to the genome that correlate to reductions in survival and fitness. As brook trout populations are often characterized by low levels of genetic diversity and small effective population sizes relative to other salmonids (Kelson et al. 2015; Ruzzante et al. 2016), they can be particularly susceptible to genetic invasion by captive populations as rates of introgression increases. Until a better understanding of the potential non-genetic effects of stocking are known and there is more investigation into the factors that influence rates of introgression, a cautionary approach to recreational stocking programs may still be warranted.

## Acknowledgements

We thank Eric Hallerman, Julian Avery, and Jason Detar, for providing comments on an earlier draft, Steve Szoke, Nathan Newton and Laurel Seemiller, for assistance in the field, Shannon Julian for laboratory support, and the PFBC for providing samples from their brook trout hatchery strains. Unpublished environmental data was provided courtesy of Jonathan M. Niles of the Freshwater Research Initiative at Susquehanna University. This research was funded by the R.K. Mellon Freshwater Research Initiative at Susquehanna University and is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1255832. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the USFWS.

## Data Archiving Statement

Data for this study are available from the Dryad Digital Repository:

<https://doi.org/10.5061/dryad.mb37t1q>.

## References

- Aarestrup, K., Jepsen, N., Koed, A., & Pedersen, S. (2005). Movement and mortality of stocked brown trout in a stream. *Journal of Fish Biology*, 66, 721-728.
- Allendorf, F. W., Leary, R. F., Hitt, N. P., Knudsen, K. L., Lundquist, L. L., & Spruell, P. (2004). Intercrosses and the U.S. Endangered Species Act: Should hybridized populations be included as westslope cutthroat trout? *Conservation Biology*, 18, 1203-1213.

Almodóvar, A., Nicola, G. G., Elvira, B., & García-Marín, J. L. (2006). Introgression variability among Iberian brown trout Evolutionary Significant Units: the influence of local management and environmental features. *Freshwater Biology*, 51, 1175-1187.

Anderson, E. C., & Thompson, E. A. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, 160, 1217-1229.

Araguas, R. M., Sanz, N., & García- Marín, J. L. (2004). Breakdown of the brown trout evolutionary history due to hybridization between native and cultivated fish. *Journal of Fish Biology*, 65, 28-37.

Araki, H., & Schmid, C. (2010). Is hatchery stocking a help or harm?: Evidence, limitations and future directions in ecological and genetic surveys. *Aquaculture*, 308, S2-S11.

Araki, H., Berejikian, B.A., Ford, M.K., & Blouin, M.S. (2008). Fitness of hatchery-reared salmonids in the wild. *Evolutionary Applications*, 1, 342-355.

Araki, H., Cooper, B., & Blouin, M. S. (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, 318, 100-103.

Askey, P. J., Parkinson, E. A., & Post, J. R. (2013) Linking fish and angler dynamics to assess stocking strategies for hatchery-dependent, open-access recreational fisheries. *North American Journal of Fisheries Management*, 33, 557-568.

Aunins, A. W., Petty, J. T., King, T. L., Schilz, M., & Mazik, P. M. (2015). River mainstem thermal regimes influence population structuring within an Appalachian brook trout populations. *Conservation Genetics*, 16, 15-29.

Baer, J., Blasel, K., & Diekmann, M. (2007). Benefits of repeated stocking with adult, hatchery-reared brown trout, *Salmo trutta*, to recreational fisheries? *Fisheries Management and Ecology*, 14, 51-59.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, <http://CRAN.R-project.org/package=lme4>

Besnier, F., & Glover, K. A. (2013). ParallelStructure: A R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PLoS ONE*, 8, e70651.

Bettinger, J. M., & Bettoli, P. W. (2002). Fate, dispersal, and persistence of recently stocked and resident rainbow trout in a Tennessee tailwater. *North American Journal of Fisheries Management*, 22, 425-432.

Bolstad, G. H., Hindar, K., Robersten, G., Jonsson, B., Sægrov, H., Diserud, O. H., Fiske, P., Jensen, A. J., Urdal, K., Næsje, T. F., Barlaup, B. T., Florø-Larsen, B., Lo, H., Niemelä, E., & Karlsson, S. (2017). Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nature Ecology and Evolution*, 1, 0124.

Bowman, J., Beauclerc, K., Farid, A. H., Fenton, H., Klütsch, C. F. C., & Schulte-Hostedde, A. I. (2017). Hybridization of domestic mink with wild American mink (*Neovison vison*) in eastern Canada. *Canadian Journal of Zoology*, 95, 443-451.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York, NY: Springer.

Christie, M. R., Ford, M. J., & Blouin, M. S. (2014). On the reproductive success of early-generation hatchery fish in the wild. *Evolutionary Applications*, 7, 883-896.

Currens, K. P., Hemmingsen, A. R., French, R. A., Buchanan, D. V., Schreck, C. B., & Li, H. W. (1997). Introgression and susceptibility to disease in a wild population of rainbow trout. *North American Journal of Fisheries Management*, 17, 1065-1078.

Davis, L. A., Wagner, T., & Bartron, M. L. (2015) Spatial and temporal movement dynamics of brook *Salvelinus fontinalis* and brown trout *Salmo trutta*. *Environmental Biology of Fishes*, 98, 2049-2065.

Evans, D. O., & Willox, C. C. (1991). Loss of exploited indigenous populations of lake trout, *Salvelinus namaycush*, by stocking of non-native stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 48, 134-147.

Fleming, I. A., & Petersson, E. (2001). The ability of released, hatchery salmonids to breed and contribute to the natural productivity of wild populations. *Nordic Journal of Freshwater Research*, 75, 71-98.

Ford, M. J. (2002). Selection in captivity during supportive breeding may reduce fitness in the wild. *Conservation Biology*, 16, 815-825.

Ford, M. J., Fuss, H., Boelts, B., LaHood, E., Hard, J., & Miller, J. (2006). Changes in run timing and natural smolt production in a naturally spawning coho salmon (*Oncorhynchus kisutch*) population after 60 years of intensive hatchery supplementation. *Canadian Journal of Fisheries and Aquatic Sciences*, 63, 2343-2355.

Fraser, D. J. (2008). How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications*, 4, 535-586.

García-Marín, J. L., Sanz, N., & Pla, C. (1999). Erosion of the native negative genetic resources of brown trout in Spain. *Ecology of Freshwater Fish*, 8, 151-158.

Glover, K. A., Pertoldi, C., Besnier, F., Wennevik, V., Kent, M., & Skaala, Ø. (2013). Atlantic salmon populations invaded by farmed escapees: quantifying genetic introgression with a Bayesian approach and SNPs. *BMC Genetics*, 14, 74.

Halbisen, M. A., & Wilson, C. C. (2009). Variable introgression from supplemental stocking in southern Ontario populations of lake trout. *Transactions of the American Fisheries Society*, 138, 699-719.

Hallerman, E. M. 2003. Coadaptation and outbreeding depression In *Population genetics: Principles and applications for fisheries scientists* (pp. 239-259). Bethesda, MD: American Fisheries Society.

Hansen, M. M., Fraser, D. J., Meier, K., & Mensberg, K. D. (2009). Sixty years of anthropogenic pressure: a spatio-temporal genetic analysis of brown trout populations subject to stocking and population declines. *Molecular Ecology*, 18, 2549-2562.

Harbicht, A. B., Alshamli, A., Wilson, C. C., & Fraser, D. J. (2014). Anthropogenic and habitat correlates of hybridization between hatchery and wild brook trout. *Canadian Journal of Fisheries and Aquatic Sciences*, 71, 688-697.

Harbicht, A., Wilson, C. C., & Fraser, D. J. (2014). Does human-induced hybridization have long-term genetic effects? Empirical testing with domesticated, wild and hybridized fish populations. *Evolutionary Applications*, 7, 1180-1191.

Hindar, K., Ryman, N., & Utter, F. 1991. Genetic effects of cultured fish on natural fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 48, 945-957.

Hitt, N. P., Frissell, C. A., Muhlfeld, C. C., & Allendorf, F. W. (2003). Spread of hybridization between native westlope cutthroat trout, *Oncorhynchus clarki lewisi*, and nonnative rainbow trout, *Oncorhynchus mykiss*. *Canadian Journal of Fisheries and Aquatic Sciences*, 60: 1440-1451.

Hudy, M., Thieling, T. M., Gillespie, M., & Smith, E. P. (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.

Humston, R., Bezold, K. A., Adkins, N. D., Elsey, R. J., Huss, J., Meekins, B. A., Cabe, P. R., & King, T. L. (2012). Consequences of stocking headwater impoundments on native populations of brook trout in tributaries. *North American Journal of Fisheries Management*, 32, 100-108.

Huntingford, F. A. (2004). Implications of domestication and rearing conditions for the behavior of cultivated fishes. *Journal of Fish Biology*, 65, 122-142.

- Kalinowski, S. T., Taper, L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16, 1099-1106.
- Kanno, Y., Vokoun, J. C., & Letcher, B. H. (2011). Fine-scale population structure and riverscape genetics of brook trout (*Salvelinus fontinalis*) distributed continuously along headwater channel networks. *Molecular Ecology*, 20, 3711-3729.
- Karlsson, S., Diserud, O. H., Moen, T., & Hindar, K. (2014). A standardized method for quantifying unidirectional genetic introgression. *Ecology and Evolution*, 16, 3256-3263.
- Kazyak, D. C., Rash, J., Lubinski, B. A., & King, T. L. (2018). Assessing the impact of stocking northern-origin hatchery brook trout on the genetics of wild populations in North Carolina. *Conservation Genetics*, 19, 207-219.
- Kelson, S. J., Kapuscinski, A. R., Timmins, D., & Arden, W. R. (2015). Fine-scale genetic structure of brook trout in a dendritic stream network. *Conservation Genetics*, 16, 31-42.
- King, T. L., Lubinski, B. A., Burnham-Curtis, M. K., Stott, W., & Morgan, R. P. (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. *Conservation Genetics Resources*, 4, 539-543.
- Lamazé, F. C., Garant, D., & Bernatchez, L. (2013). Stocking impacts the expression of candidate genes and physiological condition in introgressed brook charr (*Salvelinus fontinalis*) populations. *Evolutionary Applications*, 6, 393-407.
- Létourneau, J., Ferchaud, A-L., Le Luyer, J., Laporte, M., Garant, D., & Bernatchez, L. (2017). Predicting the genetic impact of stocking in Brook Charr (*Salvelinus fontinalis*) by combining RAD sequencing and modeling of explanatory variables. *Evolutionary Applications*, 1-16.

Lynch, M., & O'Hely, M. (2001) Captive breeding and the genetic fitness of natural populations.

*Conservation Genetics*, 2, 363-378.

Madiera, M. J., Gómez-Moliner, B. J., & Barbé, A. M. (2005). Genetic introgression on freshwater fish populations caused by restocking programmes. *Biological Invasions*, 7, 117-125.

Marie, A. D., Bernatchez, L., & Garant, D. (2010). Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, 19, 2025-2037.

Marie, A. D., Bernatchez, L., & Garant, D. (2012). Environmental factors correlate with hybridization in stocked brook charr (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences*, 69, 884-893.

McGinnity, P., Prodöhl, P., Ferguson, A., Hynes, R., ó Maoiléidigh, N., Baker, N., Cotter, D., O'Hea, B., Cooke, D., Rogan, G., Taggar, J., & Cross, T. (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society B*, 240, 2443-2450.

Muhlfeld, C. C., Kalinowski, S. T., McMahon, T. E., Taper, M. L., Painter, S., Leary, R. F., & Allendorf, F. W. (2009). Hybridization rapidly reduces fitness of native trout in the wild. *Biology Letters*, 5, 328-331.

Naish, K. A., Seamons, T. R., Dauer, M. B., Hauser, L., & Quinn, T. P. (2013). Relationship between effective population size, inbreeding and adult fitness-related traits in a steelhead (*Oncorhynchus mykiss*) population released in the wild. *Molecular Ecology*, 22, 1295-1309.

Naish, K. A., Taylor, J. E., Levin, P. S., Quinn, T. P., Winton, J. R., Huppert, D., & Hilborn, R. (2007). An evaluation of the effects of conservation and fishery enhancement hatcheries on wild populations of salmon. *Advances in Marine Biology*, 53, 61-194.

Nielsen, E. E., Bach, L. A., & Kotlicki, P. (2006). HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Resources*, 6, 971-973.



Ozerov, M. Y., Gross, R., Bruneaux, M., Vähä, J. P., Burimski, O., Pukk, L., & Vasemägi, A. (2016).

Genomewide introgressive hybridization patterns in wild Atlantic salmon influenced by inadvertent gene flow from hatchery releases. *Molecular Ecology*, 25, 1275-1293.

Peakall, R., & Smouse, P. E. (2006). GenAEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288-295.

Peakall, R., & Smouse, P. E. (2012). GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research - An update. *Bioinformatics*, 28, 2537-2539.

Piccolo, J. J., & Orlikowska, E. H. (2012). A biological risk assessment for an Atlantic salmon (*Salmo salar*) invasion in Alaskan waters. *Aquatic Invasions*, 7, 259-270.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Reisenbichler, R. R., & McIntyre, J. D. (1977). Genetic differences in growth and survival of juvenile hatchery and wild steelhead trout, *Salmo gairdneri*. *Journal of Fisheries Research Board of Canada*, 34, 123-128.

Ruzzante, D. E., McCracken, G. R., Parmelee, S., Hill, K., Corrigan, A., MacMillan, J., & Walde, S. J.

(2016). Effective number of breeders, effective population size and their relationship with census size in an iteroparous species, *Salvelinus fontinalis*. *Proceedings of the Royal Society B*, 283, 20152601.

Ryman, N., & Laikre, L. 1991. Effects of supportive breeding on the genetically effective population size. *Conservation Biology*, 5, 325-329.

Saikkonen, A., Kekäläinen, J., & Piironen, J. (2011). Rapid growth of Atlantic salmon juveniles in captivity may indicate poor performance in nature. *Biological Conservation*, 144, 2320-2327.

- Sloss, B. L., Jennings, M. J., Franckowiak, R., & Pratt, D. M. 2008. Genetic identity of brook trout in Lake Superior south shore streams: Potential for genetic monitoring of stocking and rehabilitation efforts. *Transactions of the American Fisheries Society*, 137, 1244-1251.
- Splendiani, A., Ruggeri, P., Giovannotti, M., & Caputo, V. (2013). Role of environmental factors in the spread of domestic trout in Mediterranean streams. *Freshwater Biology*, 58, 2089-2101.
- Stowell, S. M. L., Kennedy, C. M., Beals, S. C., Metcalf, J. F., & Martin, A. P. (2015). The genetic legacy of more than a century of stocking trout: a case study in Rocky Mountain National Park, Colorado, USA. *Canadian Journal of Fisheries and Aquatic Sciences*, 72, 1565-1574.
- Stringwell, R., Lock, A., Stutchbury, C. J., Baggett, E., Taylor, J., Gough, P. J., & Garcia de Leaniz, C. (2014). Maladaptation and phenotypic mismatch in hatchery-reared Atlantic salmon *Salmo salar* released in the wild. *Journal of Fish Biology*, 85, 1927-1945.
- Tufto, J. (2017). Domestication and fitness in the wild: A multivariate view. *Evolution*, 71, 2262-2270.
- Utter, F. A. 2003. Genetic impacts of fish introductions In *Population genetics: Principles and applications for fisheries scientists* (pp. 357-378). Bethesda, MD: American Fisheries Society.
- Vähä, J., & Primmer, C. R. (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, 15, 63-72.
- Valiquette, E., Perrier, C., Thibault, I., & Bernatchez, L. (2014). Loss of genetic integrity in wild lake trout populations following stocking: insights from an exhaustive study of 72 lakes from Québec, Canada. *Evolutionary Applications*, 7, 625-644.
- Van Wyk, A. M., Dalton, D., Hoban, S., Bruford, M. W., Russo, I. M., Birss, C., Grobler, P., van Vuuren, B. J., & Kotzé, A. (2016). Quantitative evolution of hybridization and the impact on biodiversity conservation. *Ecology and Evolution*, 7, 320-330.

Accepted Article

Wagner, T., Deweber, J. T., Detar, J., Kristine, D., & Sweka, J. A. (2014). Spatial and temporal dynamics in brook trout density: implications for population monitoring. *North American Journal of Fisheries Management*, 34, 258-269.

Wang, J. (2017). The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources*, 17, 981-990.

Waples, R. A. (1991). Genetic interactions between hatchery and wild salmonids: Lessons from the Pacific Northwest. *Canadian Journal of Fisheries and Aquatic Sciences*, 48, 124-133.

Weaver, D. M., & Kwak, T. J. (2013). Assessing effects of stocked trout on nongame fish assemblages in southern Appalachian mountain streams. *Transactions of the American Fisheries Society*, 142, 1495-1507.

Weber, E. D., & Fausch, K. D. (2003). Interactions between hatchery and wild salmonids in streams: differences in biology and evidence for competition. *Canadian Journal of Fisheries and Aquatic Sciences*, 60: 1018-1036.

Whiteley, A. R., Coombs, J. A., Hudy, M., Robinson, Z., Nislow, K. H., & Letcher, M. H. (2012). Sampling strategies for estimating brook trout effective population size. *Conservation Genetics*, 13, 625-637.

Wollebæk, J., Heggenes, J., & Røed, K. H. (2010). Disentangling stocking introgression and natural migration in brown trout: survival success and recruitment failure in populations with semi-supportive breeding. *Freshwater Biology*, 55, 2626-2638.

Zipkin, C. (1958). The removal method of population estimation. *Journal of Wildlife Management*, 22, 82-90.

## Figure Captions

**Figure 1.** Distribution of 30 sample sites in the Loyalsock Creek watershed (circles) in northcentral Pennsylvania, United States. The mid-point of stream sections stocked by the Pennsylvania Fish and Boat Commission (PFBC) and PFBC-Cooperative Nurseries in 2015 are indicated by triangles. See Table 1 for full site names and sample sizes and Table S1 for detailed stocking histories.

**Figure 2.** PCoA plot for pairwise  $F_{ST}$  estimates between sampled wild (open circles) and hatchery brook trout populations (open squares). Simulated wild and hatchery population centroids are shown in closed circles and squares, respectively. The first PCoA axis explained 28.6% of total sample variance.

**Figure 3.** Bar plot representing the classification of simulated hatchery and wild individuals to putative centroid populations in STRUCTURE. Colors within bars represent the probability of each simulated individual assigning to either a wild cluster (white) or captive cluster (gray). No simulated individuals were incorrectly assigned to the competing cluster and few exhibited significantly admixed genotypes.

**Figure 4.** Distribution for probability of wild descent for 500 simulated introgressed individual brook trout. Estimates were generated by analyzing the multilocus genotype of each individual fish in STRUCTURE along with simulated wild and hatchery centroid populations. Horizontal dashed lines represent the lower 2.5 and upper 97.5 percentiles. Sampled individuals that had a wild probability below the 2.5 percentile were assigned to hatchery origin and individuals above the 97.5 percentile were assigned to wild origin. Sampled individuals with  $P(\text{wild})$  between these cutoff values were assigned to introgressed origin.

**Figure 5.** Distribution of wild probabilities  $P(\text{wild})$  for 1742 wild-caught (A) and 300 hatchery (B) brook trout, sorted by ascending  $P(\text{wild})$ . Estimates were generated by analyzing the multilocus genotype of each individual fish in STRUCTURE along with simulated wild and hatchery population centroids. Horizontal dashed lines represent the lower 2.5 and upper 97.5 percentiles from the null distribution for wild probability for an introgressed individual (see Figure 3).

**Figure 6.** Wild probabilities ( $P(\text{wild})$ ), shown in grey for 1742 wild-caught brook trout. Fish are sorted by population, and in ascending order of  $P(\text{wild})$  within a population. Populations appear in the same order as Table 1. Horizontal dashed lines represent the lower 2.5 and upper 97.5 percentiles from the null distribution for wild probability for an introgressed individual (see Figure 3).

| Site Name                             | Abbr.  | Sample Size | N <sub>A</sub> | H <sub>E</sub> | Average P(wild) | Number of Wild Fish | Number of Introgressed Fish | Number of Hatchery Fish | Average Stocking Density (fish/km) |
|---------------------------------------|--------|-------------|----------------|----------------|-----------------|---------------------|-----------------------------|-------------------------|------------------------------------|
| <b>HATCHERY</b>                       |        |             |                |                |                 |                     |                             |                         |                                    |
| Bellefonte                            | BELL   | 100         | 7.00           | 0.69           | 0.11            | 0                   | 38                          | 62                      |                                    |
| Oswayo Hatchery, Oswayo Strain        | OSW    | 50          | 6.17           | 0.69           | 0.04            | 0                   | 10                          | 40                      |                                    |
| Benner Springs, Lighthouse Strain     | BNSPLP | 50          | 3.33           | 0.56           | 0.00            | 0                   | 0                           | 50                      |                                    |
| Oswayo Hatchery, Tylersville Strain   | TYL    | 50          | 3.33           | 0.54           | 0.00            | 0                   | 0                           | 50                      |                                    |
| Benner Springs, B Strain              | BNSPB  | 50          | 3.33           | 0.55           | 0.00            | 0                   | 0                           | 50                      |                                    |
| <b>WILD</b>                           |        |             |                |                |                 |                     |                             |                         |                                    |
| Mill Run- Laporte                     | MILA   | 50          | 3.92           | 0.54           | 1.00            | 50                  | 0                           | 0                       |                                    |
| Level Run                             | LEVL   | 50          | 8.17           | 0.72           | 1.00            | 50                  | 0                           | 0                       |                                    |
| Mainstem Loyalsock Creek <sup>1</sup> | MAIN16 | 2           | 2.83           | 0.79           | 1.00            | 2                   | 0                           | 0                       | 39.76                              |
| Upstream East Branch                  | EAST   | 50          | 4.67           | 0.57           | 0.99            | 50                  | 0                           | 0                       |                                    |
| Unnamed Tributary to Elk Creek        | UNT    | 48          | 7.08           | 0.70           | 0.99            | 48                  | 0                           | 0                       |                                    |
| Bear Run <sup>2</sup>                 | BEAR   | 50          | 6.33           | 0.69           | 0.99            | 50                  | 0                           | 0                       | 39.76                              |
| Streby Run                            | STRB   | 50          | 7.92           | 0.74           | 0.99            | 49                  | 1                           | 0                       |                                    |
| Huckle Run <sup>2</sup>               | HUCK   | 22          | 3.92           | 0.58           | 0.99            | 22                  | 0                           | 0                       | 31.48                              |
| Yellow Run <sup>2</sup>               | YELL   | 50          | 5.67           | 0.68           | 0.99            | 50                  | 0                           | 0                       | 126.56                             |
| Sand Spring Run <sup>2</sup>          | SSR    | 50          | 7.33           | 0.73           | 0.99            | 50                  | 0                           | 0                       | 528.97                             |
| Grandad Run                           | GRAN   | 50          | 7.50           | 0.76           | 0.99            | 49                  | 1                           | 0                       |                                    |
| Red Run <sup>2</sup>                  | RED    | 50          | 7.25           | 0.72           | 0.99            | 49                  | 1                           | 0                       | 528.97                             |
| Pole Bridge Run <sup>1,2</sup>        | POLE   | 109         | 7.42           | 0.69           | 0.99            | 106                 | 3                           | 0                       | 77.67                              |
| Scar Run <sup>2</sup>                 | SCAR   | 50          | 5.67           | 0.68           | 0.99            | 48                  | 2                           | 0                       | 31.48                              |
| Downstream East Branch <sup>2</sup>   | DSEA   | 73          | 7.92           | 0.69           | 0.99            | 70                  | 3                           | 0                       | 107.56                             |
| Shanerburg Run <sup>1,2</sup>         | SHAN   | 197         | 9.92           | 0.73           | 0.99            | 192                 | 5                           | 0                       | 81.16                              |
| Coal Run <sup>2</sup>                 | COAL   | 50          | 6.25           | 0.68           | 0.98            | 48                  | 2                           | 0                       | 39.76                              |
| Brunnerdale Run                       | BRUN   | 50          | 7.00           | 0.75           | 0.98            | 47                  | 3                           | 0                       |                                    |

|                                       |       |     |      |      |      |     |    |   |        |
|---------------------------------------|-------|-----|------|------|------|-----|----|---|--------|
| Rock Run <sup>2</sup>                 | ROCK  | 50  | 5.58 | 0.68 | 0.98 | 46  | 4  | 0 | 126.56 |
| Snake Run <sup>2</sup>                | SNAK  | 50  | 5.17 | 0.60 | 0.98 | 45  | 5  | 0 | 76.34  |
| Weed Run <sup>2</sup>                 | WEED  | 50  | 7.08 | 0.64 | 0.98 | 47  | 3  | 0 | 251.69 |
| Dry Run- Hoagland Branch <sup>2</sup> | DRHO  | 50  | 6.33 | 0.66 | 0.98 | 47  | 3  | 0 | 251.69 |
| Jacoby Hallow <sup>2</sup>            | JACO  | 50  | 7.42 | 0.72 | 0.97 | 49  | 0  | 1 | 36.42  |
| Lick Run <sup>2</sup>                 | LICK  | 50  | 5.83 | 0.72 | 0.96 | 42  | 8  | 0 | 32.03  |
| Double Run <sup>1,2</sup>             | DOUB  | 154 | 8.83 | 0.74 | 0.95 | 129 | 25 | 0 | 74.75  |
| Flag Marsh Run                        | FLAG  | 49  | 7.67 | 0.72 | 0.95 | 40  | 9  | 0 |        |
| Mill Creek- Hillsgrove <sup>1</sup>   | MIHI  | 50  | 7.67 | 0.71 | 0.95 | 46  | 2  | 2 | 65.59  |
| Conklin Run                           | CONK  | 50  | 4.33 | 0.63 | 0.94 | 36  | 14 | 0 |        |
| Swamp Run <sup>2</sup>                | SWAM  | 50  | 7.83 | 0.72 | 0.89 | 44  | 1  | 5 | 251.69 |
| Little Bear Creek <sup>1,2</sup>      | DSL B | 38  | 8.00 | 0.76 | 0.77 | 28  | 2  | 8 | 353.49 |

<sup>1</sup> Stocking occurs at the sample location

<sup>2</sup> Stocking within 2km of the sample location

Site name abbreviations, sample sizes, measures of population genetic diversity ( $N_A$ - average number of alleles per locus,  $H_E$ -Nei's unbiased estimate of heterozygosity), and average stocking densities for 30 wild populations and five hatchery strains of brook trout used to determine the degree of introgression in the Loyalsock Creek watershed. Records are sorted by descending values of  $P(\text{wild})$ . Average  $P(\text{wild})$  and number of individuals that assigned to each class were determined by individual assignment to one of two (hatchery and wild) simulated populations. See supplementary text for more detailed stocking records.









