

Title. Identifying the Appropriate Unit of Management for GRSM Brook Trout (*Salvelinus fontinalis*)

Abstract. Great Smoky Mountains National Park (GRSM) is committed to monitoring ecological and evolutionary functions and processes of park ecosystems. Brook trout (*Salvelinus fontinalis*) is the only salmonid native to the Southern Appalachians and functions as a keystone species in some headwater streams. The historic use of hatchery-reared brook trout for supplemental and restorative stocking in GRSM underscores the need to recognize the evolutionary relationship among stream populations. A recent survey of microsatellite DNA variation in GRSM brook trout indicated the presence of highly significant differentiation at all hierarchical levels which suggests that the individual stream should be considered the unit of management. Given that management resources are limited and that stream-specific management is often not practical, fisheries managers need to know whether the genetic divergence observed among GRSM brook trout reflect adaptive differences or is the variation due to stochastic processes like random genetic drift. DNA microarrays are a powerful method for the global analysis of steady-state intracellular mRNA levels, and thus identifying genes that are transcriptionally modulated as a consequence of metabolic or bioenergetic demands. The information gathered from these arrays of gene sequences can be used to characterize complex biological processes and interactions providing insight into the adaptive significance of observed genetic differentiation. This research objective, if funded, would represent the first attempt at determining whether GRSM fisheries managers should focus their resources on genetic relatedness or demographics.

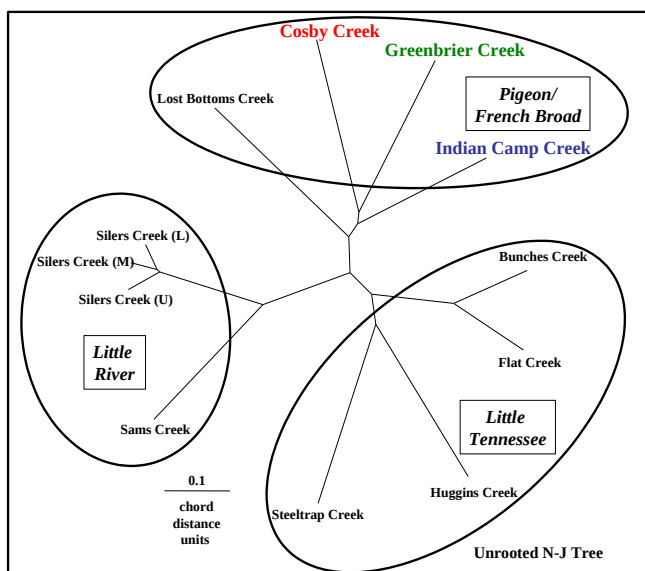
USGS Principal Investigator. Tim L. King, Research Fishery Biologist (Genetics), USGS-Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430; Phone: 304.724.4450; Facsimile: 304.724.4424; E-mail: tlking@usgs.gov.

Problem statement/justification.

Need: Great Smoky Mountains National Park (GRSM) is committed to monitoring ecological and evolutionary functions and processes of park ecosystems. Effective management of imperiled species is often complicated because inadequate information exists for identification of species, subspecies, distinct population segments, or management units.

Salmonid fishes inhabiting lower latitudes present this and other significant challenges to conservation biologists that attempt to maintain ecological and evolutionary processes within and among populations. Brook trout (*Salvelinus*

fontinalis) is the only salmonid native to the Southern Appalachians and functions as a keystone species in some headwater streams. Since the turn of the last century, this native fish has lost approximately 75 percent of its range in GRSM. The historic use of hatchery-reared brook trout for supplemental and restorative stocking in GRSM underscores the need to recognize the evolutionary relationship among stream populations. Given that management resources are limited and that stream-specific management is often not practical, fisheries managers need to know whether the genetic divergence observed among GRSM brook trout reflect adaptive differences or is the variation due to stochastic processes like random genetic drift.



priorities. Genetic variation detected at 13 microsatellite DNA loci in over 700 GRSM brook trout has shown a range of allelic diversity (including some of the lowest levels reported for vertebrate populations). Analyses have shown a high degree of genetic differentiation among GRSM's three drainages and among tributaries within those drainages (Figure 1). Similarly, analysis of molecular variance and likelihood assignment tests indicated the presence of highly significant differentiation at all hierarchical levels and implies that the individual stream should be considered the unit of management.

Genetic relationships among some GRSM brook trout populations have not been extensively investigated using contemporary molecular methods and thus constitute a critical gap in the park's developing knowledge base. The loss of

a single population within any stream is likely to result in a direct reduction in the park's existing biodiversity. These findings create somewhat of a management conundrum for GRSM - Does the genetic divergence observed among GRSM brook trout show adaptive significance or is the variation due to stochastic processes like random genetic drift? Indirect evidence from a study underway in GRSM suggests that the differentiation may, in fact, reflect localized adaptation. This recent research has found that three divergent populations from the Pigeon/French Broad river drainage, when stocked together into LeConte Creek near Gatlinburg, TN in 1999, established a reproducing population. However, fin clips sampled from LeConte Creek brook trout in 2006, genotyped at 13 microsatellite loci, and subjected to parentage assignment analysis suggested that fish from the three distinct streams may not have mated randomly but instead exhibited positive (selective) assortative mating. Alternatively, if all fish in LeConte Creek are randomly mating but only matings of fish from the same source population are viable, this suggests the presence of a post-reproductive isolating mechanism. Either of these mechanisms, if confirmed, could have a significant impact on future supplementation and restoration plans for brook trout throughout the species' range. Additional populations have been reestablished within GRSM using fish from multiple genetically divergent source populations.

The independent evolution of adaptively divergent phenotypes among closely related lineages is most likely the consequence of natural selection (King and Wilson 1975). Because natural selection ultimately acts on the genetic variation underlying character variation, identifying the genes associated with parallel evolutionary changes among recently diverged lineages is essential to uncovering candidate genes implicated in adaptive phenotypic variance (Lee and Mitchell-Olds 2006). Contemporary genomic approaches can be particularly useful for identifying genetic targets of selection and genetic mechanisms of adaptation even among recently diverged lineages (Derome et al. 2006; Roberge et al. 2006). For example, comparative genomic expression profiles generated by cDNA microarrays (or gene chips) are a major first step for identifying an initial pool of candidate genes that might be implicated in adaptation, particularly in non-model systems for which whole genomes sequences are unavailable. Specialized expression arrays containing genes related by function, tissue, or pathway are becoming widely adopted by researchers studying the adaptive significance of genetic differentiation. Given that few gene sequences are available for brook trout, an analysis of global gene regulation may be the most direct avenue to identifying diagnosable genetic polymorphisms between genetically divergent populations occurring within watersheds and to understanding the evolutionary differences that exists between recently diverged lineages (Miller et al. 2002; Rise et al. 2004).

The objectives of this research proposal are to: 1) use variation at microsatellite loci to determine whether the pattern of nonrandom assortment observed in LeConte Creek is observed in other streams within GRSM established with multiple genetically divergent populations; and to 2) utilize gene expression profiles generated from cDNA microarray technology (i.e., functional genomics) to determine whether the large genetic differences observed between and within drainages (Figure 1) are adaptive in nature or due to random genetic drift. This study will be the first of its kind for brook trout and will determine whether GRSM fisheries managers should focus their resources on genetic relatedness or demographics (i.e., determine the number of fish needed to sustain population viability). The latter objective will include a comparison of gene expression profiles among representatives of all major regional groupings of brook trout sampled throughout the species' native range. This comparison will not only allow the variation in gene expression profiles among GRSM brook trout to be observed in a range-wide context, but provide unique information to fisheries management on the adaptive nature of detectable genetic variation. Additionally, this information will insure that GRSM meets the intent of the enabling legislation of the National Park Service which is to "promote and regulate the use of Federal areas known as national parks, monuments, and reservations by such means and measures as conform to the fundamental purpose of said parks, monuments and reservations, which purpose is to conserve the scenery and the natural and historic objects and the wild life therein and to provide for the enjoyment of the same in such manner and by such means as will leave them unimpaired for the enjoyment of future generations" (NPS organic act, 16 USC 1).

Procedures/methods.

Research Design: Initially, we will focus on addressing the question of whether positive assortative mating or some other post-reproductive isolating mechanism is occurring in three other populations founded by fish of divergent origins. This will involve the collection of 50-60 fin clips of fish from each of three restored streams: 1) Sams Creek (stocked with Cosby, Greenbrier, and Indian Camp creeks fish); 2) Mannis Branch (stocked with Greenbrier and Indian Camp creeks fish); and 3) Bear Creek (stocked with Steeltrap, Forney, Little Huggins, and Bunches creeks fish). We will search for adaptive significance in the genetic divergence observed among these genetically distinct source populations via functional genomics by testing these distinct populations for differences in gene expression profiles. Genes demonstrating these differences will be identified (to function) and their sequence differences will be quantified.

Tissue Collections: Sample collections will be performed by NPS personnel with assistance of USGS personnel. A collection for microsatellite DNA analyses will consist of 50 – 60 individual brook trout sampled (via adipose fin

clipping) from each collection site. Adipose tissue will be placed into 2 ml vials containing 95% chilled ethanol. A collection for gene expression analyses will consist of 10 -15 fish from each of the seven source populations. Fish will be flash frozen in liquid nitrogen and/or placed on dry ice in the field and stored at -80 °C for extraction of RNA and subsequent microarray experimentation. Particular care will be given to minimize bias of size difference between streams because metabolic rate is expected to show negative allometry. All fish tissue will then be transferred to the USGS-Leetown Science Center, Kearneysville, WV.

DNA and RNA Isolations: Genomic DNA will be extracted from the adipose fin tissue using the Puregene DNA extraction kit (Gentra Systems, Inc., Minneapolis, MN) and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). DNA concentrations will be determined by fluorescence assay and integrity of the DNA will be visually inspected in 1% agarose gels. Tissue-specific expression will be assayed to maximize the likelihood of detecting gene expression differences between brook trout from each of the distinct source populations. Tissues (e.g., white muscle, liver, kidney, brain) will be ground using baked (220°C, 5 h) mortars and pestles under liquid nitrogen, then total RNA will be extracted in TRIzol reagent (Invitrogen) according to the manufacturer's recommendations and the methods reported by Rise et al. (2004). Quantity and quality of RNA will be assessed using an ND-1000 small-volume spectrophotometer (Nanodrop Technologies) and the 2100 Bioanalyzer microfluidics electrophoresis system (Agilent Technologies), respectively.

Microsatellite DNA Survey: Thirteen polymorphic microsatellite DNA markers developed by the principal investigator (T. King unpublished) will be screened in brook trout from the restored populations. Amplification reactions will consist of 100-200 ng genomic DNA, 1X PCR buffer (10mM Tris-HCl, pH 8.3, 50 mM KCl), 2 mM MgCl₂, 0.25 mM dNTPs, 0.5 µM forward and reverse primer, and 1.0 U AmpliTaq® DNA polymerase in a total volume of 10-20 µl. The forward primer will be 5' modified with NED, FAM, or HEX fluorescent labels. Amplifications will be carried out on PTC-200 Thermal Cyclers under the following conditions: initial denaturing at 94 °C for 2 min; 35 cycles of 94 °C denaturing for 30 sec, 56 °C annealing (Table 1) for 30 sec, 72 °C extension for 1 min; and a final extension at 72 °C for 10 min. Amplified labeled products will be subjected to capillary electrophoresis on ABI Prism™ 3100 Genetic Analyzers. Genescan™ 2.1 Analysis software and Genotyper™ 3.6 Fragment Analysis software will be used to score, bin, and output allelic (and genotypic) data as described in King et al. (2001).

Statistical Analyses: Estimates of individual pair-wise genetic distances, using the proportion of shared alleles algorithm, will be calculated using a 32-bit version of Microsat 1.5d (Eric Minch, Stanford University). GENEPOP (v. 3.1) (Raymond and Rousset 1995) will be used to test for conformation to Hardy-Weinberg expectations, genotypic linkage disequilibrium, population differentiation (e.g., re-established versus source populations), and to calculate F_{ST} statistics. All probability values will be adjusted for multiple simultaneous table-wide tests using the sequential Bonferroni technique to minimize type-I statistical inference errors (Rice 1989). CERVUS 3.0 (Tristan Marshall, Univ. of Edinburgh) will be used to conduct a lineage analysis using an offspring's (i.e., individual from re-established population) microsatellite data to test against that of candidate lineages from each source population. This analysis will select the two individuals from all source populations that are most closely related to each test individual.

RNA isolation, characterization, and labeling –RNAs obtained from these preparations will be used for generating labeled targets for microarray hybridizations. All hybridization experiments will be performed using the SuperScript Indirect cDNA Labeling System kit and instructions (Invitrogen). Briefly, 5.0 µg total RNA will be reverse transcribed using an anchored oligo d(T)20 primer in cDNA synthesis reactions that incorporates aminoallyl- and aminohexyl-modified nucleotides. The modified cDNAs will then be labeled with fluorescent dyes in reactions with the amino-functional groups in coupling buffer.

Microarray description, experimental design, and hybridization - Transcriptome profiles will be used to identify orthologous genes between brook trout and Atlantic salmon. A 16,006 cDNA gene microarray developed by a consortium comprising groups from Canada (Genomics Research on Atlantic Salmon Project; GRASP), the U.S., France, and Norway will be utilized (Rise et al. 2004). The GRASP microarray consists of Atlantic salmon (13,421) and rainbow trout (2,576) cDNA features selected for immune-specific and reproduction-relevant genes as well as a wide variety of different classes of genes with a preponderance of expressed sequence tags (ESTs) randomly selected based on EST cluster quality and uniqueness (von Schalburg et al. 2005). Microarray experiments will be designed to comply with MIAME guidelines (Brazma et al. 2001). To incorporate sufficient reproducibility in these competitive hybridization gene expression experiments, a hierarchical replication strategy will be utilized. Multiple subjects per

experimental group (e.g., tissue type/stream of origin) will provide biological replication. To minimize technical variability, all targets will be synthesized in one round and each hybridization experiment will be conducted simultaneously on slides from a single batch where possible. Each hybridization experiment will also include dye-flips. Fluorescent images of hybridized arrays will be acquired immediately at 10 μ m resolution using ScanArray Gx Microarray Scanner (PerkinElmer). For these competitive hybridization (i.e., cross-species) experiments, the hybridization performance of labeled targets to salmonid features will be assessed as a percentage of features bound from the numbers of Atlantic salmon and rainbow trout features passing a hybridization signal threshold, defined as two standard deviations above Arabidopsis signal mean (von Schalburg et al. 2005). Transformations or normalizations will be performed on these data as needed. All microarrays will be preprocessed, hybridized, and washed according to the methods of von Schalburg et al. (2005) and any additional recommendations for the GRASP microarray found at <http://web.uvic.ca/cbr/grasp/>.

Validation of Gene Expression - All brook trout genes successfully hybridizing to the GRASP microarray and exhibiting variation in expression will be validated via real-time (RT) polymerase chain reaction (PCR). Real-time PCR will be performed on an ABI7900HT (Applied Biosystems) utilizing one or more reference genes (e.g., glyceraldehyde 3-phosphate dehydrogenase, β -actin, or RNA polymerase II) for quantitation (Radonic et al. 2004).

Literature Cited.

- Brazma A and Co-authors (2001) Minimum information about a microarray experiment (MIAME): Toward standards for microarray data. *Nature Genetics* 2001, **29**, 365-371.
- Derome N, Duchesne P, Bernatchez (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Molecular Ecology*, **15**, 1239-1249.
- King M-C, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science*, **188**, 107-116.
- King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski BA (2001) Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology*, **10**, 807-821.
- Lee EC, Mitchell-Olds T (2006) Preface to the special issue: ecological and evolutionary genomics of populations in nature. *Molecular Ecology*, **15**, 1193-1196.
- Miller NA, Gong Q, Bryan R, Ruvolo M, Turner LA, LaBrie ST (2002) Cross-hybridization of closely related genes on high-density macroarrays. *BioTechniques*, **32**, 620-625.
- Radonic A, [Thulke S](#), [Mackay IM](#), [Landt O](#), [Siegert W](#), [Nitsche A](#) (2004) *Guideline to reference gene selection for quantitative real-time PCR*. *Biochem. Biophys. Res. Commun.*, **313**, 856-62.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.
- Rise ML and Co-authors (2004) Development and application of a salmonid EST database and cDNA microarray: Data mining and interspecific hybridization characteristics. *Genome Research*, **14**, 478-490.
- Roberge C, Einum S, Guderley H, Bernatchez L (2006) Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Molecular Ecology*, **15**, 9-20.
- von Schalburg KR, Rise ML, Cooper GA, Brown GD, Gibbs AR, Nelson CC, Davidson WS, Koop BF (2005) Fish and chips: Various methodologies demonstrate utility of a 16,006-gene salmonid microarray. *BMC Genomics*, **6**, 126, 1-8.

Expected outcomes or products. This research, if funded, would represent the first attempt at determining whether GRSM fisheries managers should focus their resources on genetic or demographic factors. More specifically, are the observed genetic differences due to natural selection (i.e., have an adaptive significance) or neutral processes (e.g., random genetic drift). If the former, the adaptive significance of the observed genetic differentiation should be considered on a case-by-case basis. If the latter, resources should be focused on determining the number of individuals required to restore a viable population. Multiple peer-reviewed manuscripts will result from this research including: 1) a description patterns of genetic variation among the GRSM populations surveyed for microsatellite variation; 2) a quantitative and qualitative description of the amount of cross-hybridization between brook trout and the Atlantic salmon microarray chip; and 3) a paper describing the level of differentiation between southern Appalachian and other regional forms of brook trout and a recommendation on the need (or not) of a taxonomic revision.

Technology/information transfer. This research effort should provide park resource managers with a clear delineation of management units and allow assessment of conservation priorities from an evolutionary perspective. Any data suitable for dissemination as metadata will be prepared in compliance with the Federal Geographic Data Committee's

(FGDC) Content Standards for Digital Geospatial Metadata and the National Biological Information Infrastructure (NBII) Metadata Clearinghouse. Following NBII guidelines, the results will be made available through the internet via a T1 line. All manuscripts and reports will be authored by the GRSM and USGS Principal Investigators and will conform to the peer-review requirements of each agency. Microarray data will be deposited in the NCBI Gene Expression Omnibus database at <http://www.ncbi.nlm.nih.gov/geo/>.

Work schedule. FY08: Begin tissue collections (adipose-fine) from brook trout in Sams Creek, Mannis Branch, and Bear Creek; perform DNA extractions, PCR, and genotyping at 13 microsatellite DNA loci; begin collection of whole fish to be frozen in liquid nitrogen for RNA extraction; begin hybridizations of RNA to Atlantic salmon microarray chip; provide a status report to USGS and GRSM at the end of fiscal year.

FY09: Continue and complete all tissue and specimen sampling; complete DNA/RNA extractions, and continue PCR and genotyping at microsatellite DNA loci; complete microarray hybridizations, RT-PCR, sequencing and data analysis; provide a status report at end of fiscal year.

FY10: Complete final report and manuscripts by March 2010; present results to GRSM managers and at scientific meetings.

Personnel. **Tim L. King**, Research Fishery Biologist (Genetics), USGS-Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430; Phone: 304.724.4450; Facsimile: 304.724.4424; E-mail: tlking@usgs.gov

John F. Switzer, Fishery Biologist (Genetics), USGS-Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430; Phone: 304.724.4574; Facsimile: 304.724.4424; E-mail: jswitzer@usgs.gov

Stephen E. Moore, Supervisory Fishery Biologist, Great Smoky Mountains National Park, 107 Park Headquarters Road, Gatlinburg, TN 37738; Phone: 865.436.1250; e-mail: steve_e_moore@nps.gov

Matt Kulp, Fishery Biologist, Great Smoky Mountains National Park, 107 Park Headquarters Road, Gatlinburg, TN 37738; Phone: 865.436.1254; e-mail: matt_kulp@nps.gov

Budget.

Activity/Purchase	Costs		
	FY08	FY09	FY10
Salaries (Contract employees) - molecular Fishery Biologist (3 months/year)	\$17,800	\$18,700	\$0
GRSM field assistance (3 months FY08; 1 month FY09)	\$ 8,000	\$ 3,000	
Expendable supplies			
DNA/RNA extraction, fragment and microarray analysis	\$11,000	\$ 5,600	\$0
Subtotals/Year	\$36,800	\$27,300	\$0
USGS-Leetown overhead (16.801%)	\$ 6,180	\$ 4,590	\$0
Total per year	\$42,980	\$31,890	\$0
Total request from POBS program	\$74,870		