Development of an amplified fragment length polymorphism (AFLP) test to distinguish Colorado River from Rio Grande cutthroat trout

December 12, 2011

Kevin B. Rogers[1], Janet Epp[2], and John Wood[2]

[1]Aquatic Research, Colorado Parks and Wildlife, P.O. Box 775777, Steamboat Springs, CO 80477, kevin.rogers@state.co.us; [2]Pisces Molecular, 1600 Range St., Suite 201, Boulder, CO 80301

Introduction

Development of range-wide status assessments coupled with recent advances in molecular techniques have spawned a proliferation of research geared toward resolving the taxonomic history of Colorado's native cutthroat trout (Rogers 2008a). Interest in developing ways to characterize our native subspecies was initiated with the listing of the greenback cutthroat trout (*Oncorhynchus clarkii stomias*) under the Endangered Species Act in 1973, and has continued unabated since. Reports provide taxonomic information on subspecies identification and purity that has been used to guide management of the three remaining subspecies native to Colorado. These assessments employed a variety of techniques, ranging from morphometric and meristic work in the early years, to characterization of mitochondrial DNA exploring a host of different of genes. More recent work has focused on allozyme and nuclear DNA tests including paired interspersed nuclear elements (PINEs), bi-allelic markers (BIAMs), restriction fragment length polymorphisms (RFLPs), microsatellites, and amplified fragment length polymorphisms (AFLPs).

Early molecular methods proved capable of identifying introgression with rainbow trout (*O. mykiss*) or Yellowstone cutthroat trout (*O. c. bouvieri)*, but separation of the three native cutthroat trout subspecies found in Colorado remained elusive. Morphometric and meristic treatments were often inconclusive, and diagnostic molecular markers could not be identified with traditional taxonomic approaches. It was hoped that with their ability to assess variability at thousands of anonymous loci, that AFLPs (Vos et al. 1995) could serve as a powerful tool for detecting divergence among closely related populations (Bernatchez and Duchesne 2000). In fact, AFLPs are ideal when low genetic variability such as would be expected with intraspecific comparisons (Holland et al. 2008) as is the case with cutthroat trout. Since the variable loci are distributed throughout the genome, AFLPs are well suited to represent that diversity, as well as provide resolution when results from traditional DNA methods are unsatisfactory (Flannery et al. 2007). In addition, AFLPs have been shown to be more efficient than

microsatellite loci in discriminating the source of an individual among putative populations making them especially beneficial in systems characterized by weak population structuring (Campbell et al. 2003). Colorado Parks and Wildlife (CPW) has used the AFLP test developed at Pisces Molecular (Boulder, CO) to assess taxonomic status and purity since 2007 (Rogers 2008a). This test was developed specifically to distinguish Lineage CR from Lineage GB along with any Yellowstone cutthroat trout or rainbow trout admixture. As such, primers were selected that provided substantial polymorphisms between the two. Unfortunately, these markers do not provide good separation between Lineage CR and Rio Grande cutthroat trout (Figure 1). A new test targeted for making this distinction was needed, and is the focus of this report.

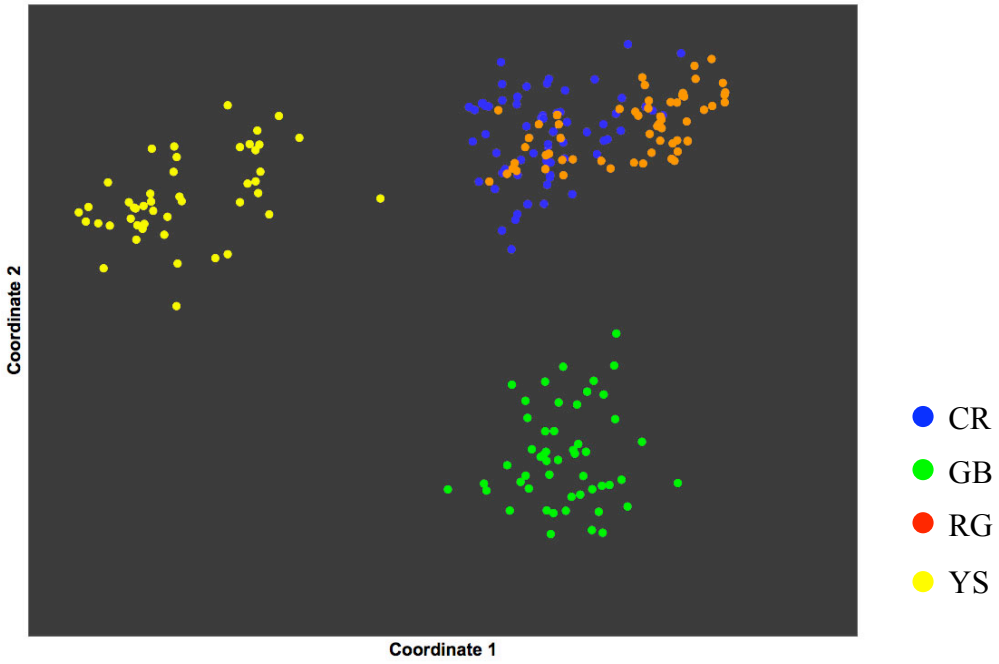

FIGURE 1: Separation of cutthroat trout subspecies across the primary and secondary principle components using AFLPs and similar reference populations (from Metcalf 2007). Notice the tight clustering of Lineage CR (blue) with Rio Grande cutthroat trout (red) but good separation from Lineage GB (green) and Yellowstone cutthroat trout (yellow).

Methods

Cutthroat trout tissues were obtained during routine sampling of populations across Colorado using standard protocols (Rogers 2007). Trout DNA was extracted from fin clips using a proteinase K tissue lysis and spin-column DNA purification protocol (Qiagen DNeasy). AFLP marker fragments were generated using restriction digested DNA (EcoR1 and MseI) and a single pair of +3 PCR primers (ACT for the FAM-labeled

forward primer; CAG for the reverse primer). Fragments were separated and sized on an ABI 3130 DNA sequencer (Applied Biosystems, Foster City, California). Using the program Genemapper 4.0 (Applied Biosystems), a genetic fingerprint was produced for each individual sample by scoring for the presence or absence of a standardized set of 149 markers between 50 and 450 base pairs in size generated from reference cutthroat trout populations (Table 1). This approach is analogous to the procedure used to develop the traditional $AFLP_{Standard}$ test (Rogers 2008a) that compares a population of interest to reference populations of rainbow trout, Yellowstone cutthroat trout, Rio Grande cutthroat trout, and both varieties of Colorado River cutthroat trout (Lineage CR and Lineage GB). The genetic fingerprints of individuals in the test population were compared to those found in the reference populations (Rogers 2008a) using a Bayesian approach for identifying population clusters (Pritchard et al. 2000). An advantage of this approach is that population membership and admixture can be assessed for each individual (Anderson et al. 2008). Reference populations were selected based on previous molecular work to determine purity and historic stocking data. The similarity or dissimilarity with reference populations was scored as the probability (q) that each test individual shares a genetic background with the cutthroat subspecies reference population groups used (Table 1) with the program STRUCTURE 2.2 (Falush et al. 2007; Pritchard et al. 2007). Average q values from the run with the highest log likelihood (Pritchard and Cowley 2007) were used to generate the admixture proportions for the unknown population.

TABLE 1: Reference populations used to distinguish between Lineage CR and Rio Grande cutthroat trout with AFLPs.

Trout Subspecies	Water	County	Water Code	Collection Date	Sample Size
Lineage CR[a]	Williamson Lake #3	Inyo	NA[b]	08/31/06	29
	Piedra, E Fk	Hinsdale	42096	02/07/06	20[c]
	Slater Crk, S Fk	Routt	23286	NA	14[d]
	Parachute Crk, E Fk	Garfield	21460	NA	10[d]
Rio Grande cutthroat	Canones Crk	Rio Arriba	329	03/29/06	19
	Columbine Crk	Taos	1026	09/17/02	20[e]
	Osier Crk	Conejos	44444	09/22/04	11
	Cuates Crk	Costilla	38141	07/25/05	10

[a]Lineage reported as Colorado River cutthroat trout by Metcalf et al. (2007)
[b]These tissues collected from the Williamson Lakes in California by K. Rogers
[d]These DNAs obtained from J. Metcalf and were used in her dissertation work at CU Boulder.
[d]These DNAs obtained from D. Shiozawa via J. Metcalf, and used in her dissertation work at CU
[e]These DNAs obtained from V. Pritchard and were used in her dissertation work at New Mexico State University

Populations selected for further study in this report (Table 2), were believed to be aboriginal based on previous work, but that suggested either RGCT admixture in putative CRCT populations, or CRCT admixture in putative RGCT populations based on recent $AFLP_{Standard}$ work (Rogers 2008a). Confidence intervals were assessed around admixture with the software application QSTRAP Version 3.1 (Rogers 2008b). This program uses a bootstrapping approach to derive confidence intervals around mean values of q by taking a sample of n q-values drawn with replacement from the pool of q-values generated by STRUCTURE for the unknown population of interest based on known reference populations (Table 1) with which admixture is to be evaluated (Table 2). This process was iterated 10,000 times, and the ordered mean values were plotted. The bounds of the middle 95% of values then reflect the upper and lower confidence limits around the mean. Lucid discussions of this approach can be found in Efron and Tibshirani (1986) and Manly (1997).

TABLE 2: Populations used to test the efficacy of the new $AFLP_{RG-CR}$ test.

Subspecies	Water	County	Water Code	Collection Date	Sample Size
CRCT	Bunker Creek	Rio Blanco	19364	08/30/05	20
	Deer Creek	Moffat	20185	08/17/05	24
	Johnson Creek	Routt	20802	02/07/06	30[a]
RGCT	Alder Creek, W. Fk	Rio Grande	47755	10/05/05	13
	Columbine Creek	Taos	1026	09/17/02	20[b]
	Rhodes Gulch	Conejos	43840	05/24/04	14

[a]These contain seven fish from Burton Creek, a tributary to Johnson Creek
[b]These DNAs obtained from V. Pritchard and were used in her dissertation work at New Mexico State University

Given the heuristic nature of AFLPs, microsatellites, and the program STRUCTURE used to analyze both, it was felt that obtaining sequence data from the NADH dehydrogenase subunit 2 (ND2) mitochondrial gene would also be beneficial in determining if mild admixture could be attributed to diversity within a given lineage or whether it represented true admixture. This gene was demonstrated to contain considerable diversity among Colorado's native cutthroat trout and is useful for separating the subspecies (Metcalf et al. 2007). An aliquot of each sample DNA from the Rhodes Gulch collection was amplified using PCR primers specific to the ND2 mitochondrial gene of cutthroat trout (*Oncorhynchus clarkii*), generating a 648 bp fragment. After amplification, residual primers and dNTPs were removed or inactivated using Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP). Fluorescently-labeled DNA sequences in the forward and reverse direction for each sample were generated using a BigDye Terminator v3.1 kit (Applied Biosystems). After the BigDye reactions were completed, unincorporated fluorescently labeled nucleotides were removed using BigDye Xterminator (Applied Biosystems) according to the manufacturer's instructions. Samples were run on a capillary sequencer (Applied Biosystems 3130 Genetic Analyzer,

POP7 polymer, 36cm array). Sequence reads generated from the forward and reverse strands of each sample DNA were assembled using the Contig Express program (Vector NTI 11, Invitrogen). The assembled contiguous sequence chromatograms were examined for sequence quality and accuracy, and the primer sequences edited (removed) from the ends of the fragments. The sequence from each sample was reported in FASTA file format. Sequences were aligned in ClustalW (Chenna et al. 2003) and compared to a half dozen populations (Table 2) that represent the suite of genetic diversity found in Colorado's cutthroat trout in MEGA4 (Tamura et al. 2007). The evolutionary history was inferred using the Minimum Evolution (ME) method (Rzhetsky and Nei 1992). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) were calculated (Felsenstein 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004). The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (Nei and Kumar 2000) at a search level of one. The Neighbor-joining algorithm (Saitou and Nei 1987) was used to generate the initial tree.

TABLE 3.- A key to reference populations used to build an ND2 phylogenetic tree in which Rhodes Gulch samples could be compared.

Reference population	Pisces Code	Drainage	Water Code
Lineage CR[1]			
O. *clarkii pleuriticus* (Nanita)*	-	-	-
O. *clarkii pleuriticus* (Graneros)*	-	-	-
Williamson Lake #3	WLM	White	-
Piedra River, E Fk	EFP	San Juan	Durango[b]
Navajo River	NAV	San Juan	Durango[b]
Lineage GB			
O. *clarkii stomias* (Como)*	-	-	
Severy Creek	SEV	Arkansas	31312
Bobtail Creek	BTC	Williams Fork	23016
Cunningham Creek	CNC	Roaring Fork	23957
Bear Creek			
O. *clarkii stomias* (Bear)*	-	-	
Bear Creek	BEA2	Arkansas	29157
Rio Grande			
O. *clarkii virginalis* (Ricardo)*	-	-	-
O. *clarkii virginalis* (Torcido)*	-	-	-
Yellowstone			
O. *clarkii bouvieri* (LeHardy)*	-	-	-
O. *clarkii bouvieri**	-	-	-
Yellowstone River (LeHardy)	LEH	Yellowstone	-
Rhodes Gulch[c]	RDG	Rio Grande	43840

*These sequences obtained from GenBank

[a]These represent the lineage of Colorado River cutthroat trout similar to those native to the White and Yampa River basins (Trappers Lake fingerprint)

[b]From broodstock housed at the CPW's Durango Hatchery
[c]Only a limited amount of DNA remained in these samples obtained from Marlis Douglas at CSU, some of which was dehydrated

Results and Discussion

Using the reference populations described above (Table 1), we found 149 polymorphic fragments (loci) between RGCT and CRCT populations, which were grouped as a binset used to score all subsequent samples. When reference populations were analyzed as unknowns in GeneMapper and STRUCTURE using this binset, all individual fish assigned to their respective taxa (Figure 2) with only very minor probability of assigning to the opposite taxa in a single sample from Columbine Creek (RGCT) and a few from South Fork Slater Creek (CRCT).

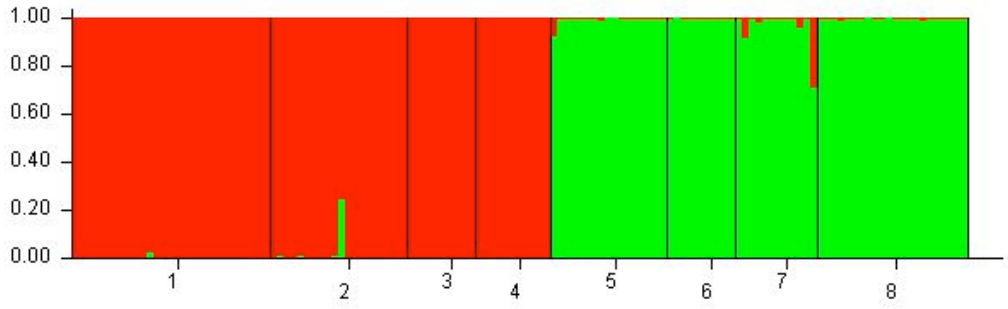

Figure 2: RGCT (Populations 1-4) and CRCT (Populations 5-8) reference populations analyzed as unknowns with the 149-allele binset in GeneMapper

When using the same reference populations (Table 1) and the 149 allele binset to investigate apparently spurious results obtained with the traditional 119 allele binset, we obtained results more in line with expectation. Putative CRCT populations that showed some level of RGCT admixture under the $AFLP_{Standard}$ 119 allele binset (Figures 3-5) looked to be pure CRCT when subjected to the specific $AFLP_{RG\text{-}CR}$ test (149 allele binset). Results were not as clear when putative RGCT populations were investigated. The $AFLP_{RG\text{-}CR}$ while much improved, failed to remove all evidence of CRCT admixture (Figures 6-8). Results for Alder and Columbine Creek were much improved, Rhodes Gulch remained predominantly CRCT (Table 4). Although a putative population of pure RGCT, suspicion of genuine admixture led us to sequence the ND2 mitochondrial gene to determine if any CRCT haplotypes were present. Adequate DNA was only available from 11 fish, but they did provide reliable sequence data despite substantial desiccation in some samples. Indeed, a common CRCT haplotype was detected in one of the samples (Figure 9), suggesting that legitimate admixture with CRCT is present.

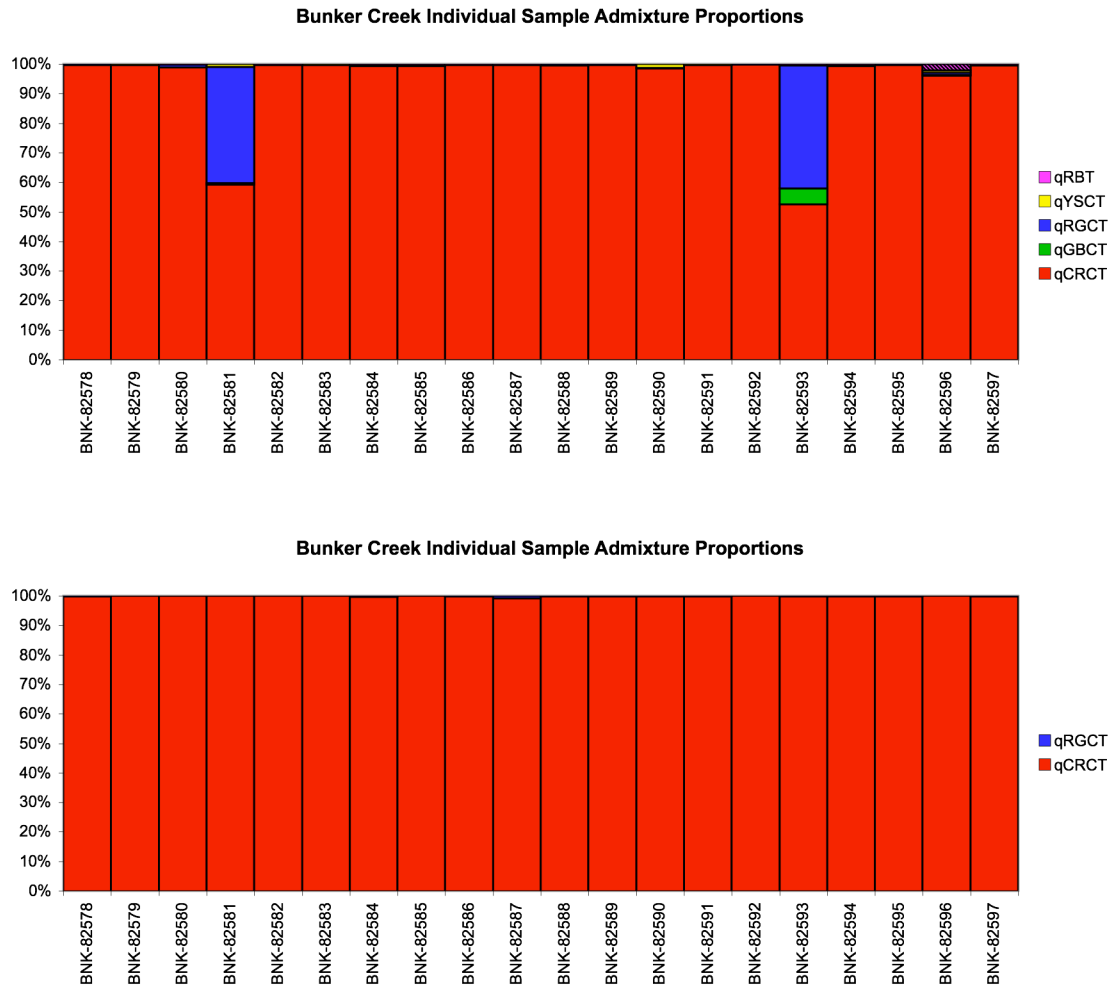

FIGURE 3: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 20 samples collected on 8/30/05 from Bunker Creek (Pisces sample numbers 82578-82597) are shown on top while results for the same samples tested with the $AFLP_{RG\text{-}CR}$ test using the 149 allele binset are shown below.

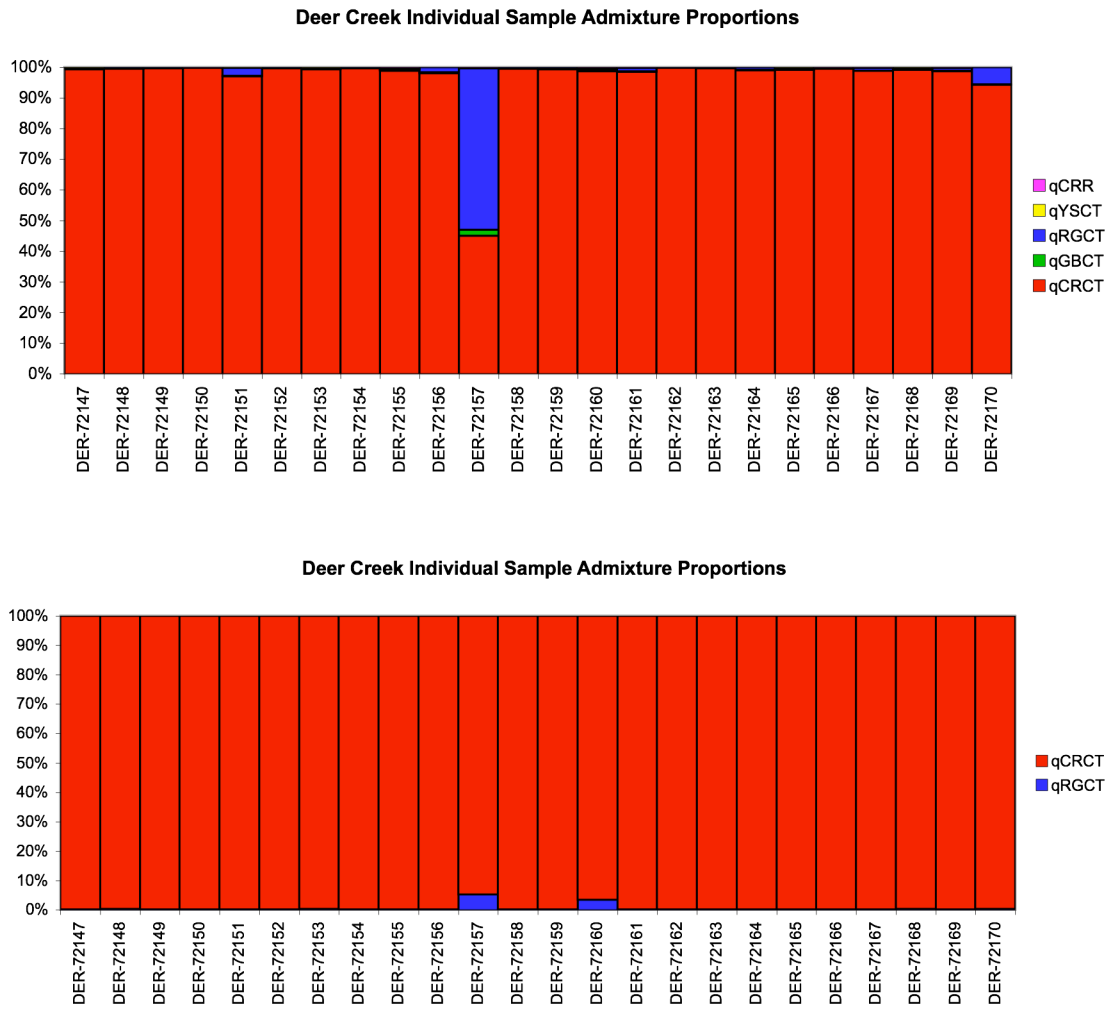

FIGURE 4: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 24 samples collected on 8/17/05 from Deer Creek (Pisces sample numbers 72147-72170) are shown on top while results for the same samples tested with the $AFLP_{RG\text{-}CR}$ test using the 149 allele binset are shown below.

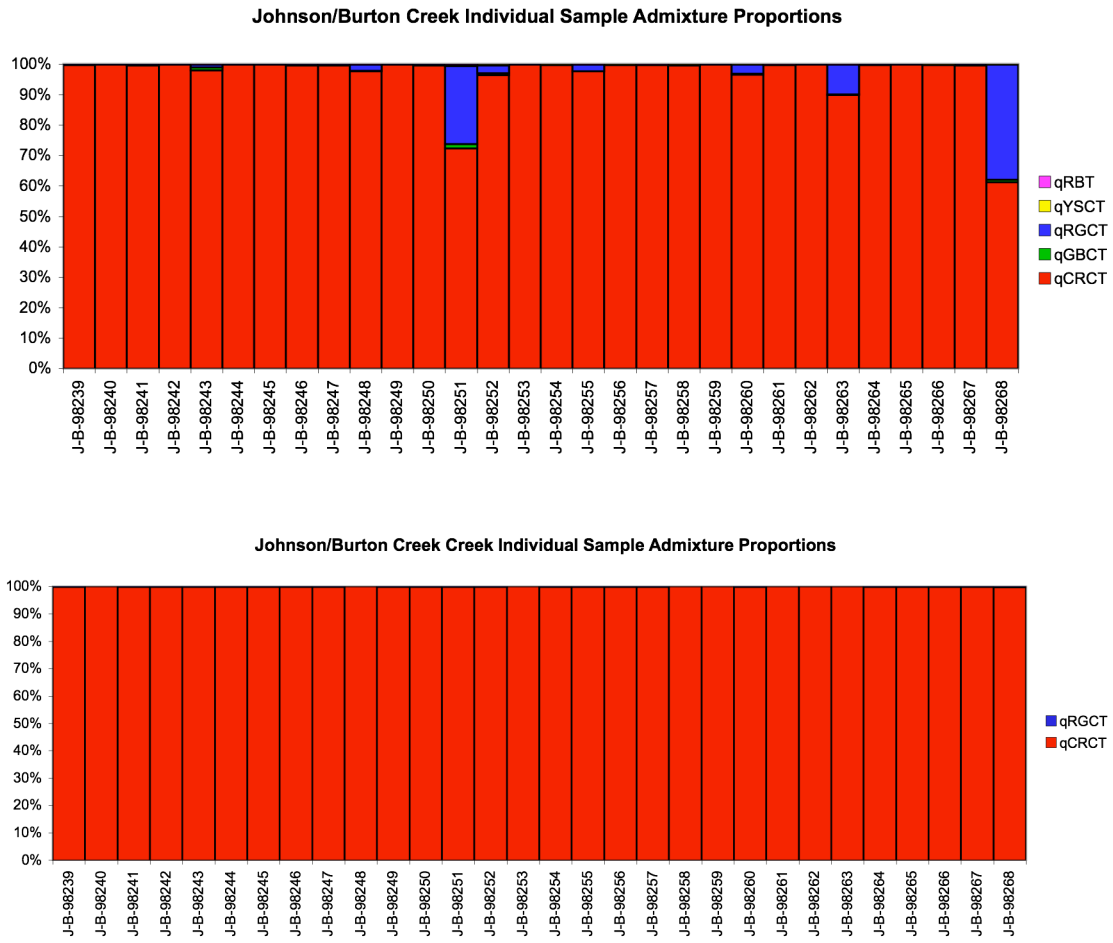

FIGURE 5: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 30 samples collected on 10/2/08 from Johnson and Burton Creeks (Pisces sample numbers 98239-98268) are shown on top while results for the same samples tested with the $AFLP_{RG\text{-}CR}$ test using the 149 allele binset are shown below. The first 23 samples came from Johnson Creek while the last seven came from its tributary, Burton Creek.

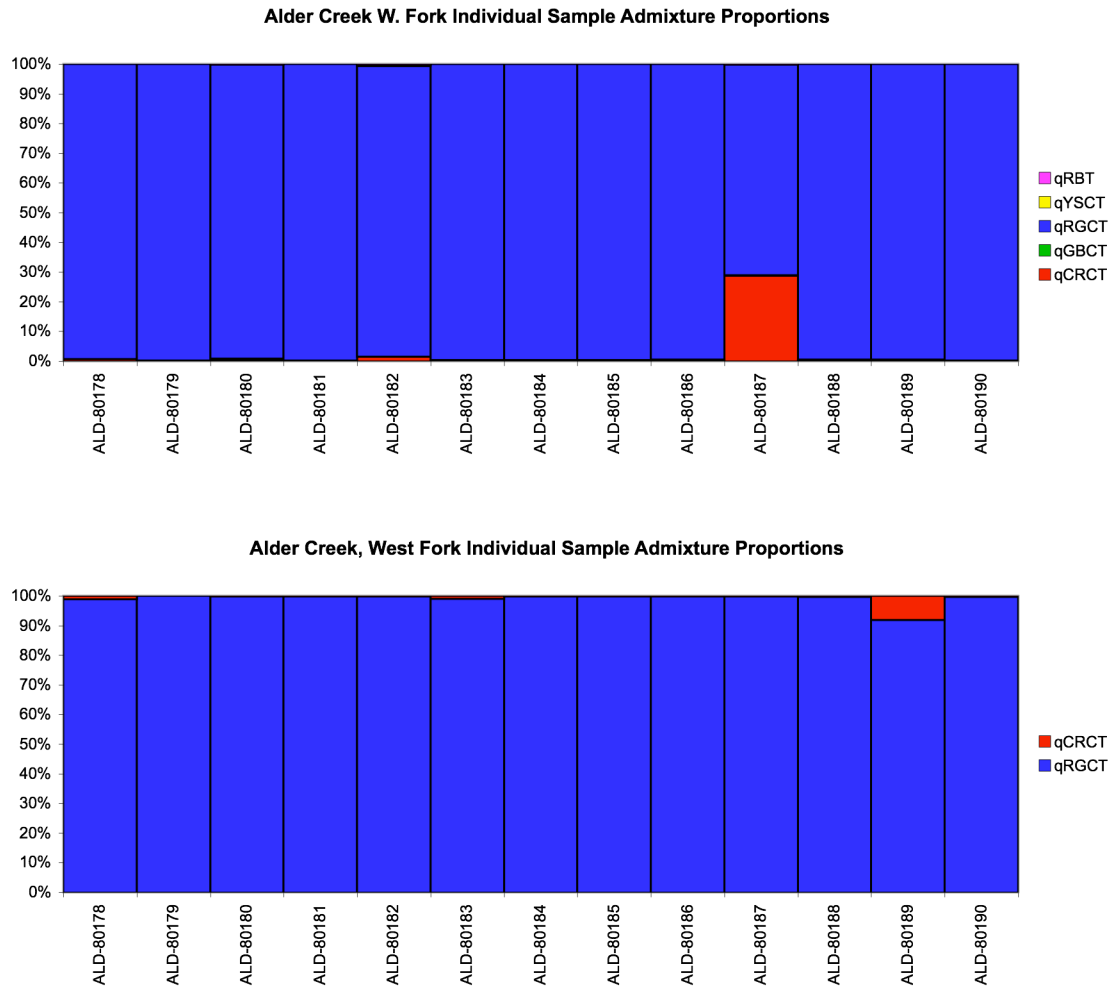

FIGURE 6: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 13 samples collected on 10/5/05 from the West Fork of Alder Creek (Pisces sample numbers 80175-80190) are shown on top while results for the same samples tested with the $AFLP_{RG\text{-}CR}$ test using the 149 allele binset are shown below. Three samples were excluded from this collection due to DNA degradation.

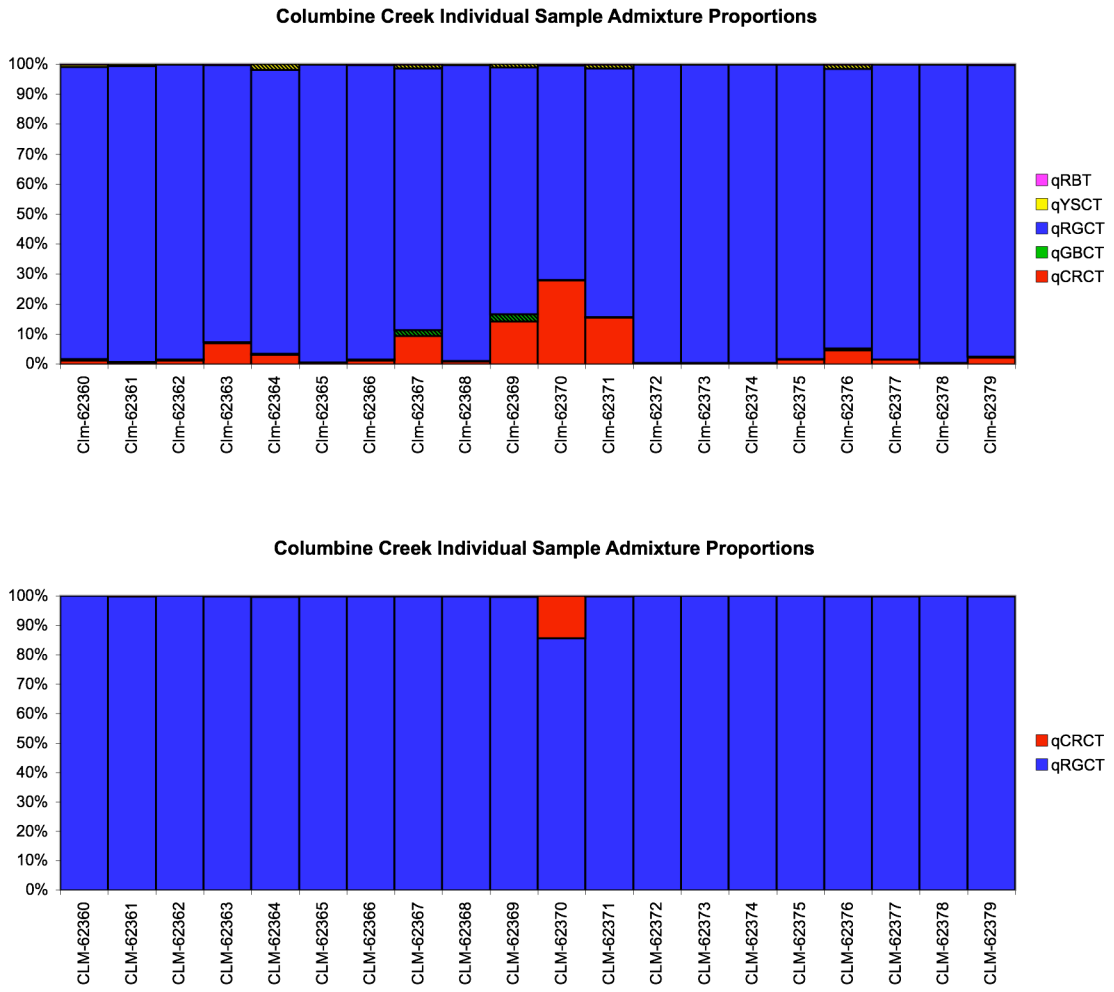

FIGURE 7: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 20 samples collected on 9/17/02 from Columbine Creek (Pisces sample numbers 62360-62379) are shown on top while results for the same samples tested with the $AFLP_{RG\text{-}CR}$ test using the 149 allele binset are shown below.

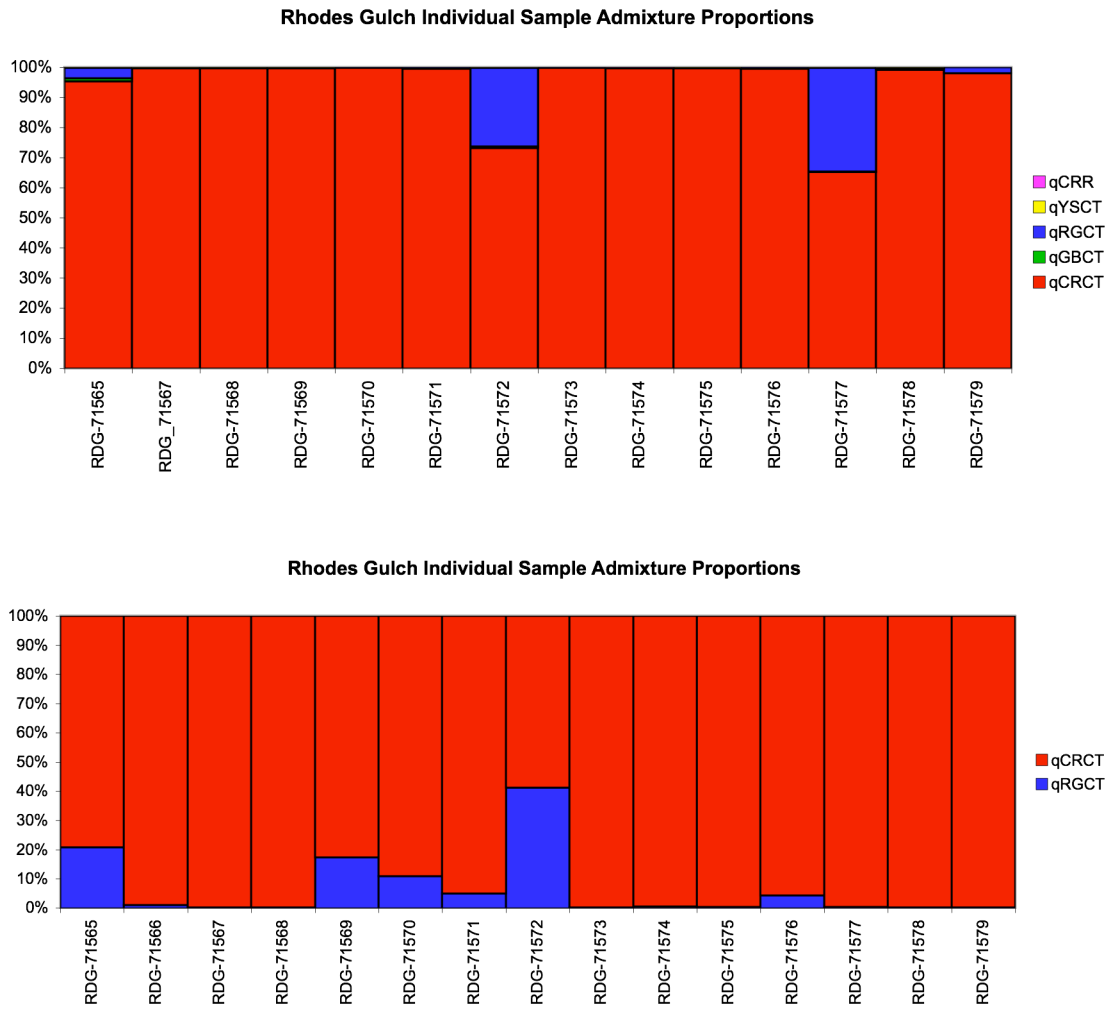

FIGURE 8: $AFLP_{Standard}$ results obtained using the traditional 119 allele binset from 14 samples collected on 5/24/04 from Rhodes Gulch (Pisces sample numbers 71565-71579) are shown on top while results for the same samples tested with the $AFLP_{RG-CR}$ test using the 149 allele binset are shown below..

TABLE 4: A summary of populations used to test the efficacy of the new $AFLP_{RG\text{-}CR}$ test compared to the traditional $AFLP_{Standard}$ test using population mean q-values and associated 95% confidence intervals calculated with QSTRAP Version 3.1.

Water	**qCRCT (95% CI)**	**qRGCT (95% CI)**
Bunker Creek AFLP	95 (88-99)	4 (0-10)
Bunker Creek RG-CR	100 (100-100)	0 (0-0)
Deer Creek AFLP	97 (92-99)	3 (1-8)
Deer Creek RG-CR	99 (99-100)	1 (0-1)
Johnson Creek[a] AFLP	97 (93-99)	3 (1-6)
Johnson Creek[a] RG-CR	100 (100-100)	0 (0-0)
Alder Creek, W Fk AFLP	3 (0-7)	97 (93-100)
Alder Creek, W Fk RG-CR	1 (0-2)	99 (98-100)
Columbine Creek AFLP	5 (2-8)	94 (91-97)
Columbine Creek RG-CR	1 (0-2)	99 (98-100)
Rhodes Gulch AFLP	95 (88-99)	5 (0-11)
Rhodes Gulch RG-CR	93 (87-98)	7 (2-13)

[a]The last seven samples came from the tributary Burton Creek

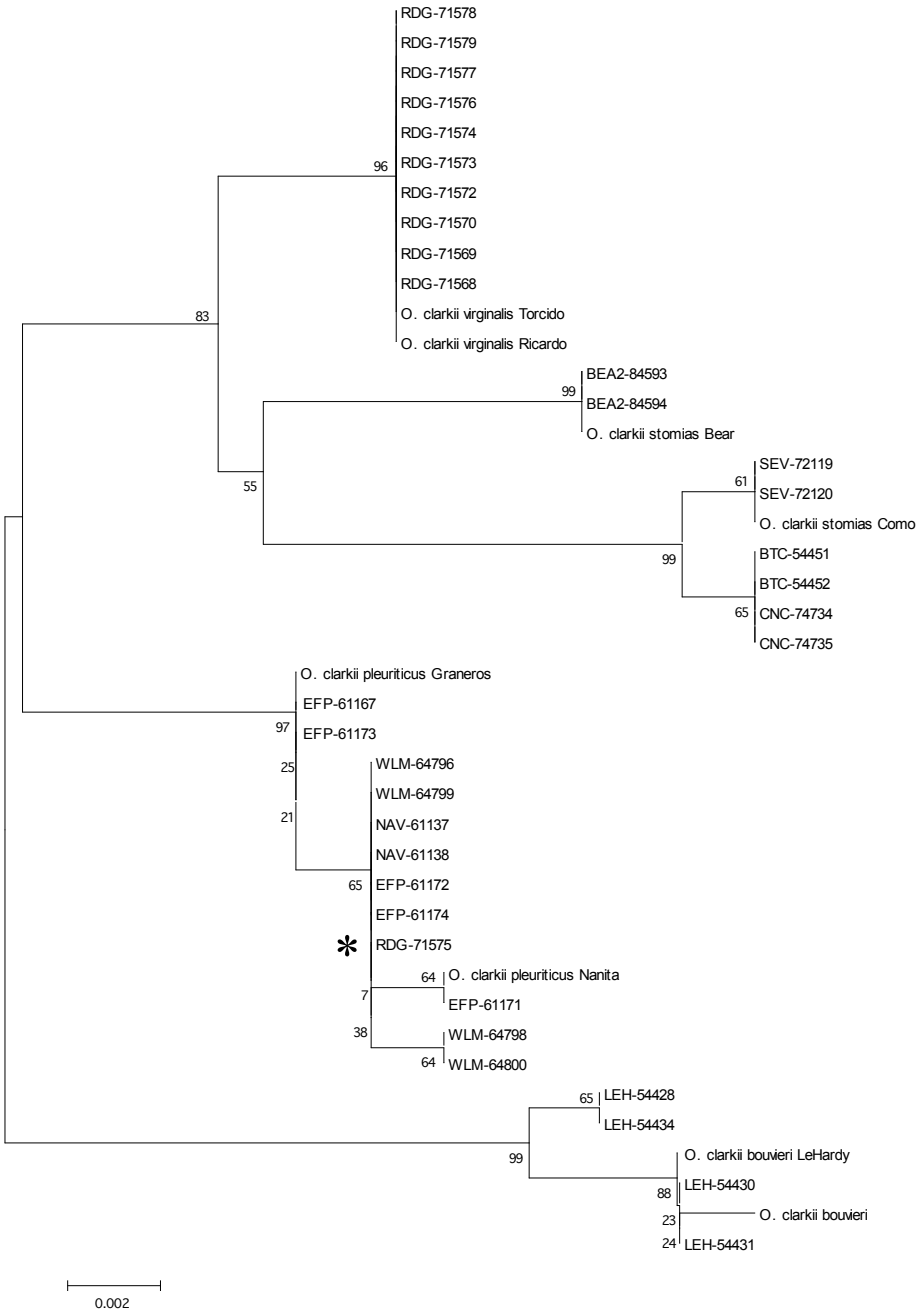

FIGURE 9: The evolutionary history of these samples (Table 3) was inferred using the Minimum Evolution method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method, and are in the units of the number of base substitutions per site. Pisces sample number RDG-71575 is labeled with an asterisk.

Conclusion

As expected, development of a targeted binset resulted in the ability to distinguish Colorado River cutthroat trout from Rio Grande cutthroat trout using the AFLP methodology. While evidence of RGCT admixture in putative CRCT populations was erased by this new test, two of the three putative RGCT populations (Table 2) still showed very slight evidence of introgression, suggesting that further testing should be conducted to ensure confidence in the ability of this new test to accurately distinguish apparent CRCT admixture in RGCT populations from background noise. The third population (Rhodes Gulch) registered substantial CRCT influence (Figure 8) and also harbored a CRCT mitochondrial haplotype in the ND2 gene (Figure 9) suggesting that this population indeed harbored CRCT admixture and was therefore not a suitable test case for this study.

Literature cited

ANDERSON, J. D., W. J. KAREL, K. A. ANDERSON, AND P. A. ROPER-FOO. 2008. Genetic assessment of sheepshead stock structure in the northern Gulf of Mexico: morphological divergence in the face of gene flow. North American Journal of Fisheries Management 28:592-606.

BERNATCHEZ, L, AND P. DUCHESNE. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? Canadian Journal of Fisheries and Aquatic Sciences 57:1-12.

CAMPBELL, D., P. DUCHESNE, AND L. BERNATCHEZ. 2003. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. Molecular Ecology 12:1979-1991.

CHENNA, R., H. SUGAWARA, T. KOIKE, R. LOPEZ, T. J. GIBSON, D. G. HIGGINS, AND J. D. THOMPSON. 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research 31:3497-3500.

EFRON, B. AND TIBSHIRANI, R. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science 1:54-77

FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.

FLANNERY, B. G., J. K. WENBURG, A. J. GHARRETT. 2007. Variation of amplified fragment length polymorphisms in Yukon River chum salmon: population structure and application to mixed-stock analysis. Transactions of the Americaon Fisheries Society 136:911-925.

HOLLAND, B. R., A. C. CLARKE, H. M. MEUDT. 2008. Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. Systematic Biology 57:347-366.

MANLY, B. J. F. 1997. Randomization, bootstrap, and Monte Carlo methods in biology,2nd edition. Chapman and Hall, London.

METCALF, J. L., V. L. PRITCHARD, S. M. SILVESTRI, J. B. JENKINS, J. S. WOOD, D. E. COWLEY, R. P. EVANS, D. K. SHIOZAWA, AND A. P. MARTIN. 2007. Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. Molecular Ecology 16:4445-4454.

NEI, M., AND S. KUMAR. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

PRITCHARD, V. L., K. JONES, D. E. COWLEY. 2007. Estimation of introgression in cutthroat trout populations using microsatellites. Conservation Genetics 8:1311-1329.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure from multilocus genotype data. Genetics 155:945-959.

PRITCHARD, J. K., X. WEN, AND D. FALUSH. 2007. Documentation for structure software: Version 2.2. University of Chicago. Available: http://pritch.bsd.uchicago.edu/software. (June 2007).

ROGERS, K. B. 2007. A suggested protocol for collecting cutthroat trout tissues for subsequent genetic analysis. Colorado Divison of Wildlife. Fort Collins, Colorado.

ROGERS, K. B. 2008a. Using amplified fragment length polymorphisms to characterize purity of cutthroat trout in Colorado: results from 2007. Colorado Division of Wildlife, Fort Collins.

ROGERS, K. B. 2008b. QStrap: bootstrapping confidence intervals around admixture estimates from STRUCTURE output. Version 3.1. Colorado Division of Wildlife. Available: http://wildlife.state.co.us/research/aquatic/software. (April 2008).

RZHETSKY, A., AND M. NEI. 1992. A simple method for estimating and testing minimum evolution trees. Molecular Biology and Evolution 9:945-967.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425.

TAMURA K., J. DUDLEY, M. NEI M, AND S. KUMAR. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software Version 4.0. Molecular Biology and Evolution 24:1596-1599.

TAMURA, K., M. NEI, AND S. KUMAR. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences (USA) 101:11030-11035.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.