

To the University of Wyoming:

The members of the Committee approve the dissertation of William C. Rosenthal
presented on November 19, 2025.

Dr. Catherine Wagner, Chairperson

Dr. Annika Walters, Outside Member

Dr. David Tank

Dr. Christopher Weiss-Lehman

Dr. Robert Al-Chokhachy

Mark Smith, External Member

APPROVED:

Dr. Melanie Murphy, Director, Program in Ecology and Evolution

Dr. Carolyn Pepper, Interim Dean, Graduate Education

Rosenthal, William C. *Looking back to strive forward: genomics-informed history and conservation of Yellowstone cutthroat trout*, Ecology and Evolutionary Biology Ph.D., Program in Ecology and Evolution, August, 2026.

Genomic data opens many doors for exploring the history and current status of populations. In this document I utilize genomic data to explore the recent and ancient history of Yellowstone cutthroat trout: an imperiled subspecies of cutthroat trout native to Wyoming. First I explore the legacy of over a century of fish stocking efforts across the Yellowstone cutthroat trout range and find that these actions have left an indelible mark on many populations. However, I also uncover evidence that the admixture from stocking sources may not have affected parts of the genome involved in local adaptation, thereby retaining the ecological integrity of admixed populations. Next, I explore the history of Yellowstone cutthroat trout populations and their relationships with other cutthroat trout lineages. In this investigation I uncover a cryptic isolated lineage of Yellowstone cutthroat trout and decipher the admixed history of two important Yellowstone cutthroat trout hatchery lines. Lastly, I attempt to understand potential environmental drivers of population genetic diversity and find results that corroborate riverscape theory. Altogether this body of work represents a tremendous step forward in understanding Yellowstone cutthroat trout and is an example of the utility genomic data can have for both conservation and evolutionary biology.

**LOOKING BACK TO STRIVE FORWARD: GENOMICS-INFORMED
HISTORY AND CONSERVATION OF YELLOWSTONE CUTTHROAT
TROUT**

By
William C. Rosenthal

A dissertation submitted to the Program in Ecology and Evolution.
and the University of Wyoming
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in
Ecology and Evolutionary Biology
Laramie, WY
August, 2026

ACKNOWLEDGMENTS

I am extremely grateful for the help of many people and organizations that made completing this work possible. More specifically, I appreciate the support from the Wyoming Game & Fish Department, Idaho Department of Fish & Game, Montana Fish, Wildlife & Parks, Greater Yellowstone Coordinating Committee, Western Native Trout Initiative, U.S. Fish and Wildlife Service, National Parks Service, Jackson Hole Trout Unlimited, Jackson Hole One Fly, and Wyoming NASA Space Grant. Specifically, Jason Burckhardt, Paul Gerrity, and Mike Mazur were especially generous with their time and assistance in helping me understand the wealth of existing knowledge of Yellowstone cutthroat trout and their range. I owe Katie Wagner a huge debt of gratitude, not only for her mentorship but also for her trust in me to complete such an ambitious project. Assistance from Dr. Sean Harrington and Dr. Josh Jahner was also crucial to developing the analyses presented here.

I also have to thank my friends and family that helped me persevere through the difficulties of higher education. My parents never wavered in their support for my non-traditional career path, and I cannot thank them enough for that. I will never forget the many nights spent playing cards with friends in Laramie: my appreciation for the warmth of that friendship across so many cold Wyoming winters cannot be understated. Maria Rodriguez also deserves mention because she never stops being an incredibly thoughtful and inspirational person.

Dr. Claudia Richbourg deserves the biggest thanks of all. Beyond a general brightening of my life, she helped tremendously in keeping me centered, focused, and motivated despite the many difficulties thrown at the both of us. I would not be where or who I am today without her support.

DEDICATION

To those working to create a better world for all

Contents

1	Introduction	1
2	Natural selection resists genetic swamping in admixed trout populations	3
2.1	Introduction	4
2.2	Materials and Methods	8
2.3	Results	22
2.4	Discussion	27
2.5	Figures	34
2.6	Tables	41
2.7	Appendix	81
2.8	Additional figures	83
3	History and inter-basin dispersal of Yellowstone cutthroat trout	115
3.1	Introduction	116
3.2	Materials and Methods	118
3.3	Results	124
3.4	Discussion	126
3.5	Figures	132
4	Roles of history and environment on population genetic diversity in Yellowstone cutthroat trout	140

4.1	Introduction	141
4.2	Materials and Methods	144
4.3	Results	148
4.4	Discussion	149
4.5	Figures	153
4.6	Tables	154
5	Management recommendations and data availability	161

List of Figures

2.1	Stocking records and broodstock use timeline	35
2.2	Conceptual figure for identifying introduced ancestry	36
2.3	Ancestry results	37
2.4	Homogenization results	38
2.5	Introduced ancestry and genetic diversity	39
2.6	Admixture and putatively adaptive loci	40
2.7	Interspecific admixture results	84
2.8	Interspecific admixture map	85
2.9	Admixture program results	86
2.10	Admixture and entropy comparison	87
2.10	Uncollapsed phylogenetic tree	89
2.11	Entropy results for Yellowstone Headwaters	90
2.12	Entropy results for Upper Yellowstone	91
2.13	Entropy results for Shields	91
2.14	Entropy results for Stillwater	91
2.15	Entropy results for Clarks Fork Yellowstone	92
2.16	Entropy results for Upper Wind alpine lakes	93
2.17	Entropy results for Upper Wind	94
2.18	Entropy results for Little Wind	95
2.19	Entropy results for Popo Agie	96

2.20	Entropy results for Upper Bighorn	96
2.21	Entropy results for Nowood	97
2.22	Entropy results for Greybull	98
2.23	Entropy results for Bighorn Lake	99
2.24	Entropy results for North Fork Shoshone	99
2.25	Entropy results for South Fork Shoshone	100
2.26	Entropy results for Little Bighorn	100
2.27	Entropy results for Upper Tongue	100
2.28	Entropy results for New Fork	101
2.29	Entropy results for Snake Headwaters	102
2.30	Entropy results for Gros Ventre	103
2.31	Entropy results for Greys-Hoback	104
2.32	Entropy results for Palisades	105
2.33	Entropy results for Salt	105
2.34	Entropy results for Upper Henrys	106
2.35	Entropy results for Teton	106
2.36	Entropy results for Willow	106
2.37	Entropy results for American Falls	107
2.38	Entropy results for Blackfoot	107
2.39	Entropy results for Portneuf	107
2.40	Entropy results for Raft	108
2.41	Entropy results for Goose	108
2.42	Entropy results for Upper Snake-Rock	109
2.43	Entropy results for Beaver-Camas	109
2.44	snmf cross-entropy	110
2.45	Watterson's theta and prop. introduced posterior predictive check	111
2.46	Tajima's pi and prop. introduced posterior predictive check	112

2.47	Watterson's theta and prop. Snake posterior predictive check	113
2.48	Tajima's pi and prop. Snake posterior predictive check	114
3.1	Sample locations map	133
3.2	Snake and Yellowstone Headwaters sample map	134
3.3	Cutthroat trout phylogeny	135
3.4	<i>O. virginalis</i> phylogeny	136
3.5	Two-Ocean Pass divergence estimates	137
3.6	Henrys Lake and S. Paint Rock Creek admixture	138
3.7	Spotting phenotype evolution phylogeny	139
4.1	Historic summer air temperature map	154
4.2	Stream branching probability map	155
4.3	Genetic diversity distributions	156
4.4	Trait evolution trees	157
4.5	Regional genetic diversity t-tests	158

List of Tables

2.1	Table of sample locations	41
2.2	Env. associated loci information	59
2.3	Gene ontology information	62
2.4	Allele freq. shift results	71
2.5	Genetic diversities	72
4.1	Phylogenetic GLMM results	159
4.2	PGLS model results	160

Chapter 1

Introduction

Genomic data has been appreciated for its utility to understanding evolutionary biology for almost as long as we have been able to generate it (Ommen [2005](#)). As the quality and density of genomic data produced has skyrocketed over the past few decades, so has our ability to understand the effects of events from more recent timescales (Supple and Shapiro [2018](#)). Improved precision paired with new resources for non-model organisms has paved the way for conservation genomics to flourish (Ouborg et al. [2010](#)), and now genomic data can provide data critical to improving species conservation outlooks (Hohenlohe, Funk, and Rajora [2021](#)).

Yellowstone cutthroat trout (*Oncorhynchus virginalis bouvieri*) are an iconic freshwater fish of the American West, native to the Snake and Yellowstone river basins across Wyoming, Idaho, Montana, Utah, and Nevada. Despite their large native range, however, the subspecies is under considerable threat and now occupies less than half of this area (Endicott et al. [2016](#)). Chief among these threats is the introduction of other trout species: brook trout (*Salvelinus fontinalis*) and brown trout (*Salmo trutta*) can outcompete Yellowstone cutthroat trout, while rainbow trout (*O. mykiss*) can hybridize with cutthroat trout and convert an entire population into hybrids over the course of only a few generations (Mandeville et al. [2019](#)). Such a drastic reduction in range size has motivated population managers to better understand what threats face remaining populations of

Yellowstone cutthroat trout.

Historically, a tremendous number of Yellowstone cutthroat trout were raised in fish hatcheries and distributed across the continent: one hatchery on the shores of Yellowstone Lake exported over 50 million eggs each year during its operation (Franke [1997](#)). In addition to the Yellowstone Lake hatchery, several other hatchery lines were maintained and stocked throughout the subspecies' range. All of this stocking may have substantially altered the natural patterns of biodiversity for Yellowstone cutthroat trout, which could impact the conservation status of otherwise intact populations. Thankfully, genomic data provides the necessary resolution to quantify the effects of historic stocking for Yellowstone cutthroat trout populations.

Genomic data can also provide important information for understanding the history of Yellowstone cutthroat trout and related lineages. The geologic history of the Snake River basin is especially complex, which has prevented previous studies using less-dense molecular data from arriving at solid conclusions (Shiozawa, Evans, Houston, et al. [2018](#)). A better understanding of cutthroat trout history is important for accurately describing evolutionarily significant units and also provides important context for how hydrology is shaped by geologic processes. Lastly, genomic data from many populations across a diversity of habitats, like the set of population used in this body of work, can help uncover what role environmental characteristics may have in shaping evolutionary outcomes.

Chapter 2

Natural selection resists genetic swamping in admixed trout populations

Abstract

Translocating individuals has long been used by natural resource managers to re-establish extirpated populations and bolster existing ones. However, historic usage of this tool was done without the precision of modern efforts and may have affected contemporary population status. Disentangling the genetic legacy of this history stands to improve the efficacy of future translocation efforts. We used DNA sequence data from over 5,000 individual fish to understand 1) how much native diversity has been replaced 2) if putatively adaptive loci resist replacement, and 3) if patterns of genetic diversity have been altered in Yellowstone cutthroat trout (*Oncorhynchus virginalis bouvieri*). Our analyses uncovered genetic homogenization of Yellowstone cutthroat trout populations, but not total loss of local ancestry. We also detected a lineage-specific association between introduced ancestry and genetic diversity; introduced ancestry from one source increased population genetic diversity, which drove a pattern of increased diversity associated with

total introduced ancestry. Genetic diversity of the stocking source is therefore a key consideration if increasing genetic diversity of the recipient population is the goal. Genotype-environment associations identified loci associated with average summer temperatures. Examining these loci in admixed populations in high temperature areas showed greater departure from introgression source allele frequencies than other loci, suggesting that they resisted swamping by introgression. Retention of putatively adaptive variation in the wake of introgression indicates that gene flow does not always erode local adaptation, and that the two could work synergistically to promote adaptation in some circumstances.

2.1 Introduction

Movement of individuals from one location to another is a commonly used technique in modern conservation and natural resource management to boost population abundance in specific locations, facilitate range shifts, and restore extirpated populations. The effectiveness of translocations has been demonstrated in many cases (Griffith et al. 1989; B. J. Novak, Phelan, and Weber 2021; Morris et al. 2021), and it may play a crucial role as climate change and human activities continue to alter habitats for species around the globe (Thomas 2011; Butt et al. 2021). Translocations can improve the genetic status of populations as well by supplementing genetic diversity and facilitating genetic rescue (Whiteley et al. 2015; Fitzpatrick et al. 2020; Clarke, A. C. Smith, and Cullingham 2024). The utility of translocations is not without caveats, though, as understanding factors affecting translocation success, and its potential pitfalls, are crucial to effectively use limited resources and minimize harm to ecosystems and populations (Schwartz and T. G. Martin 2013). In particular, the genetic consequences of human-mediated movement of individuals into wild populations has the potential to homogenize genetic diversity and erode locally adapted ancestry (Olden et al. 2004; Weeks et al. 2011). Understanding how

translocation practices affect genetic variation within and between wild populations requires large-scale and in-depth surveys (Joshua P. Jahner et al. [2019](#); Zimmerman et al. [2019](#); Erdman et al. [2022](#)) so that spatial variability in outcomes can be adequately and accurately captured.

Fisheries management has extensively utilized translocations from wild and captive sources for over a century: the intentional introduction of fish to a waterbody, often referred to as stocking, is one of the most common practices in fisheries management worldwide. Motivations for stocking have varied through time; historically, stocking was nearly exclusively performed with the goal of supplementing wild fish stocks to bolster recreational or consumptive value. Often, this involved the introduction of non-native species, with impacts detectable at scales ranging from local communities (Claussen and Philipp [2023](#)) to continental patterns of biodiversity (Rahel [2000](#)). In modern times, stocking is also widely used in conservation initiatives to bolster and/or restore wild populations of native species, and has lead to positive outcomes, ranging from local reintroduction of extirpated species (Pham, West, and Closs [2013](#); Hayes and Banish [2017](#); Riaz et al. [2020](#); Buckley et al. [2024](#)) to increasing populations of focal species over a large area (Johnson, He, and Fielder [2015](#)).

However, use of this management tool is not without its pitfalls. A review by (McMillan et al. [2023](#)) found that over 80% of studies examining the effects of captive-bred individuals on conspecific wild salmonid populations identified some degree of negative effect on the wild fish with genetic effects as the most commonly described negative impact. The theoretical basis for these effects has been long discussed around three central processes: 1) hatchery domestication (Harada, Yokota, and Iizuka [1998](#)), wherein selection acting on hatchery populations decreases their fitness and genetic diversity, 2) dilution of locally adapted ancestry (Carvalho et al. [1991](#)) , i.e., the replacement of potentially adaptive native alleles with introduced alleles, and 3) erosion of population differentiation (Satake and Araki [2012](#)), a process where alleles from introduction sources become

widespread and replace alleles that would otherwise differentiate populations. Considered together, these three processes can alter the fitness, genetic diversity, and genetic differentiation of wild populations.

More often discussed than negative genetic effects of conservation-motivated translocations is the potential to increase fitness through genetic rescue (Whiteley et al. 2015). Genetic rescue aims to increase the genetic diversity of populations facing inbreeding depression, and is often implemented via intentional translocations. Genetic diversity has relatively recently become an important consideration for conservation efforts, and genetic rescue is a key tool for increasing population genetic diversity when other interventions are not possible (Richard Frankham 2022). Increasing genetic diversity is done to increase fitness, prevent future inbreeding depression, and/or prevent the loss of adaptive potential (Kardos et al. 2021). Fitness decreases due to introgression from translocated individuals can also occur (Gilk et al. 2004; Huff et al. 2011), though a review of intentional translocation studies found that increases in fitness were more common (Clarke, A. C. Smith, and Cullingham 2024). However, in most cases, we simply do not know if translocated individuals successfully interbreed with the resident population, nor do we understand how any interbreeding that does occur affects the genetic diversity or ancestry of recipient populations. Studies to assess these factors quantitatively at broad scales are needed to understand human-mediated effects on the genetic status of populations of conservation concern.

The Rocky Mountain Cutthroat Trout (*Oncorhynchus virginalis*) is a geographically widespread freshwater fish native to western North America. Like many other trout, it is a recreationally and culturally important fish that has been the target of many management actions, including translocations (Gresswell 2011; J. L. Metcalf et al. 2012). Morphological differences between Rocky Mountain Cutthroat Trout lineages, of which there are nine currently recognized, are mostly subtle despite molecular evidence for substantial differentiation between them (Trotter et al. 2018). Historic translocation records for these

fish reflect the neglect of this diversity, as records in the late 19th and earlier 20th centuries were often labeled “blackspotted trout” or “cutthroat trout” without any record of location of origin (Nordberg et al. 2021). Despite an incomplete record, genetic data have proved crucial to disentangling the puzzle of historic fish movement, including identifying critical misidentifications (J. L. Metcalf et al. 2012). In the case of Colorado River Cutthroat Trout (*O. v. pleuriticus*), one subspecies of Rocky Mountain Cutthroat Trout, human-mediated translocations of fish have extensively reshaped the genetic diversity of the subspecies across its range (J. L. Metcalf et al. 2012; S. M. Love Stowell et al. 2015). However, for other Rocky Mountain Cutthroat Trout subspecies, the genetic effects of translocations are only sparsely understood (Eaton, Janet L. Loxterman, and Ernest R. Keeley 2018).

Yellowstone Cutthroat Trout (*O. v. bouvieri*), a lineage of Rocky Mountain Cutthroat Trout native to the Upper Snake and Yellowstone River basins, were extensively stocked in North America during the early-mid 20th century; the fish hatchery operating in Yellowstone National Park produced over 50 million Yellowstone Cutthroat Trout eggs every year and had exported over 800 million fry and eggs by the time it closed in 1953 (Franke 1997; Trotter 2008). Translocations of Yellowstone Cutthroat Trout from state-managed hatcheries continues to this day, with over 1.4 million individuals introduced within the state of Wyoming alone in 2024 (Wyoming Game & Fish Department 2024). Given the tendency for cutthroat trout to readily differentiate over geographic space (R. J. Behnke 1992; Janet L Loxterman and Ernest R Keeley 2012; J. L. Metcalf et al. 2012; Shiozawa, Evans, Houston, et al. 2018), the question stands whether these translocations have affected the differentiation, diversity, and local adaptation of Yellowstone Cutthroat Trout populations.

Genomic data provides newfound power to distinguish the genetic impacts of translocations, even within the same subspecies (Joshua P. Jahner et al. 2019). Here we use a genomic dataset of over 5,000 samples from 231 wild Yellowstone Cutthroat Trout populations across their range in conjunction with 12 collections of hatchery lines and

other cutthroat trout species and lineages to document 1) the extent to which native ancestry of Yellowstone Cutthroat Trout has been replaced by over a century of intensive translocations, 2) if any replacement has affected adaptive and neutral diversity equally, and 3) whether or not this replacement has altered population genetic diversity. We use a combination of genetic clustering and phylogenetic approaches to precisely identify and quantify admixture from non-native sources, genotype-environment associations to identify putatively adaptive loci, and generalized linear models to assess changes in genetic diversity due to introgression.

2.2 Materials and Methods

System background and sampling

Yellowstone Cutthroat Trout are native to the Snake and Yellowstone River basins in western North America, spanning the states of Montana, Nevada, Idaho, Utah, and Wyoming. The range of the subspecies includes elevations from approximately 1,000 meters above sea level to over 2,400 meters, spans climates from subarctic to cold desert, and includes waterways ranging in size from headwater alpine streams to large rivers. Some Yellowstone Cutthroat Trout within the Snake River Headwaters near Jackson, Wyoming have distinctly smaller and more numerous spots than other Yellowstone Cutthroat Trout, known colloquially as the “finespotted” form (R. J. Behnke 1992). This distinction led to their description as a separate subspecies, *Oncorhynchus clarkii behnkei*, recently revised to *Oncorhynchus virginalis behnkei* (Markle 2018). This designation is controversial as molecular evidence has not supported the distinction of the subspecies (M. A. Novak, Kershner, and Mock 2005), so we include finespotted Yellowstone Cutthroat Trout from the Snake River Headwaters as *O. v. bouvieri* here for simplicity. We collected samples – either from wild fish or archived DNA from previously sampled fish – from 231 different wild populations across this area (Table 2.1). Sampling was primarily conducted via

backpack electrofishing, though other methods such as rod and reel angling, raft electrofishing, gillnets, and weirs were also used. Samples acquired specifically for this project were collected under University of Wyoming IACUC permit number 20220810CW00569-01 and Yellowstone National Park sampling permit number 8264. Archived DNA was obtained from the Montana Parks & Wildlife Conservation Genetics Lab at the University of Montana, the Idaho Department of Fish & Game’s Eagle Genetics Lab, and from Pisces Molecular, LLC (Boulder, Colorado, USA). Populations previously identified as having minimal hybridization with introduced Rainbow Trout (i.e., Conservation Populations; Endicott et al. (2016)) or those with unknown genetic status were prioritized. Each population was represented by an average of 23 samples, with a minimum of 1 and maximum of 146, and the name used to represent each population was that submitted by the individual who provided the samples.

To identify any hybridization with other trout species, we included samples from Lahontan Cutthroat Trout (*O. henshawi*; from United States Fish & Wildlife Service’s Lahontan National Fish Hatchery), Westslope Cutthroat Trout (*O. lewisi*; from Montana Department of Fish, Wildlife & Parks’ Washoe Hatchery), Rainbow Trout (*O. mykiss*; from Eagle Lake and Arlee strains), and Golden Trout (*O. aguabonita*; from Wyoming Game & Fish Department’s Daniel hatchery). To assess intraspecific introgression, we included samples from three known stocking sources from lineages within Yellowstone Cutthroat Trout: the LeHardy Rapids strain (Yellowstone River drainage, Wyoming, USA; Figure 2.1), the Henrys Lake strain (Snake River drainage, Idaho, USA), and the Bar BC strain (Snake River drainage, Wyoming, USA). The LeHardy Rapids strain was sourced from Wyoming Game & Fish Department’s Tensleep Hatchery, the Henrys Lake strain was sourced from the Idaho Department of Fish & Game’s Henrys Lake Egg Collection Facility, and the Bar BC strain was sourced from the Wyoming Game & Fish Department’s Auburn Hatchery and the United States Fish & Wildlife Service’s Jackson National Fish Hatchery. The Bar BC strain fish possess the finespotted phenotype of Snake River Yellowstone

Cutthroat Trout. Samples from three historical hatchery lines of Yellowstone Cutthroat Trout could not be obtained: Yellowstone Lake, South Paintrock Creek, and McBride Lake. We use the LeHardy Rapids strain as a representative of both contemporary Yellowstone Cutthroat Trout stocking and of historic stocking from the Yellowstone Lake hatcheries; we were unable to source historic samples of Yellowstone Lake origin, and preliminary analyses showed that LeHardy Rapids strain fish were genetically similar, though not identical to, contemporary samples from Yellowstone Lake.

To prevent misidentifying ancestry from *O. virginalis* outside of Yellowstone Cutthroat Trout, we also included samples representing other *O. virginalis* lineages: Bear River Cutthroat Trout (*O. v. utah*; BRC; from Wyoming Game & Fish Department’s Wigwam hatchery), Colorado River Cutthroat Trout (CRC; from Wyoming Game & Fish Department’s Daniel hatchery), and Bonneville Cutthroat Trout (*O. v. utah*; BON; from Utah Division of Wildlife Resources’ Glenwood hatchery, originating from the southern Bonneville basin). Each of these hatchery source populations was represented by an average of 33 individuals (minimum 5, maximum 52).

We compiled records of cutthroat trout stocking within Idaho (<https://idfg.idaho.gov/fish/stocking>; accessed 29 May 2025), Montana (<https://myfwp.mt.gov/fishMT/plants/plantreport>; accessed 29 May 2025), and Wyoming (<https://wgfapps.wyo.gov/FishStock/FishStock>; accessed 29 May 2025) using each state’s online stocking records database (Figure 2.1). We only included records of Yellowstone Cutthroat Trout, Bear River Cutthroat Trout, Snake River Cutthroat Trout, or unspecified cutthroat trout. We also tabulated the number of stocking events and total number of fish stocked for each population present in the records that, after completing our genetic analyses, was found to have no introduced ancestry.

DNA extraction and sequencing

We extracted DNA from tissue samples using Qiagen DNeasy blood and tissue extraction kits. We prepared twenty-two reduced representation genomic libraries for high throughput DNA sequencing at the University of Wyoming Genome Technologies Laboratory using the genotyping-by-sequencing protocol in (Thomas L. Parchman et al. 2012), with 288 or 384 individuals multiplexed per library. We digested DNA using restriction enzymes EcoRI and MseI before ligating unique 8-10 base pair nucleotide barcodes to the resulting fragments. We then multiplexed and amplified samples in two PCR replicates. We size-selected each library to retain only fragments 325-400 base pairs in length via BluePippin (Sage Science). We sequenced each library multiplexed to 288 individuals on one Illumina NextSeq200 P3 flowcell to produce approximately 1.2 billion, 100 base pair, single-end reads. We sequenced each library multiplexed to 384 individuals on one lane of an Illumina NovaSeq 1.5B flowcell to produce approximately 1.5 billion, 150 base pair, single-end reads. Sequencing on the Illumina NextSeq platform was performed at the University of Wyoming Integrated Microscopy Core (Laramie, Wyoming, USA) and sequencing on the Illumina NovaSeq platform was performed at Admera Health (South Plainfield, New Jersey, USA).

Sequence assembly and variant calling

We assigned reads to each individual sample by parsing barcodes and then aligned each sample to the Rainbow Trout reference genome (Gao et al. 2021) (accession GCA_013265735.3) using the bwa-mem algorithm (Li 2013). For libraries with paired-end reads, we only retained the forward read as our library preparation protocol does not attach reverse adapters. To ameliorate library effects from the inconsistencies inherent in size selection across so many libraries and from different sequencing effort and read lengths, we identified variable loci on each library separately using `bcftools` (Danecek, Bonfield, et al. 2021), removing indels. Each resultant vcf was then filtered to keep only biallelic sites (SNPs), sites with at least 3 reads per individual, and sites with missing data less than

65%. We then used `vcftools` (Danecek, Auton, et al. 2011) to identify which sites were found across all filtered library vcfs and create a new vcf containing all individuals and only those sites. We filtered this vcf to retain only sites with a minor allele count of 3 or greater, sites with maximum 10% missing data, sites with quality scores greater than 19, sites with genotype quality scores greater than 9, average depth less than 50, and individuals with less than 75% missing data. To account for potential PCR error duplication, we then removed heterozygous genotype calls with significantly unlikely allelic imbalance using the `allelicBalance.py` script (from <https://github.com/joanam/scripts/blob/master/allelicBalance.py>) with a p-value threshold of 0.01. All sequence processing was performed on the University of Wyoming’s Medicine Bow Computing Cluster.

Quantifying hatchery introgression

We quantified hatchery introgression in a multi-stage approach using **entropy** (Gompert and Buerkle 2013; Shastry et al. 2021), which precisely assigns proportions of an individual’s ancestry to one or more genetic clusters. The number of clusters used is set *a priori* and typically denoted by K . Here we use **entropy** to identify native and introduced ancestry in a manner conceptually (though not mathematically) similar to genetic stock identification approaches. Genetic stock identification programs can be powerful and useful, but do not provide precise estimates of admixture between lineages and require prior knowledge of the diversity present (Moran and Anderson 2019). We use samples with known ancestry, primarily from hatchery sources, as an analog for the reference samples used in GSI approaches with the goal of identifying genetic clusters that correspond to different lineages within our hatchery source populations (and any native lineages present in wild samples). The proportional assignment of individual ancestries to different clusters then allows us to quantify the contribution of hatchery sources to wild population ancestry, even if small in magnitude.

Because introgression from trout lineages outside of *O. virginalis* into Yellowstone Cutthroat Trout would be more easily detected than intraspecific introgression, we first used **entropy** to quantify introgression with Rainbow Trout, Golden Trout, Lahontan Cutthroat Trout, and Westslope Cutthroat Trout. We did this by constructing a vcf file for each population plus “reference” individuals from each of the aforementioned trout species, converting this into the **entropy** input format (mpg1), and running the program on each file separately at K=5 clusters (four reference populations plus the focal population with no prior information about population assignment). Populations with less than five individuals were pooled together to avoid misassignment due to imbalanced sample sizes. Some focal populations showed unexpected admixture with other species; we also ran **entropy** on these populations at K=6 to see if there was support for these populations containing ancestry from two lineages not included among the hatchery sources (e.g., *O. virginalis* and another, unrepresented species). Each **entropy** run had a total chain length of 20,000 steps, discarded the first 10,000 steps as burn-in, and recorded each tenth step in the post-burn-in chain. In cases where we ran both K=5 and K=6 for a population and both models converged, we selected the best model using the deviance information criterion (DIC) and considered only the results for those models. To prevent cluster switching between replicate chains, we set the starting values for each individual according to their origin; we gave each reference population its own cluster (e.g., cluster 1 for Rainbow Trout, cluster 2 for Golden Trout) and set the starting value at 1 for the cluster corresponding to each sample’s origin and 0 for all other clusters. For the focal wild population in each file, we set the starting value equal to 0 for all reference clusters but modified the starting values for non-reference clusters depending on the number of clusters. In K=5 models we set the starting value for cluster 5 to 1, and in K=6 models we set the starting values for clusters 5 and 6 to 0.75 and 0.25, respectively. We ran five replicate chains for each population and each value of K. We determined model convergence by examining effective sample size and potential scale-reduction factor scores – if 95% of potential scale-reduction

factor estimates were under 1.3, we considered a model converged. To avoid assigning ancestry to species unlikely to be present, we then conducted a second run of **entropy** analyses on populations with any individuals inferred to be less than 98.5% Rocky Mountain cutthroat trout, but including only those references assigned above 2.5% of the population’s total ancestry in the first run. We then removed any individuals with total proportion of ancestry assigned to other species $> 1.5\%$ to create a filtered set of *O. virginalis* individuals. Preliminarily, we ran **admixture** (Alexander, Novembre, and Lange 2009) on this set of individuals with K=1 through K=14 clusters.

Intraspecific introgression was identified in a manner similar to that described above, though we pooled wild populations found within the same USGS 8-digit hydrologic unit code (HUC 8) area (U.S. Geological Survey 2018). Computational constraint necessitated that we first narrow down the number of hatchery source populations analyzed for each HUC 8. To do this, we ran a modified version of entropy that allows the allele frequencies of the ancestral populations modeled by the program to be set a priori. We used plink (Purcell et al. 2007; Chang et al. 2015) to calculate allele frequencies for all hatchery reference samples from each of the five major river basins (corresponding to five lineages) within *O. virginalis*: the Yellowstone River, Snake River, Bear River, Bonneville Basin, and Colorado River. We excluded the Henrys Lake strain Yellowstone Cutthroat Trout reference, as our models did not consistently identify it either identical to or a separate cluster from our LeHardy Rapids strain reference. We then split our samples into files containing only the samples from one population and ran entropy on each file with K=5 clusters and with the pre-set allele frequencies for each cluster specified to one of the previously mentioned lineages. We again ran five replicate chains for each file; each chain was 85,000 steps long and had the first 1,000 steps discarded to burn-in and recorded each tenth step in the post-burn-in chain. We assessed convergence in the same manner as described above. We recorded which clusters represented $\geq 15\%$ of a population’s ancestry, as these would be included as reference populations in the next step. One exception to this

is the Portneuf HUC 8 (17040208), where historic admixture from Bonneville basin ancestry had been detected (M. Campbell et al. 2018) and so we included the Bonneville reference samples for this HUC 8 in addition to the other reference groups identified. Another exception is the inclusion of South Paint Rock Creek as a reference for HUC8s within the Wind-Bighorn River basin with preliminary evidence for Bear River ancestry. South Paint Rock Creek was historically used as a broodstock source within Wyoming and preliminary analyses showed that it likely contained ancestry from the Bear River basin.

We also constructed another set of files, one for each HUC 8, that included all of the same samples except for reference populations from the same region (e.g., no Yellowstone River reference for any HUC 8 within the Missouri River region), as attempting to split these closely related groups might cause issues for model convergence. Two HUC 8s, the Upper Yellowstone and Snake River Headwaters, contained too many individuals to be included in one file; we split these HUC 8s by population into four and two pieces, respectively, before continuing. We then ran entropy on each HUC 8 file at $K=N$ and $K=N+1$ clusters, where N is the number of reference populations included in that file. We tested different values of K in order to discern whether or not there was support for a cluster of local or otherwise unrepresented ancestry; some HUC 8s showed patterns consistent with a potential second local or unrepresented cluster, and so we also ran these at $K=N+2$ clusters. We again ran five replicate chains for each model and ran each chain for 70,000 MCMC steps with the first 50,000 discarded as burn-in and recording each tenth step in the post-burn-in chain. Burn-in length was set based on computational constraints and the methods of (Joshua P Jahner et al. 2026); this prior work also used entropy to parse evolutionary relationships below the level of subspecies. We assessed model convergence in the same manner as described above. We expected some models to not converge at $K=N$, given the documented difficulties of attaining convergence in ancestry estimation models as model complexity increases (Rosenberg et al. 2002; Corander et al. 2008; Skotte, Korneliussen, and Albrechtsen 2013), and so we continued only with models

that converged. If multiple models converged for the same dataset, we selected the best model using DIC and continued with the results from that model. If models both with and without reference populations from the same region (HUC 2) as the focal HUC 8 converged, we continued only with the results from models including the reference populations from the same region (HUC 2). We also conducted **admixture** analyses on the same individuals with the same number of clusters to compare these results to those from **entropy**.

We identified hatchery introgression from model results on a genetic cluster-by-cluster basis for each HUC 8 (Figure 2.2 step 2). We denoted a genetic cluster as “introduced” if it a) was assigned nearly entirely ($\geq 95\%$) to one or more hatchery source populations and b) any hatchery source populations assigned were not derived from the focal HUC 8 (e.g. hatchery strains derived from the same HUC 8 as the focal populations). In the case of the Auburn Hatchery and Jackson National Fish Hatchery lines, because their source population is only 600 meters from the intersection with two other HUC 8s, we extended the local area for these lines to include all three of these HUC 8s. Criterion a) was necessary to prevent assigning ancestry of ambiguous status as introduced, and was set at 95% to be conservative while also allowing for variation due to allele frequency changes post-introduction or MCMC stochasticity. Criterion b) was necessary to avoid miscategorizing shared ancestry due to relatedness as introduced, introduced ancestry. We also recorded the inferred basin of origin for clusters of introduced ancestry, and included uncertainty in this inference by documenting all hatchery source populations assigned $\geq 65\%$ to the cluster in question. The 65% cutoff was chosen because it was anti-conservative but still required the majority of ancestry to be assigned identically.

Overall, because we require positive evidence for the presence of introduced ancestry, these likely represent conservative criteria for inferring introduced ancestry at a fine geographic scale. Detecting introduced ancestry from closely-related introgression sources is still difficult, however, and our methods may still fail to detect this. After we designated clusters as “introduced” and assigned them to a basin of origin using these criteria, we

designated all other clusters as “native” and summarized across populations and HUC8s by taking the average native and introduced ancestry across individuals. Uncertainty in these estimates was conservatively propagated from cluster proportion estimates by adding together the 95% credible interval widths across all clusters within the same category (e.g., native ancestry, introduced ancestry).

Phylogenetic tree estimation

To identify instances of admixture from sources not represented in our **entropy** analyses and begin to understand the relationships between lineages within Yellowstone cutthroat trout, we estimated a phylogenetic tree using **CASTER-site** (Zhang, Nielsen, and Mirarab 2025) and **IQ-TREE** (Minh et al. 2020). We subsetting our samples to include a maximum of 5 individuals each from every population with less than 5% inferred introduced ancestry, individuals from cutthroat trout species (all *O. virginalis* lineages, Westslope cutthroat trout, and Lahontan cutthroat trout) hatchery source populations used in our **entropy** analyses, samples from the Henrys Lake Egg Collection Facility (Henrys Lake, Idaho, USA), and one Arlee strain Rainbow Trout sample to be used as an outgroup (Figure 2.2 step 3). We also included any populations that **entropy** inference identified as entirely introduced but would not converge at $K > 1$ and populations from basins with $\geq 95\%$ of their total ancestry assigned to the same hatchery source; in both of these situations, we wanted to ensure that unrepresented hatchery sources and/or more complex introduction histories that might cause difficulties for **entropy** model convergence would not cause false negative results for identifying introduced ancestry.

We used a similar set of SNPs for phylogenetic estimation, but increased the allowable missing data proportion per site from 0.1 to 0.2 and removed any sites that were invariant within this smaller set of individuals. We used **CASTER-site** to estimate the best-fit tree topology with 100 bootstrap replicates, and then used ModelFinder Plus (Kalyaanamoorthy et al. 2017) to identify the best fitting substitution model to estimate

branch lengths for that topology. The results from this analysis were used to correct misidentifications of native ancestry in the **entropy** results (Figure 2.2 step 4). If any wild populations deviated substantially from their expected placement within the tree based on geography, particularly if their position was estimated close to any hatchery source population despite them not being geographically proximate to these populations, we designated their ancestry to introduced (Figure 2.2). Alternatively, if their phylogenetic position reflected geography and hydrologic connectivity, we designated their ancestry as native. If placement for individuals within a population varied enough to constitute potentially different ancestry designations (i.e., a mix of native and introduced ancestry), we assigned ancestry for all individuals in those populations individually. To aid visualization, we generated ultrametric trees from the **CASTER-site** topology and IQtree branch lengths using **chronos** in the R package **ape** (Paradis and Schliep 2019) with the “clock” method and $\lambda = 1.0$.

Homogenization

We assessed homogenization by comparing the average pairwise F_{ST} between populations with and without introduced ancestry. We first removed individuals with $< 98.5\%$ of their ancestry assigned to *O. virginalis* before removing all individuals from populations with less than five individuals retained after filtering, populations from historically fishless alpine lakes, and populations with $> 5\%$ of their average ancestry assigned to *O. virginalis* lineages from outside the range of Yellowstone cutthroat trout (e.g., Bear, Colorado, and Bonneville watersheds). We estimated F_{ST} using the Reich-Patterson F_{ST} estimator (Reich et al. 2009) and excluded comparisons between populations within the same basin (HUC8). Populations were categorized as “unadmixed” if their total proportion introduced ancestry was $\leq 10\%$ or “admixed” if their total proportion introduced ancestry was $> 10\%$. We then tested for statistically significant differences in average pairwise F_{ST} between groups using a permutation test, where we randomly shuffled the group designation (“admixed” or

“unadmixed”) for each F_{ST} value and then calculated the difference in F_{ST} estimates between groups. We repeated this procedure 10,000 times and compared the distribution of permuted F_{ST} differences to our observed F_{ST} difference.

Population genetic diversity

For populations with greater than five individuals retained after filtering, we used **ANGSD** (Korneliussen, Albrechtsen, and Nielsen 2014) to calculate three metrics of genetic diversity: Watterson’s estimator of θ (θ_W), Tajima’s π , and observed heterozygosity. We calculated θ_W and Tajima’s π at the population level and calculated observed heterozygosity at the individual level before taking the mean across individuals within each population. We estimated multiple metrics of population genetic diversity to assess whether any observed patterns would hold across different facets of genetic diversity. We also wanted to examine the effects of introduced ancestry from specific lineages on genetic diversity, and so we tabulated the amount of introduced Snake River and introduced Yellowstone River ancestry estimated within each population as alternative metrics of introduced ancestry.

We examined the relationship between genetic diversity and all measures of introduced ancestry using generalized linear mixed models in a Bayesian framework via the R package **brms** (Bürkner 2017; Bürkner 2018). We constructed models with the same covariates for each of the population genetic diversity estimates and for each measure of introgression, each with a random intercept for the basin origin of each population. We used gamma regression for each model, as Shapiro-Wilk tests confirmed that our population genetic diversity estimates were not normally distributed across populations. Each model was run with three replicate chains, a total chain length of 3000 and the first 1000 iterations discarded as burn-in. Prior probability distributions were normal distributions with means = 0 and standard deviations = 10. Model convergence was assessed by Gelman-Rubin statistics below 1.01, and model fit was assessed via posterior predictive checks. We inferred

statistical significance of a parameter if its 95% credible interval did not overlap zero.

Genotype-environment associations

To identify loci potentially involved with local adaption to environmental conditions, we searched for genotype-environment associations using latent factor mixed models as implemented in the R package `lfmm` (Caye et al. 2019; Frichot, Schoville, et al. 2013). We included all individuals from all wild populations except those with ancestry from other species. We compiled average summer air temperature at the HUC 10 level for each population. Average summer air temperature was calculated from the monthly 30 year averaged air temperature data from PRISM (PRISM Group, Oregon State University 2023) for the months of June, July, and August. We identified the number of latent factors to include by examining the cross-entropy values for runs of `snmf` (Frichot and François 2015) initialized with $K = 1$ through $K = 20$ populations: we selected the value after which the slope of cross-entropy drastically decreased. Once the number of latent factors was selected, we imputed the remaining missing data within our genotype matrix using the `impute` functionality of `snmf`, using results for that value of K . We fit models using the `lfmm_lasso` function and corrected the resultant calibrated p-values for multiple testing bias using the Bonferroni-Hochberg method. We considered a SNP significantly associated with an environmental variable if its corrected p-value for that variable was < 0.05 .

After identifying loci associated with summer temperatures, we searched for signs that natural selection was acting on those loci in the wake of introgression. To do this, we identified populations with $> 10\%$ and $< 95\%$ introduced ancestry occurring in environments in the top 30% of summer temperatures. We chose the degree of introgression such that there might be enough introgression to at least partially alter local allele frequencies but not enough introgression to replace the local ancestry entirely. We chose populations only within the top 30% of values for each environmental variable to increase the likelihood that natural selection would be acting on our environmentally

associated loci. To identify signs of natural selection at environmentally associated loci in the selected populations, we used an approach similar to that of (Fitzpatrick et al. 2020) and calculated the shifts (i.e., absolute value of the difference) in allele frequency between the inferred introduction source and the observed population allele frequencies. To account for admixture from multiple sources, we combined the allele frequencies of all stocking sources inferred within a population (at $\geq 1\%$ ancestry) with weights according to their relative contribution. For example, if ancestry from one source accounted for 50% of a population’s ancestry and another 25% came from another source, we would weight the frequencies of those sources at 66.6% and 33.3%, respectively. We then compared the allele frequency difference of the inferred source population and the wild population between the environmentally associated loci random non-associated loci. We did this by comparing the allele frequency difference of each environmentally associated locus to two non-associated loci occurring at a similar frequency ($\pm 5\%$) to the associated locus in the stocking source population. To account for potential non-independence resulting from close physical proximity of environmentally associated SNPs, we used genome-wide, sex-average recombination rates from (Brekke and Knutsen 2026) to thin these loci to only retain one locus per centimorgan of chromosome. For each population and for all populations combined, we tested for statistically significant differences in allele frequency shifts using t-tests and corrected the resultant p-values for multiple testing bias through the Benjamini-Hochberg correction (Benjamini and Hochberg 1995). To estimate the proportion of environmentally associated loci lost to admixture (i.e., swamped), we calculated the proportion of associated loci fixed in both the inferred introduction source and the wild population.

To further validate that loci identified by **1fmm** were influenced by selection, we also used **PCAdapt** (Luu, Bazin, and Blum 2017) to identify outlier loci without use of environmental data. We ran **PCAdapt** with $K = 13$ and again corrected p-values though the Benjamini-Hochberg correction. We considered any loci with corrected p-values < 0.05 as

outliers. To see if loci identified by `lfmm` were more likely to be identified by `PCAdapt`, we used Fisher’s exact tests with an alternative hypothesis of an odds ratio greater than one. We also aggregated gene ontology information for genes containing or starting within 100bp of environmentally associated loci according to the (Gao et al. 2021) genome annotation.

2.3 Results

DNA sequencing, sequence assembly, and variant calling

DNA sequencing generated an average of 2.41 million reads per individual; on average, 98% were successfully aligned to the reference genome. After removing 1,314 individuals with high proportions of missing data, 6,059 individuals remained. Before filtering, we retained 109,362 sites across all libraries. After filtering, we retained 32,282 sites.

Quantifying hatchery introgression and homogenization

The first round of genetic clustering models to identify introgression from other trout species identified 554 individuals across 71 populations with at least 1.5% of their ancestry outside of *O. virginalis* (Figure 2.7, Figure 2.8). Of these samples, 424 had ancestry from rainbow trout (*O. mykiss*), 142 had ancestry from golden trout (*O. aguabonita*), 72 had ancestry from Lahontan cutthroat trout (*O. henshawi*), and 27 had ancestry from Westslope cutthroat trout (*O. lewisi*). Ancestry coefficient estimates, which range from 0 to 1, were precise with 95% credible interval widths of 0.002 on average. In all but a few instances, introgression with other species was confined to alpine lake populations and populations with known hybridization (Endicott et al. 2016), primarily associated with Rainbow Trout.

Our preliminary examination of population genetic structure via `admixture` (Alexander, Novembre, and Lange 2009) found $K=6$ as the best supported (as inferred by cross-validation errors) number of clusters to represent all *O. virginalis* samples. At this

number of clusters, ancestry coefficients show genetic differentiation across watersheds and potential inter-basin gene flow (Figure 2.9). The second round of genetic clustering models to identify relevant hatchery source populations using the modified, fixed-reference version of **entropy** (Gompert and Buerkle 2013; Shastry et al. 2021) identified between zero and four hatchery sources at $\geq 15\%$ (mean 2.24) across all populations. In the third round of clustering, **entropy** models for specific hydrologic basins (USGS HUC8 level) identified a genetic cluster not associated with introduced ancestry in 25 of 31 basins. Results from **admixture** were overall similar to those from **entropy**. Average deviation from **entropy**-estimated ancestry coefficients in **admixture** models was 1.42×10^{-7} . One **admixture** model showed larger departures from **entropy** ancestry coefficient estimates due to differences in how the models identified clusters (Figure 2.10); we attribute this discrepancy to differences in model fitting, as **admixture** does not use fully Bayesian MCMC sampling nor does it incorporate uncertainty from genotype likelihoods.

The set of SNPs used for phylogenetic analysis contained 85,035 SNPs. We identified TVM+F+ASC+R10 as the best substitution model. The best-fitting tree showed overall strong support for many nodes, including basal nodes (Figure 2.3; see Figure 2.10 tree without collapsed clades). We identified several instances of phylogenetic relationships not reflecting geography, which implied that these populations originated from stocking sources. This was most frequently seen as populations within the Big Horn basin being phylogenetically placed within the Upper Yellowstone basin populations, which may reflect stocking of fish originating from McBride Lake or early twentieth century Yellowstone Lake brood sources (both unrepresented in our hatchery source populations). One Upper Yellowstone population, East Rosebud Creek (Stillwater River drainage), was found to be sister to populations from the Lamar River drainage (Upper Yellowstone basin), which we also determined was caused by stocking. Populations from the Henrys Fork (Upper Snake basin), including the Henrys Lake Egg Collection Facility, were shown to be sister to the Missouri River lineage; a pattern consistent with the Yellowstone River ancestry of the

Henrys Lake Egg Collection Facility hatchery source found in preliminary analyses. Therefore, wild populations within that clade were designated as introduced and having Yellowstone basin ancestry. Populations from the Palisades Reservoir (Snake Headwaters basin) and Willow (Upper Snake) HUC 8s were inferred by **entropy** to be introduced, but were found to be sister to fish from the Greys and Salt Rivers (Snake Headwaters basin) and not hatchery populations from the same basin. We therefore designated these populations as native in origin. Some individuals from Dry Creek, Goose Creek, Upper Tincup Creek, and the Salt River (all populations where all sampled individuals were included in tree estimation) placed near known stocking sources: these individuals were assigned introduced ancestry and the remainder from those populations assigned native ancestry. Samples from South Paint Rock Creek, a historic stocking source within the Bighorn River basin, showed aberrant phylogenetic placement likely indicative of unexpected admixture with a more distantly related (e.g., Bear River) source: see chapter 2 for further investigation of this result.

After using clustering and phylogenetic approaches, ancestry determined to be native in origin was found in 80% of basins, and ancestry determined to be introduced was found in 84% (Figure 2.3, Figure 2.11-2.43). The distribution of ancestry proportions was relatively consistent across individuals within a population, suggesting little active hybridization between lineages. Across all populations the mean proportion of introduced ancestry was 0.469. Introduced ancestry was most frequently associated with hatchery sources from the Yellowstone River watershed (Figure 2.4). Introduced ancestry was found in 65% of populations surveyed and 52% of populations within the native range of the subspecies (i.e., after excluding historically fishless habitats). Within almost every basin examined, we identified at least one individual fish with greater than 98% native ancestry. Results at the population scale were much more variable: across *O. virginalis* populations, only 47% contained individuals with > 98% native ancestry.

In our assessment of genetic homogenization, 129 populations were retained for

comparison: 93 were categorized as “unadmixed” 36 as “admixed”. Average pairwise F_{ST} among “unadmixed” populations was 0.152 and average pairwise F_{ST} among “admixed” populations was 0.113: the permutation test found that none of the 10,000 permuted datasets had as large of a difference between groups (Figure 2.4C).

After searching available stocking records for information on each population with no introduced ancestry, 23 populations had at least one stocking event. Of the populations without introduced ancestry and with a record of stocking, the median number of fish stocked was 198,594 and the median number of stocking events was 13. We do not have access to a comprehensive set of stocking records for the entire Yellowstone Cutthroat Trout range, and so these numbers represent a conservative minimum number of fish stocked into this subset of populations. Similarly, the incompleteness of our stocking records limits our ability to rigorously assess stocking efficacy across all sampled populations.

Population genetic diversity

Genetic diversity estimates ranged from $5.44 * 10^{-4}$ to $3.22 * 10^{-3}$ for Watterson’s theta (θ_w), $9.77 * 10^{-4}$ to $4.72 * 10^{-3}$ for Tajima’s π , and $2.80 * 10^{-4}$ to $1.69 * 10^{-3}$ for observed heterozygosity (Figure 2.5a-c, Table 2.5). Bayesian generalized linear mixed models to quantify the effects of admixture on genetic diversity metrics had effective sample sizes for each parameter of > 1000 . Posterior predictive checks confirmed that a gamma regression fit our data reasonably well (Figure 2.45-2.48). There was a significant positive effect of introduced ancestry on genetic diversity as measured by θ_W (mean estimate 0.20, 95% credible interval bounds 0.08 and 0.32; Figure 2.5), and Tajima’s π (mean estimate 0.13, 95% credible interval bounds of 0.03 and 0.23), but not for observed heterozygosity. When the effect of introduced Snake River ancestry on population genetic diversity was modeled, there was a significantly non-zero effect found in models using both θ_W (mean estimate 0.35, 95% credible interval bounds 0.19 and 0.53; Figure 2.5) and Tajima’s π (mean estimate 0.27, 95% credible interval bounds 0.13 and 0.43). The parameter estimates in

these models were higher than the parameter estimate for the model of proportion introduced ancestry on θ_W . Models examining the effect of the proportion introduced Yellowstone River ancestry found no statistically significant effect across all three measures of population genetic diversity. The effect of folded proportion introduced ancestry, or proportion admixed, was also not significantly non-zero in any model (Figure 2.5).

Genotype-environment associations

To account for population structure in our genotype-environment association analysis, cross-entropy values from **snmf** (Frichot and François 2015) showed $K=11$ as the ideal number of latent factors to include (Figure 2.44) in our LFMM models (Caye et al. 2019; Frichot, Schoville, et al. 2013). After model fitting and corrections for multiple testing, we identified 63 SNPs significantly associated with average summer temperature (Table 2.2). To compare allele frequency shifts away from introduced allele frequencies of environmentally associated loci to those of non-associated loci, we selected 6 populations found in areas with high summer temperatures (Figure 2.6a). After thinning environmentally associated loci for close proximity (911 kbp), 54 of 62 were retained for the allele frequency shift comparisons. Average allele frequency shifts away from introduced allele frequencies were higher at environmentally associated loci than at non-associated loci in all examined populations (all corrected p-values < 0.01 ; Table 2.4). Across all populations examined for average summer temperature, the average allele frequency shift at environmentally associated loci was 0.451 and the shift at non-associated loci was 0.159 (Figure 2.6b). The proportion of loci estimated to have lost alleles due to introgression varied from 0 to 0.167 in focal populations, with an average of 0.08. Of the 527 loci identified as outliers by **PCAdapt** (Luu, Bazin, and Blum 2017), 12 were also identified by **lfmm**; a Fisher’s exact test confirmed that loci identified by one method were indeed more likely to be identified by the other (odds ratio of 14.774, p-value < 0.001). Environmentally associated loci were found in or 100bp upstream of 21 different genes (see Table 2.3 for

genes and associated GO terms).

2.4 Discussion

As global climate change and habitat alteration require more direct action to conserve species, understanding the genetic effects of translocation programs has become more pressing. Documenting these effects generations after translocation can provide important insight into how human-mediated movements can be most effective in achieving conservation goals. We used collections from 231 wild and 12 hatchery populations of Yellowstone Cutthroat Trout to identify the extent of introgression from stocking programs into wild populations and consequences for population genetic diversity. We found substantial alterations of natural diversity, as introduced ancestry was found in 65% of populations surveyed (Figure 2.3) and 52% of populations within the native range of the subspecies (i.e., after excluding historically fishless habitats). Introgression of ancestry from the Yellowstone basin was the most frequently observed (Figure 2.4A-B), perhaps attributable to the long history of hatchery broodstocks from that area (Figure 2.1). We also detected associations between increased introgression and genetic diversity, though this is dependent on the introgression source and genetic diversity metric examined. Overall, we find dramatic effects of human-mediated fish movement on the landscape of genetic diversity and differentiation across the Yellowstone Cutthroat Trout range. Translocations can introduce new alleles into wild populations that may increase overall population genetic diversity, but these shared, introduced alleles also erode genetic differentiation between populations.

Our use of phylogenetic inference in conjunction with Bayesian ancestry estimation that utilizes genotype likelihoods should bolster our ability to accurately infer ancestry, even in cases of introgression from a closely-related lineage. However, the difficulty of parsing geographic differentiation versus introgression from translocations in regions

without unaltered populations to compare against may still mislead our inferences.

Translocation introgression

Our sampling efforts targeted Yellowstone Cutthroat Trout populations; it was therefore surprising to identify ancestry from other *Oncorhynchus* species in several populations. Many of these instances were from alpine lakes, which are often naturally fishless and stocked, but notable exceptions include previously unidentified Westslope Cutthroat Trout ancestry in Lodge Grass Creek (Big Horn basin), and Lahontan Cutthroat Trout ancestry in Pumpkin Creek (Big Horn basin), and Hope Creek (Upper Snake basin). The often cryptic nature of introgression between cutthroat trout species speaks to the importance of genetic assessments in making informed decisions regarding conservation action prioritization. We identified few individuals with unexpected introgression from Rainbow Trout as our sampling primarily targeted populations without known Rainbow Trout ancestry. No discovery of additional populations with Rainbow Trout ancestry is a positive finding for cutthroat trout conservation, as there are well known deleterious effects from the widespread hybridization with Rainbow Trout on Yellowstone Cutthroat Trout populations (Gresswell 2011; Al-Chokhachy et al. 2017).

Some degree of introduced ancestry was detected in 65% *O. virginalis* populations without admixture from other species (Figure 2.4). Over a century of translocations have therefore replaced over 40% of native, locally evolved ancestry with widespread, hatchery-sourced ancestry. Hatchery introgression has eroded population structure across several salmonid lineages, including Atlantic Salmon (*Salmo salar*; Ozerov et al. (2016), Glover et al. (2012), and Besnier et al. (2022)), Brook Trout (*Salvelinus fontinalis*; White et al. (2018), L  tourneau et al. (2018), Lehnert et al. (2020), Bruce et al. (2020a), and Erdman et al. (2022)), Lake Trout (*Salvelinus namaycush*; Halbisen and C. C. Wilson (2009) and Valiquette et al. (2014)), Brown Trout (*Salmo trutta*; Largiader and Scholl (1996), SANTOS et al. (2006), Hansen et al. (2009), and Casanova et al. (2022)), Chinook

Salmon (*Oncorhynchus tshawytscha*; Hess et al. (2011) and Van Doornik et al. (2013)), Coho Salmon (*Oncorhynchus kisutch*; William H. Eldridge and Kerry A. Naish (2007) and W H Eldridge, Myers, and K A Naish (2009)), and other lineages within *O. virginialis*, notably Colorado River Cutthroat Trout (*O. v. pleuriticus*; Jessica L. Metcalf, Pritchard, et al. (2007)). In general, it appears that the degree of hatchery introgression observed here is high among studied salmonid lineages, with rates of introgression similar to that observed in Lake Trout in Québec, Canada (Valiquette et al. 2014), lower than that observed in Brook Trout in New York, USA (Bruce et al. 2020b), and higher than that observed in Brook Trout in Québec, Canada (Létourneau et al. 2018).

Native biodiversity in Yellowstone Cutthroat Trout populations has been eroded but not erased. Within almost every HUC 8 basin examined, we identified at least one individual fish with greater than 98% native ancestry. Results at the population scale were much more variable: across populations containing at least one unadmixed *O. virginialis* individual, only approximately 47% contained at least one individual fish with greater than 98% native ancestry. Given the huge number of fish translocated in these waterways over the past century, including an average of over 198,000 fish into populations without introduced ancestry, the presence of any native ancestry speaks to a generally low, or at least highly variable, persistence of introduced ancestry. Studies attempting to quantify the drivers of introgression from captive-bred sources in fishes have not arrived at the same answers: habitat disturbance (Erdman et al. 2022), abiotic environmental variables (Bruce et al. 2020b), and characteristics of stocking events (Marie, L. Bernatchez, and Garant 2010; Valiquette et al. 2014; Létourneau et al. 2018; Erdman et al. 2022) have all been implicated in determining the degree of hatchery introgression observed in wild populations. Lower fitness from captive-bred and captive-raised fish has been documented in salmonids (Araki et al. 2008; Christie, Ford, and Blouin 2014) and a meta-analysis of translocations in animals found significantly lower performance of translocated individuals relative to their resident counterparts (Gross, A. E. Wilson, and Wolak 2024), so we might

expect the majority of introduced fish to reproductively contribute very little to the recipient population.

Introgression, even from less-fit, captive-bred individuals, evidently occurs in some cases and might be facilitated by ecological instability induced by large introduction events. Recent research has demonstrated that higher numbers of individuals translocated annually corresponded to higher intraspecific competition, thereby lowering fitness and decreasing average population density, in Masou Salmon (*Oncorhynchus masou*; Terui, Urabe, et al. (2023)). If fitness decreases from increased intraspecific competition after stocking events do indeed facilitate introgression, we might expect larger, more connected, and more complex meta-population networks to resist introgression more than their small, isolated counterparts (Wang, Haegeman, and Loreau 2015; Terui, Ishiyama, et al. 2018). Our results seem to support this, though we lack the comprehensive data necessary to rigorously assess it: introgression into the well-connected, large, and complex Snake River Headwaters meta-population is limited almost entirely to alpine lakes (which have limited meta-population connectivity), whereas the more fragmented and smaller populations in the Wind-Bighorn River system show significantly higher rates of introduced introgression. Altogether, introgression is expected to be unlikely unless fitness differences between local and translocated fish are minimal; this could occur through low fitness of local individuals (e.g., isolated and inbred populations in need of genetic rescue), less reduced fitness of translocated individuals (e.g., translocated fish from similar environments or with minimal time in captivity), or decreased local fitness due to increased competition (e.g., large releases relative to local population size).

Homogenization, local adaptation, and genetic diversity

Introgression from widely-spread hatchery sources has homogenized the genetic differentiation between wild Yellowstone Cutthroat Trout populations (Figure 2.4C). Inland cutthroat trouts exhibit high degrees of spatial population structure, most often

organized by hydrologic basins (R. J. Behnke 1992; Janet L Loxterman and Ernest R Keeley 2012; Jessica L. Metcalf, S. Love Stowell, et al. 2012; Shiozawa, Evans, Houston, et al. 2018), which likely explains the dramatic extent of homogenization observed here. Selecting translocation sources closely related to the recipient population could mitigate some of this phenomenon. Homogenization occurs through replacement of geographically restricted alleles with (now widespread) alleles from translocation sources; if viewed through the perspective of Whittaker’s diversity statistics (Whittaker 1972), homogenization reduces the average beta diversity and gamma diversity of the subspecies, even though it increases alpha diversity for some populations. Reduction in the total number of genetic variants present across the subspecies’ range theoretically corresponds to a reduction in adaptive potential, though the magnitude of this correlation is not yet well understood (Kardos et al. 2021).

Along similar lines, our analyses can only estimate the total proportion of native diversity lost and cannot determine how many locally adapted alleles have been lost. However, our results showing increased allele frequency shifts away from introduced ancestry allele frequencies at putatively adaptive loci (Figure 2.6) joins the body of research suggesting that introduced maladaptive alleles can be rapidly sorted from locally adapted ones provided sufficient selective pressure (Fitzpatrick et al. 2020; McLennan et al. 2025). Because we do not have genetic data from populations before admixture occurred, we cannot say that the putatively adaptive alleles were not absent in the wild populations, subsequently introduced into via stocking, and then selected for. We consider this to be the overall less likely option, as the putatively adaptive loci were often fixed for the allele being selected against in our hatchery references. If the adaptive alleles were introduced by stocking, then our results would instead be an example of adaptive introgression and have similar implications. Though further study of these dynamics in other contexts and taxa is sorely needed, it appears that local adaptation strong enough to at least partially overcome swamping can arise within lineages that have been distinct for no more than 600,000 years

(Shiozawa, Evans, Houston, et al. 2018) containing sub-lineages that arose as little as 10,000 years ago (R. Behnke 2010) (chapter 2 of this dissertation). If commonplace, this degree of local adaptation addresses some concerns regarding outbreeding depression in conservation-motivated translocations (Weeks et al. 2011). Sourcing individuals for translocation efforts from more closely-related populations should still be prioritized, however, as significant outbreeding depression is less likely and fewer alleles are likely to be lost due to shared ancestry. Natural selection can only retain alleles currently under selection, and therefore has no concern for preserving adaptive potential.

We observed a statistically significant, though small in magnitude, increase in population genetic diversity associated with introgression from fish stocking programs (Figure 2.5). Increased genetic diversity in populations with hatchery introgression (as measured by allelic richness) has been observed in other systems (Marie, L. Bernatchez, and Garant 2010; Valiquette et al. 2014), and is largely attributed to the introduction of new alleles into the population. In this system, increased genetic diversity appears to be driven by introgression from the higher diversity Snake River hatchery source (specifically the Jackson source; Figure 2.5d), which in turn drives the global pattern of increased genetic diversity with increased introduced ancestry. The mechanism suggested by these results, where mixing lineages with genetic diversities acts similar to mixing solutions with different solute concentrations, implies that translocations could also lower genetic diversity if utilizing a low-diversity source.

Conclusions

We found evidence for homogenization of Yellowstone Cutthroat Trout diversity from translocations over the past 150 years. Some amount of native ancestry persists in almost every subbasin examined despite widespread translocations, which suggests that introductions do not always lead to introgression and that the vast majority of translocated fish do not introduce new alleles into recipient populations. introduced ancestry within O .

virginalis is most often attributed to Snake River and Yellowstone River drainage sources, though we uncovered previously unknown introgression from other lineages and other species within *Oncorhynchus* as well. These results highlight the importance of utilizing genetic monitoring tools for species like cutthroat trout whose phenotypic differences are subtle and not sufficiently reliable for identification, especially in cases of hybridization. Our widespread survey of Yellowstone Cutthroat Trout populations also provides population managers with accurate designations of population status that can be used for conservation prioritization and identification of translocation sources over the coming decades.

Comparisons within 6 populations at the upper end of temperatures inhabited by Yellowstone cutthroat trout all showed evidence for increased deviation from introduced allele frequencies at putatively adaptive loci, which is consistent with sorting of adaptive alleles by natural selection. Local adaptation may therefore not be as jeopardized by gene flow as previously thought, which opens enormous potential for utilizing translocations to facilitate adaptation to climate change at multiple scales of conservation action (Meek et al. 2023).

Genetic diversity, as estimated by Watterson's θ , was found to increase as a function of proportion introduced ancestry within populations. However, the effect on genetic diversity was stronger and also observed for diversity measured by Tajima's π when examining introduced ancestry from the higher-diversity Snake River hatchery source. These results suggest that introductions from a higher diversity source population may be required to reliably increase the genetic diversity of the recipient population and further highlights the importance of genetic assessments for maximizing conservation action efficacy. Increasing genetic diversity for Yellowstone Cutthroat Trout populations may be an important benefit to their future resilience, as introduced species and habitat alteration have driven their once large, connected, and often migratory populations to primarily higher elevation, smaller, and more isolated stream segments. If increasing population

adaptive potential is the goal of management action, more consideration with regards to translocation source and introduction method could be taken to minimize genetic homogenization, and use of resources on translocations with low probability of success.

Our findings demonstrate that conservation prioritization can greatly benefit from large-scale genetic assessments to validate assumptions of genetic integrity, assess genetic diversity, and identify cryptic diversity. Genetic assessments can also benefit conservation implementation by providing the information necessary to identify source populations for translocation efforts that maximize resultant genetic diversity and minimize genetic homogenization. For systems with ample introgression, genetic assessments also enable investigation of how alleles and populations respond to gene flow which can in turn benefit future translocation efforts.

2.5 Figures



Figure 2.1: Movement of cutthroat trout through stocking has been extensive since the early 20th century, moving tens of millions of fish from hatcheries to locations across the native and introduced range of Yellowstone cutthroat trout. Yellowstone cutthroat trout stocks have been derived from different sources through the history of their stocking. Panel A shows the number of cutthroat trout (either unspecified, Bear River, Snake River, or Yellowstone) stocked in the states of Idaho, Montana, and Wyoming, from compilations of these data available from state fisheries agencies. Data from Wyoming was only available after 1985. Panel B shows the history of known Yellowstone Cutthroat Trout hatchery lines and the time periods they were used as sources for stocking.

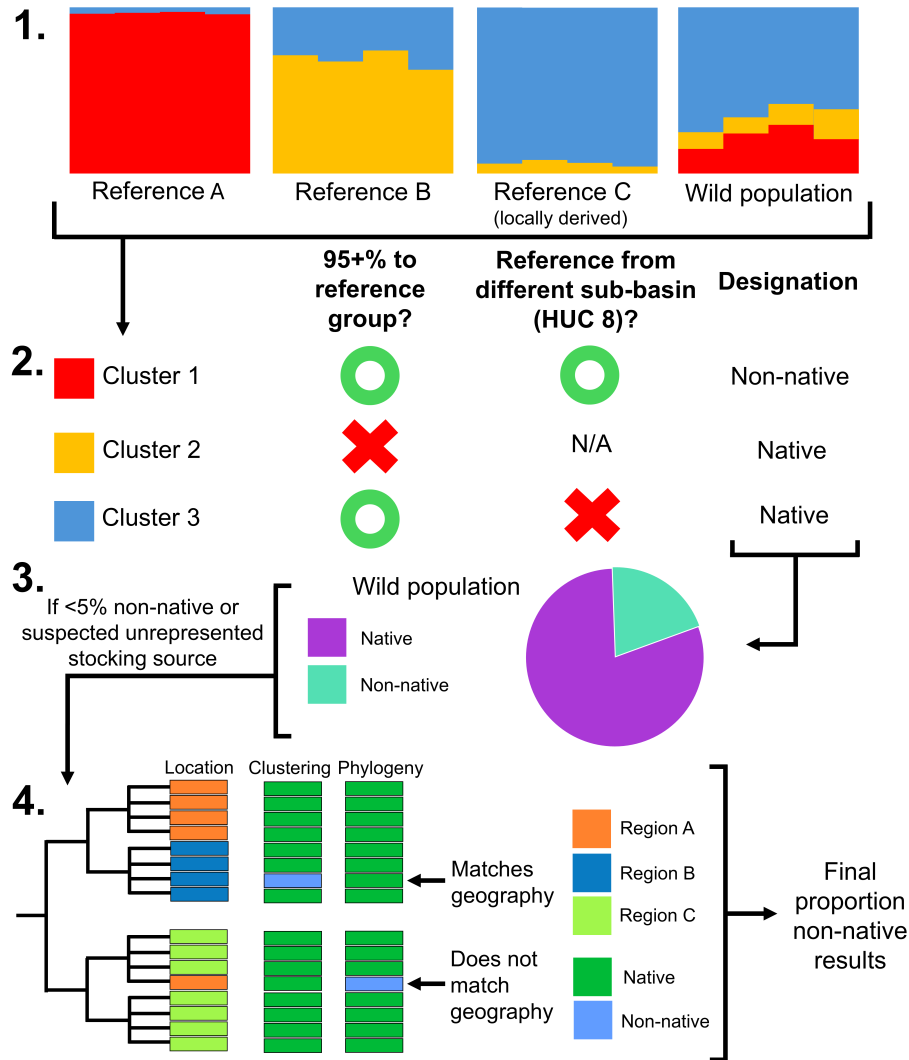


Figure 2.2: Conceptual figure of the process by which **entropy** results are used to identify introduced and native ancestry in wild populations. The methods corresponding to each of the four numbered steps are highlighted in the methods section. In the example shown, three genetic clusters are inferred from analysis of a wild population and three potential reference populations (step 1). Cluster 1 assigns strongly to Reference A, so it is designated as introduced ancestry in the focal wild population. Cluster 2 is not assigned as strongly to a reference (<95 percent), and is therefore designated as native. Cluster 3 is assigned strongly to Reference B, but this reference derives from the same HUC 8 as the focal wild population so we designate this ancestry as native in origin (step 2). After clustering, populations with < 5 percent introduced ancestry and those suspected of containing ancestry from an unrepresented stocking source are included with hatchery reference populations in a phylogenetic analysis (step 3). Designations from the clustering analysis are then revised based on the phylogenetic analysis; populations that show phylogenetic relationships consistent with geography are designated as native and those showing relationships inconsistent with geography are designated as introduced (step 4).

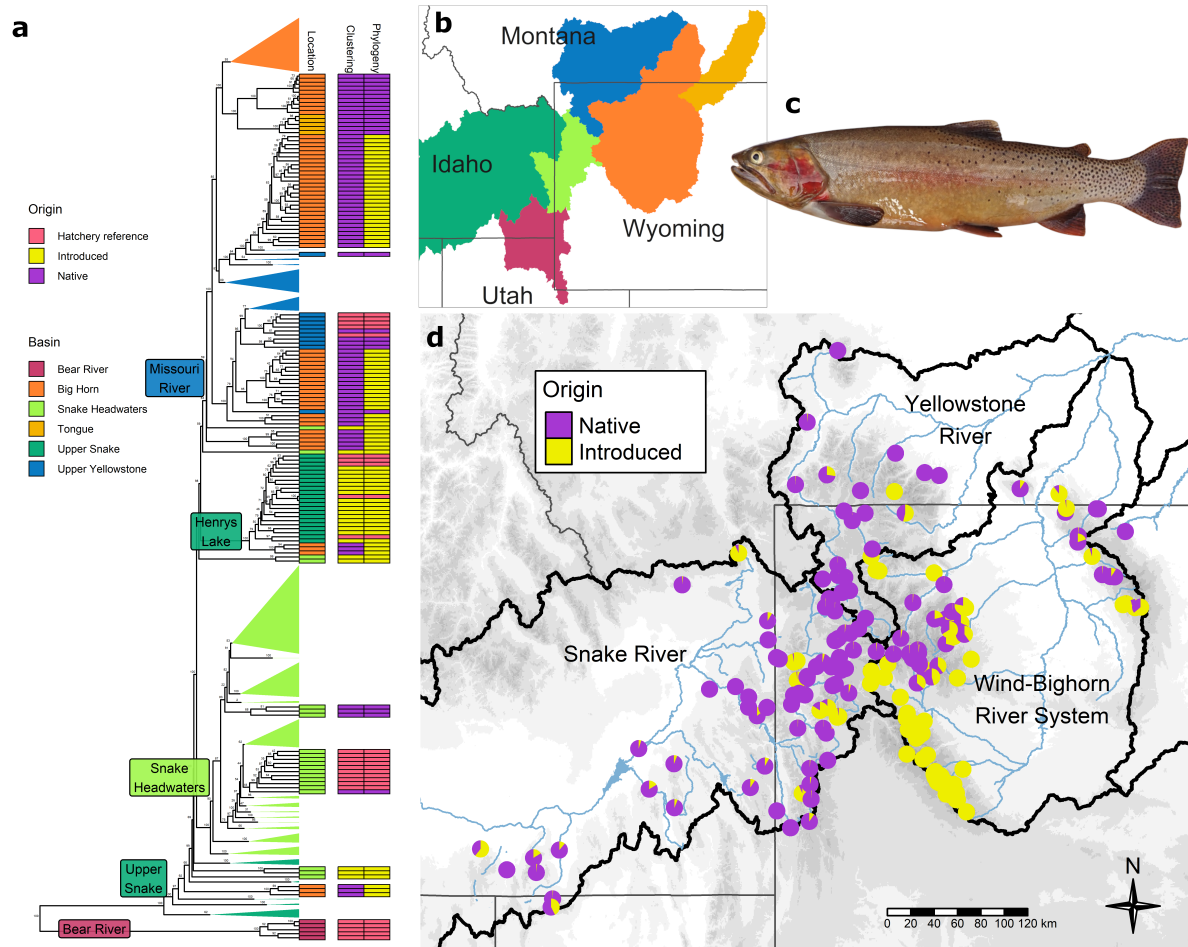


Figure 2.3: Overview of study area and introduced ancestry results. (a) Results demonstrating the phylogenetic component of introduced ancestry quantification, with colored boxes at tips corresponding to sample origin and status as inferred via clustering and phylogenetic methods. Significant geographic units are also labeled. (b) A map showing the major watersheds relevant to this study, with colors corresponding to the legend in panel a. (c) A photo of a Yellowstone cutthroat trout. (d) A map showing native ancestry distribution after both clustering and phylogenetic methods. Each pie chart represents a population and its ancestry designation.

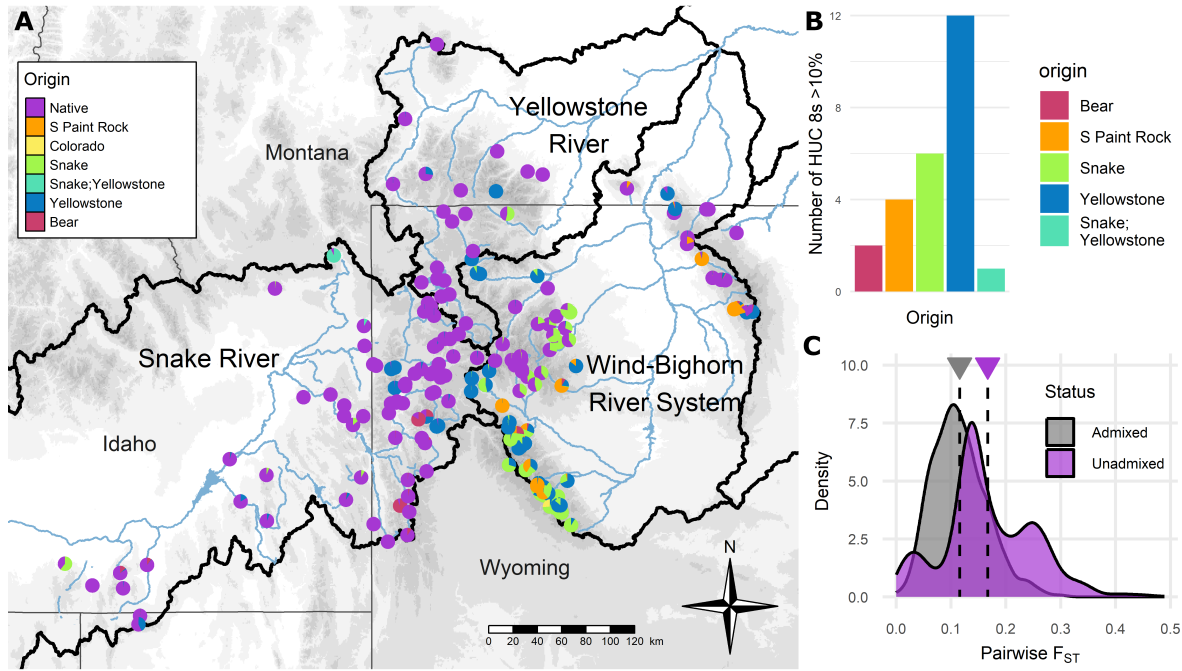


Figure 2.4: Introduced ancestry is dominantly from the Yellowstone and Snake Rivers, but ancestry from the Bear River is detectable within the Yellowstone cutthroat trout range. Results are from **entropy** and phylogenetic analyses to quantify admixture from introduced or hatchery sources pooled at the 8-digit hydrologic unit code (HUC 8) level. Panel A shows pie charts representing an average of the results across all populations within each HUC 8 area. Colors denote the proportions of native, and introduced ancestry, with introduced sources being further categorized by their river/lineage of origin. Cases where multiple lineages were matched to a cluster (i.e. Snake and Yellowstone) are denoted with blue color. Panel B shows the number of HUC 8s with > 10% ancestry assigned to each introduced ancestry cluster. Panel C shows distributions of pairwise F_{ST} estimates between all admixed ($\geq 10\%$ introduced ancestry) populations and all unadmixed ($< 10\%$ introduced ancestry) populations not from the same HUC 8 basin.

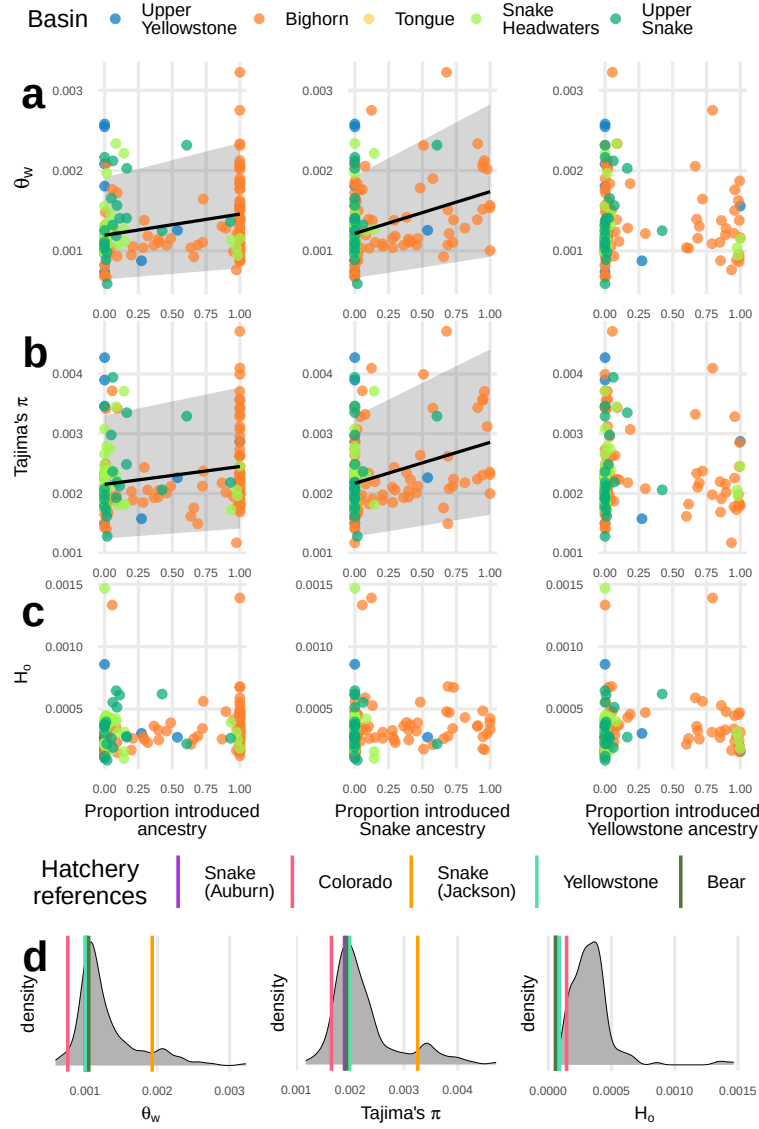


Figure 2.5: Effects of historic stocking on genetic diversity. (a) Scatter plots of three measures of introduced ancestry and on population genetic diversity as measured by Watterson's estimator (θ_W): lines and confidence intervals are shown where a significant relationship was recovered. Point colors in panels a-c correspond to population basin of origin. (b) Plots of three measures of introduced ancestry and population genetic diversity as measured by Tajima's π . Again, trend lines and confidence intervals are shown where a significant relationship was recovered. (c) Plots of three measures of introduced ancestry and population genetic diversity as measured by observed heterozygosity. (d) Density plots of three measures of wild population genetic diversity. Colored lines denote the diversity for different hatchery stocking sources.

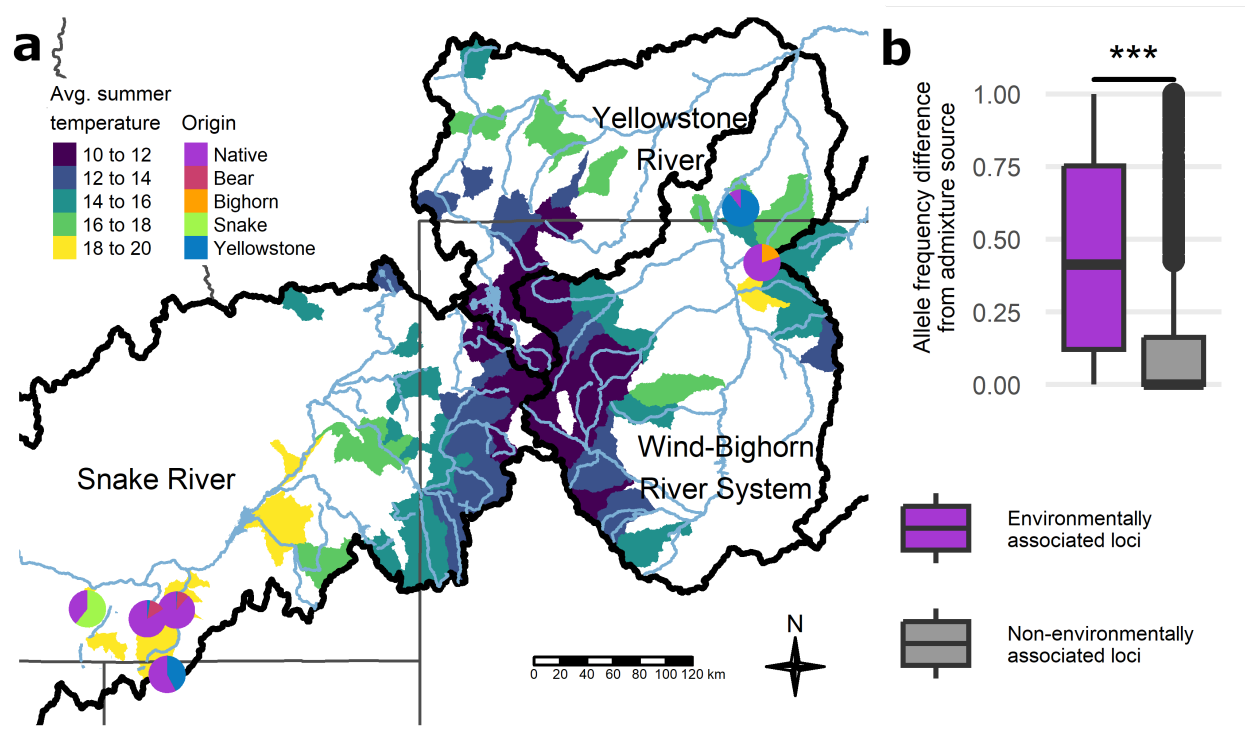


Figure 2.6: Effects of admixture on putatively adaptive loci for responses to high summer temperatures. (a) A map showing local (USGS 10-digit hydrologic unit code) 30-year average summer air temperatures. Also shown are pie charts for populations within the top 30% of average summer temperature with introduced ancestry $> 10\%$ and $< 95\%$ used for the comparisons in panel b. (b) Results from a t-test examining the allele frequency shift away from inferred admixture source frequencies for loci associated with summer air temperature and loci not associated with summer air temperature. Putatively adaptive loci exhibited a greater shift away from introduced allele frequencies than frequency-matched, non-adaptive counterparts, suggesting maintenance of adapted alleles via natural selection. Differences between groups were statistically significant across all populations and within each population.

2.6 Tables

Table 2.1: Sampling locations for wild trout populations within the range of *O. virginalis* included in this study.

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Aragon Lake	USFWS	2021	8	42.9511	-109.32
Atlantic Creek	Univ. Wyoming	2022	34	44.04964	-110.152
Baptiste Creek	WGFD	2017	17	42.85996	-109.276
Baptiste Lake	USFWS & WGFD	2021 & 2017	9	42.87435	-109.307
Baptiste Lake Creek	WGFD	2018	11	42.87278	-109.31
Barstow Lake	WGFD	2022	8	42.78446	-110.755
Basin Creek	Univ. Wyoming	2022	30	44.21644	-110.507
Bear Creek	WGFD	2019	38	43.73538	-109.446
Bears Ears Lake	WGFD	2017	30	42.85231	-109.133
Beaver Creek	Univ. Wyoming	2022	33	44.27962	-110.467
Big Bull Elk Creek	USFWS & MTFWP	2014	30	45.08749	-108.007
Bitch Creek	IDFG	2020	30	43.96359	-111.12

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Blackfoot River	IDFG	2000	30	43.1275	-112.506
Blackrock Ck	Univ. Wyoming	2020	30	43.8289	-110.357
Blacktail Springs	Univ. Wyoming	2020	30	43.66683	-110.703
Blueberry Lake	WGFD	2020	13	43.27253	-109.613
Bob Creek drainage	USFWS	2021	41	43.186	-109.443
Brushy Fork Willow Cr	MTFWP	2016	30	45.22775	-109.289
Buffalo Creek	MSU	2021	2	44.95445	-110.309
Bug Creek	WGFD	2021	30	43.77468	-109.501
Burns Canyon	IDFG	2018	30	43.60445	-111.47
Burwell Creek	WGFD	2021	10	43.88463	-109.514
Caldwell Creek	WGFD	2021	30	43.78912	-109.504
Cassia Creek	IDFG	2001	21	42.3475	-113.355
Cedar Creek	WGFD & Pisces Mol.	2006	30	44.60544	-107.654

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Cliff Ck	Univ. Wyoming	2020	30	43.24668	-110.499
Cliff Lake	WGFD & Pisces Mol.	2014	15	42.78829	-109.128
Corral Creek	IDFG	2021	52	44.38665	-112.041
Corwin Springs Sec- tion	MTFWP	2021	40	45.15776	-110.828
Cottonwood Ck	Univ. Wyoming	2020	30	43.67658	-110.707
Cottonwood Ck GV	Univ. Wyoming	2020	30	43.5553	-110.257
Cottonwood Creek	WGFD & Pisces Mol.	2011	30	43.8148	-108.947
Cottonwood Creek	IDFG	2011	30	42.2878	-113.631
Crooked Cr	MTFWP	2016	30	45.1258	-108.424
Crystal Ck	Univ. Wyoming	2020	30	43.61187	-110.429
Cub Lake	WGFD	2018	15	43.2076	-109.535
Darren Lake	WGFD	2021	23	42.90714	-109.334
Deadman Creek	Univ. Wyoming	2020	30	42.97465	-110.675

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Deadman Lake	USFWS	2021	10	43.04971	-109.405
Deadman's Bar Spring	Univ. Wyoming	2020	30	43.75892	-110.603
Death Canyon	NPS	2021	28	43.65291	-110.808
Deer Creek	WGFD & Pisces Mol.	2009	30	44.94495	-107.94
Diamond Creek	IDFG	2000	30	42.82817	-111.311
Dick Creek	WGFD	2019	14	43.9898	-109.131
Dinwoody Creek	WGFD	2018 & 2020	12	43.28453	-109.554
Dinwoody Creek WR Reservation	USFWS	2021	4	43.33915	-109.45
Dog Ck	Univ. Wyoming	2020	30	43.28527	-110.803
Dons Lake	WGFD	2018 & 2019	15	43.21065	-109.547
Double Lake	WGFD	2020	2	43.31629	-109.574
Down Meadows Dinwoody Creek	WGFD	2020	12	43.28129	-109.554

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Downs Fork Creek	WGFD	2020	2	43.28258	-109.556
Dry Creek	IDFG	1999	28	42.35605	-114.198
Dry Medicine Lodge Creek	WGFD	2021	35	44.44796	-107.412
Duck Creek	IDFG	2021	12	44.63049	-111.436
Dugout Creek	MTFWP	2016	27	46.18851	-110.375
E Fk of E Fk Wind River	WGFD	2019	13	43.7638	-109.304
E Fk Wind River	WGFD	2019	30	43.67841	-109.369
East Fork Greys River	WGFD	2021	13	42.56818	-110.669
East Fork West Red Lodge Cr	MTFWP	2016	14	45.24986	-109.444
East Rosebud Creek	MTFWP	2007	12	45.104	-109.769
East Shore Yellowstone Lake	NPS	2021	12	44.44536	-110.3
East Tensleep Creek	WGFD	2018	31	44.21543	-107.134

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Elkhorn Creek	WGFD & Pisces Mol.	2009	30	44.971	-107.607
Emerald Creek	WGFD	2021	20	43.88598	-109.52
Fall Creek	IDFG	2014	51	43.4372	-111.387
Flat Ck	Univ. Wyoming	2020	30	43.53982	-110.726
Frontier Creek	WGFD	2020	9	43.82436	-109.577
George Creek	UDWR	2022	30	41.90899	-113.443
Gibson Jack Creek	IDFG	2000	30	42.81571	-112.393
Goose Creek	IDFG	1999	30	42.19575	-113.917
Granite Ck	Univ. Wyoming	2020	30	43.28315	-110.532
Granite Creek below Turquoise lake	WGFD	2023	8	43.4419	-110.485
Grape Creek	IDFG	2016	9	42.17474	-113.601
Grassy Lake	WGFD	2019	5	43.23696	-109.494
Grave Creek	WGFD	2017	16	42.86111	-109.234
Grave Lake	WGFD	2017	6	42.86403	-109.257

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Gravel Creek	Univ. Wyoming	2022	5	43.97317	-110.373
Greybull River	WGFD	2017, 2019, 2020	24	44.094	-109.058
Grizzly Bear Lake	NPS	2021	28	43.80269	-110.811
Grouse Creek	USGS	2021	31	43.74323	-110.292
Harkness Creek	IDFG	1999	17	42.67283	-112.123
Hatchet Lake	USFWS	2021	9	43.04233	-109.364
Heart River	Univ. Wyoming	2022	35	44.24648	-110.443
Hidden Lake	WGFD	2019	30	43.39903	-109.631
High Meadow Creek	WGFD & Pisces Mol.	2014	10	42.78697	-109.105
High Meadow Lake	WGFD & Pisces Mol.	2014	15	42.79259	-109.129
Honeymoon Lake	WGFD	2020	25	43.29385	-109.569
Hope Creek Henrys Fork	IDFG	2021	25	44.61308	-111.426
Horse Creek	WGFD	2021	2	44.22637	-109.033
Horseshoe Lake	WGFD	2018	15	43.24353	-109.472

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Howard Creek	IDFG	2021	3	44.62319	-111.36
Hunter Lake	WGFD & Pisces Mol.	2014	29	44.60376	-110.019
Ishawooa Creek	WGFD	2017 & 2019	50	44.25071	-109.571
Island Lake	WGFD	2022	26	43.08516	-109.636
Kagevah Lake	Univ. Wyoming	2021	7	42.92942	-109.344
Lake 24	USFWS	2021	6	42.96161	-109.299
Lake 26	USFWS	2021	7	43.06973	-109.378
Lake 47a	USFWS	2021	6	42.93803	-109.344
Lake 74	USFWS	2021	10	43.06254	-109.41
Lake 84	USFWS	2021	8	43.00615	-109.414
Lake 85	USFWS	2021	6	43.00908	-109.406
Lake 86	USFWS	2021	4	43.0095	-109.404
Lake 87	USFWS	2021	5	42.98601	-109.411
Lake 87a	USFWS	2021	5	42.9912	-109.405
Lake 88	USFWS	2021	2	42.98677	-109.415
Lake 90	USFWS	2021	1	42.98936	-109.417
Lake 92	USFWS	2021	3	42.97983	-109.41
Lake 99	USFWS	2021	7	42.95415	-109.383
Lake Solitude	NPS	2021	25	43.79319	-110.845
Lamar River	MSU	2021	23	44.8809	-110.22

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Leidy Creek	USGS	2021	75	43.73024	-110.338
Little Lamar	MTFWP	2016	30	44.66252	-110.006
Little Rocky Creek	MTFWP	2020	25	45.39915	-109.757
Little Warm Springs Creek	WGFD	2018	33	43.52246	-109.705
Lodgegrass Creek	WGFD & Pisces Mol.	2006	29	44.98352	-107.84
Lonesome Lake	WGFD & Pisces Mol.	2014	15	42.77762	-109.215
Long Creek	WGFD	2021	10	43.67531	-109.876
Lower Bar BC	Univ. Wyoming	2020	30	43.54541	-110.792
Lower Black Joe Lake	Univ. Wyoming	2019	30	42.73918	-109.172
Lower Deep Creek	WGFD	2018	13	42.73369	-109.094
Lower Fish Ck	Univ. Wyoming	2020	30	43.46734	-110.87
Lower Frontier Creek	WGFD	2020 & 2021	21	43.82026	-109.574
Lower Glacier Lake	WGFD	2018	15	43.19774	-109.561

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Lower Ice Lake	Univ. Wyoming	2019	7	42.92942	-109.344
Lower Lake Creek	WGFD	2018	19	44.2113	-107.193
Lower Ross Lake	WGFD	2019	19	43.3866	-109.656
Lower Wiggins Fork Creek	WGFD	2019	30	43.63158	-109.526
Lydle Lake	USFWS	2021	10	43.07231	-109.328
MacLeod Lake	WGFD	2023	30	43.36863	-110.379
Main Basin Yellowstone Lake	NPS	2021	12	44.45982	-110.371
Marked Tree Lake	WGFD	2018	15	43.04711	-109.463
Marquette Creek	WGFD	2021	26	44.38864	-109.238
Marten Creek	WGFD	2022	29	42.73899	-110.66
Meeteetse Creek	WGFD	2022	5	44.21055	-109.007
Middle Deep Creek	WGFD	2018	13	42.73449	-109.106

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Middle Fork Dry Creek	WGFD	2019	10	43.23514	-109.51
Middle Wiggins Fork Creek	WGFD	2021	15	43.81744	-109.55
Mill Creek	WGFD	2021	20	44.45068	-107.457
Miner Cr	MTFWP	2009	25	45.6381	-110.703
Mink Creek	Univ. Wyoming	2022	2	44.01652	-110.299
Moose Lake	WGFD	2019	15	43.22134	-109.514
Muddy Creek	WGFD	2021	23	44.93828	-109.653
N Beaver Creek	WGFD	2021	32	44.76255	-107.801
N Buffalo Fk	Univ. Wyoming	2020	30	43.88506	-110.208
Native Lake	WGFD	2019	4	43.23793	-109.499
NF Little Wind River	USFWS	2016	1	43.02576	-109.109
NFS Grinnell Creek	Univ. Wyoming	2016	1	44.4937	-109.934
NFS Middle Creek	Univ. Wyoming	2020	30	44.5004	-109.966
NFS Trout Creek	Univ. Wyoming	2019	30	44.479	-109.344

Population Name	Sample provider	Year	Number of Samples	Lat	Long
North Fork Popo Agie River	WGFD & Pisces Mol.	2014 & 2021	15	42.76645	-109.116
North Leigh Creek Spring Creek	IDFG	2005	19	43.82755	-111.031
North Yel- lowstone Lake	NPS	2021	13	44.54058	-110.374
Onemile Creek	UDWR	2022	30	41.97157	-113.427
Pacific above Mink	Univ. Wyoming	2022	16	44.01508	-110.297
Pacific Creek	Univ. Wyoming	2022	4	44.01414	-110.244
Palisades Creek	IDFG	2018	30	43.37986	-111.238
Palisades Creek Tribu- taries	IDFG	2018	74	43.44437	-111.134
Papoose Lake	WGFD	2014	4	42.76789	-109.169
Phillips Lake	WGFD	2019	4	43.24042	-109.491
Pickett Creek	WGFD	2021	30	44.12175	-109.224
Pine Creek	IDFG	2020	30	43.52384	-111.329

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Piney Creek	WGFD	2021	22	44.12421	-109.338
Pumpkin Creek	WGFD & Pisces Mol.	2007	26	44.96245	-107.716
Raft Creek	WGFD	2017	19	42.87546	-109.232
Rainey Creek	IDFG	2018	30	43.44616	-111.33
Rawhide Creek	WGFD	2021	25	44.1817	-109.15
Red Gulch Creek	WGFD & Pisces Mol.	2007	29	44.97108	-107.586
Robert's Lake	USFWS	2021	23	42.92042	-109.277
Robert's Lake Drainage	USFWS	2021	35	42.91528	-109.269
Rock Creek	USGS	2021	42	43.75577	-110.441
Ross Fork	ISU	2003	33	43.01087	-112.128
S Beaver Creek	WGFD	2021	29	44.71541	-107.806
S. Fork Little Tongue River	WGFD & Pisces Mol.	2006	28	44.79674	-107.297
Salt River	WGFD	2021 & 2022	30	42.51865	-110.88
SF Bull Lake Ck	USFWS	2021	3	43.07721	-109.415

Population Name	Sample provider	Year	Number of Samples	Lat	Long
SF Hoback River	WGFD	2022	50	43.03752	-110.482
SF Little Wind Ranger Park	WGFD	2017	8	42.84453	-109.217
SF Owl Creek	USFWS	2016	29	43.67024	-109.094
SF Wood River	WGFD	2021	15	43.93132	-109.133
Shangrala Dinwoody Creek	WGFD	2020	6	43.31867	-109.556
Sheep Creek	WGFD	2022	30	42.85405	-110.672
Sheridan Creek	WGFD	2018	32	43.62812	-110.025
Shoal Lake	WGFD	2023	25	43.37514	-110.363
Sickle Creek	Univ. Wyoming	2022	30	44.186	-110.407
Slate Ck	Univ. Wyoming	2020	30	43.62104	-110.401
Slough Creek	MSU	2021	52	45.11123	-110.131
Soda Butte Creek	MSU	2021	3	44.93789	-110.083
Soda Fork Buffalo River	MSU	2021	5	43.87486	-109.963

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Soldier Creek	WGFD	2021	34	44.24169	-107.288
South Arm Yellowstone Lake	NPS	2021	13	44.3277	-110.33
South Fork Shoshone River	WGFD	2021	19	43.97637	-109.701
South Leigh Creek	IDFG	2015	30	43.81686	-111.001
South Paint Rock Creek	WGFD & Pisces Mol.	2006	29	44.23392	-107.321
Southeast Arm Yellow- stone Lake	NPS	2021	13	44.34313	-110.248
Spider Lake	Univ. Wyoming	2019	21	42.95063	-109.331
Spread Creek	USGS	2021	2	43.772	-110.401
Spread Creek ditch by mouth	WGFD	2021	14	43.79174	-110.538
Spring Creek	USFS	2022	30	42.64851	-111.027
SR side chan- nel	Univ. Wyoming	2020	30	43.68275	-110.701

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Steamboat Lake	USFWS	2021	4	43.06498	-109.354
Steer Creek	WGFD	2021	30	43.09386	-110.7
Stonefly Creek	WGFD	2018	31	43.72351	-110.018
Stough Creek	WGFD	2016	5	42.6408	-108.997
Surprise Creek	Univ. Wyoming	2022	24	44.25348	-110.439
Thorofare Creek	WGFD	2018	2	44.12988	-110.078
Three Channel Spring	Univ. Wyoming	2020	30	43.5527	-110.789
Timber Creek Henrys Fork	IDFG	2021	3	44.66979	-111.429
Trapper Creek	WGFD	2021	30	44.46536	-107.543
Trout Creek Bighorn Lake	WGFD	2022	56	44.97442	-107.928
Turquoise lakes	WGFD	2023	9	43.42135	-110.566
Twin Lake	WGFD	2020	15	43.27457	-109.62
Unnamed Graves Lakes	WGFD	2017	14	42.88331	-109.281

Population Name	Sample provider	Year	Number of Samples	Lat	Long
Upper Bar BC	Univ. Wyoming	2020	28	43.67816	-110.705
Upper Fish Ck	Univ. Wyoming	2020	30	43.53827	-110.848
Upper Frontier Creek	WGFD	2021	9	43.85848	-109.607
Upper Ross Lake	WGFD	2019	24	43.36265	-109.642
Upper SF Shoshone River	WGFD	2021	13	43.85442	-109.779
Upper Tin-cup Creek	USFS	2022	39	42.99297	-111.155
Upper Wiggins Fork Creek	WGFD	2021	15	43.86161	-109.51
Upper Wood River	WGFD	2019	8	43.93355	-109.216
Wallace Creek	MTFWP	2013	17	45.23262	-110.491
Washakie Reservoir	USFWS	2021	13	42.96857	-109.035
Watkins Lake	WGFD	2021	24	43.77774	-109.84

Population Name	Sample provider	Year	Number of Samples	Lat	Long
West Fork Greys River	WGFD	2021	17	42.56747	-110.679
West Thumb Yellowstone Lake	NPS	2021	12	44.43156	-110.535
Whetstone Creek	Univ. Wyoming	2022	8	43.958	-110.405
Willow Creek	WGFD	2021	30	43.25777	-110.682
Willow Creek ID	IDFG	2020	30	43.5845	-111.748
Wittombona Lake	USFWS	2021	13	43.081	-109.371
Wood River	WGFD	2019	30	44.0063	-109.02
Wykee Lake	USFWS	2021	1	42.94125	-109.317
Wyoming Creek	IDFG	2006	20	44.11032	-111.126

Table 2.2: Information for loci identified as environmentally associated by LFMM. Chromosome and position information are from the Rainbow Trout reference genome, accession GCA_013265735.3.

CHROM	POS	Corrected p-value
NW_023493667.1	1044444	0.04183
NC_048584.1	8122406	0.011022
NC_048584.1	8404897	0.024201
NC_048584.1	16852143	0.000378
NC_048584.1	40479071	0.042119
NC_050572.1	17170273	0.014688
NC_050572.1	36286484	0.030357
NC_050570.1	31332037	0.032622
NC_050570.1	31602887	0.00621
NC_048593.1	22883413	0.000521
NC_048586.1	14617107	0.020235
NC_048589.1	10032406	0.010472
NC_048589.1	13967305	0.000865
NC_048585.1	22730677	0.01324
NC_048585.1	54718824	0.031606
NC_048587.1	12603956	0.024583
NC_048587.1	17473290	0.001521
NC_048587.1	24167726	0.001248
NC_048577.1	54439067	0.024201
NC_048577.1	54439138	0.024201
NC_048573.1	45339082	0.020235
NC_048579.1	14969287	0.000865
NC_048579.1	35478119	0.011378

CHROM	POS	Corrected p-value
NC_048579.1	35521191	0.000521
NC_048579.1	57620966	0.025574
NC_048579.1	71689040	0.044672
NC_048567.1	15638190	0.006075
NC_048567.1	32347824	0.025943
NC_048567.1	56742244	0.014688
NC_048567.1	57926451	0.035219
NC_048580.1	27039078	0.000378
NC_048580.1	45285130	0.020235
NC_048574.1	15265504	0.009388
NC_048574.1	15265508	0.002834
NC_048574.1	15265530	0.002834
NC_048574.1	68644818	0.024201
NC_048574.1	74074024	0.021254
NC_048574.1	74074029	0.005786
NC_048571.1	24812155	0.014688
NC_048571.1	51203094	0.014688
NC_048571.1	79060618	0.020235
NC_048571.1	79343047	0.014948
NC_048571.1	88880138	0.024201
NC_048571.1	88880143	0.024201
NC_048571.1	88880149	0.024201
NC_048571.1	88880163	0.019942
NC_048572.1	20151393	0.009388
NC_048572.1	20218511	0.01623

CHROM	POS	Corrected p-value
NC_048570.1	21742297	0.042119
NC_048570.1	43672308	0.035219
NC_048570.1	43859472	0.045862
NC_048570.1	45457736	0.009388
NC_048570.1	71996204	0.024201
NC_048565.1	32308309	0.00621
NC_048566.1	37888355	0.012078
NC_048569.1	36635579	0.030357
NC_048569.1	40996993	0.002106
NC_048569.1	44575962	0.029366
NC_048569.1	78420280	1.60E-05
NC_048569.1	97343239	0.024201
NC_048576.1	29144507	0.044672
NC_048576.1	75015876	0.022427

Table 2.3: Gene information and gene ontology terms for loci found within (or 100bp downstream of) annotated genes in the rainbow trout reference genome.

Entrez Gene ID	GO ID	GO name	Gene Name
100301703	GO:0000166	nucleotide binding	ube2c
100301703	GO:0005524	ATP binding	ube2c
100301703	GO:0016740	transferase activity	ube2c
100301703	GO:0051301	cell division	ube2c
110488023	GO:0042147	retrograde transport, endosome to Golgi	LOC110488023
110488023	GO:0032456		LOC110488023
110488023	GO:1990745	EARP complex	LOC110488023
110488023	GO:0015031	protein transport	LOC110488023
110488023	GO:0005829	cytosol	LOC110488023
110488023	GO:0000149	SNARE binding	LOC110488023
110489035	GO:0045944	positive regulation of transcription by RNA polymerase II	LOC110489035
110489035	GO:0016592	mediator complex	LOC110489035
110489035	GO:0003677	DNA binding	LOC110489035
110489035	GO:0003712	transcription coregulator activity	LOC110489035
110489035	GO:0006397	mRNA processing	LOC110489035
110489035	GO:0035145	exon-exon junction complex	LOC110489035
110489035	GO:0003729	mRNA binding	LOC110489035
110489035	GO:0005634	nucleus	LOC110489035
110489035	GO:0005737	cytoplasm	LOC110489035
110489035	GO:0008380	RNA splicing	LOC110489035
110489035	GO:0003723	RNA binding	LOC110489035

Entrez Gene ID	GO ID	GO name	Gene Name
110489035	GO:0000184		LOC110489035
110489035	GO:0006417	regulation of translation	LOC110489035
110489035	GO:0051028	mRNA transport	LOC110489035
110490714	GO:0045892	negative regulation of DNA-templated transcription	LOC110490714
110490714	GO:0005654	nucleoplasm	LOC110490714
110490714	GO:0045814	negative regulation of gene expression, epigenetic	LOC110490714
110490714	GO:0097355	protein localization to heterochromatin	LOC110490714
110499548	GO:0003824	catalytic activity	LOC110499548
110499548	GO:0016787	hydrolase activity	LOC110499548
110499548	GO:0016020	membrane	LOC110499548
110499548	GO:0005783	endoplasmic reticulum	LOC110499548
110499548	GO:0005789	endoplasmic reticulum membrane	LOC110499548
110499548	GO:0033961	cis-stilbene-oxide hydrolase activity	LOC110499548
110499548	GO:0009636	response to toxic substance	LOC110499548
110499548	GO:0006629	lipid metabolic process	LOC110499548
110499548	GO:0016803	ether hydrolase activity	LOC110499548
110499548	GO:0019369	arachidonate metabolic process	LOC110499548
110499548	GO:0004301	epoxide hydrolase activity	LOC110499548
110499548	GO:0097176	epoxide metabolic process	LOC110499548
110502275	GO:0005515	protein binding	LOC110502275
110502275	GO:0070161	anchoring junction	LOC110502275
110502275	GO:0005737	cytoplasm	LOC110502275

Entrez Gene ID	GO ID	GO name	Gene Name
110502275	GO:0005634	nucleus	LOC110502275
110502275	GO:0030055	cell-substrate junction	LOC110502275
110502275	GO:0031589	cell-substrate adhesion	LOC110502275
110504185	GO:0005515	protein binding	kidins220a
110504185	GO:0016020	membrane	kidins220a
110504185	GO:0019887	protein kinase regulator activity	kidins220a
110504185	GO:0030165	PDZ domain binding	kidins220a
110509819	GO:0045666	positive regulation of neuron differentia- tion	LOC110509819
110509819	GO:0045930	negative regulation of mitotic cell cycle	LOC110509819
110509819	GO:0005737	cytoplasm	LOC110509819
110509819	GO:0051726	regulation of cell cycle	LOC110509819
110509819	GO:0007399	nervous system development	LOC110509819
110509819	GO:0030425	dendrite	LOC110509819
110509819	GO:0043025	neuronal cell body	LOC110509819
110509819	GO:0045786	negative regulation of cell cycle	LOC110509819
110509819	GO:0071300	cellular response to retinoic acid	LOC110509819
110512261	GO:0006779	porphyrin-containing compound biosyn- thetic process	ppox
110512261	GO:0016491	oxidoreductase activity	ppox
110512261	GO:0004729	oxygen-dependent protoporphyrinogen oxidase activity	ppox
110512261	GO:0016020	membrane	ppox
110512261	GO:0005739	mitochondrion	ppox

Entrez Gene ID	GO ID	GO name	Gene Name
110512261	GO:0005743	mitochondrial inner membrane	ppox
110512261	GO:0031966	mitochondrial membrane	ppox
110512261	GO:0005758	mitochondrial intermembrane space	ppox
110512261	GO:0006782	protoporphyrinogen IX biosynthetic process	ppox
110512261	GO:0006783	heme biosynthetic process	ppox
110520268	GO:0005886	plasma membrane	adcy5
110520268	GO:0035556	intracellular signal transduction	adcy5
110520268	GO:0009190	cyclic nucleotide biosynthetic process	adcy5
110520268	GO:0004016	adenylate cyclase activity	adcy5
110520268	GO:0006171	cAMP biosynthetic process	adcy5
110520268	GO:0016849	phosphorus-oxygen lyase activity	adcy5
110520268	GO:0046872	metal ion binding	adcy5
110520268	GO:0016020	membrane	adcy5
110520268	GO:0000166	nucleotide binding	adcy5
110520268	GO:0005524	ATP binding	adcy5
110520268	GO:0016829	lyase activity	adcy5
110520268	GO:0042995	cell projection	adcy5
110520268	GO:0005929	cilium	adcy5
110520268	GO:0007189	adenylate cyclase-activating G protein-coupled receptor signaling pathway	adcy5
110521325	GO:0006368	transcription elongation by RNA polymerase II	LOC110521325
110521325	GO:0008023	transcription elongation factor complex	LOC110521325

Entrez Gene ID	GO ID	GO name	Gene Name
110521325	GO:0005634	nucleus	LOC110521325
110521325	GO:0006366	transcription by RNA polymerase II	LOC110521325
110521325	GO:0000987	cis-regulatory region sequence-specific DNA binding	LOC110521325
110521325	GO:0032968	positive regulation of transcription elon- gation by RNA polymerase II	LOC110521325
110521325	GO:0042795	snRNA transcription by RNA polymerase II	LOC110521325
110523827	GO:0055085	transmembrane transport	LOC110523827
110523827	GO:0016020	membrane	LOC110523827
110523827	GO:0005739	mitochondrion	LOC110523827
110523827	GO:0005743	mitochondrial inner membrane	LOC110523827
110523827	GO:0015297	antiporter activity	LOC110523827
110523827	GO:0071913	citrate secondary active transmembrane transporter activity	LOC110523827
110523827	GO:0006843	mitochondrial citrate transmembrane transport	LOC110523827
110523827	GO:0007528	neuromuscular junction development	LOC110523827
110523916	GO:0005737	cytoplasm	LOC110523916
110523916	GO:0005884	actin filament	LOC110523916
110523916	GO:0005886	plasma membrane	LOC110523916
110523916	GO:0051015	actin filament binding	LOC110523916
110523916	GO:0030041	actin filament polymerization	LOC110523916
110526032	GO:0016020	membrane	kremen1

Entrez Gene ID	GO ID	GO name	Gene Name
110526032	GO:0016055	Wnt signaling pathway	kremen1
110526032	GO:0005886	plasma membrane	kremen1
110526506	GO:0006886	intracellular protein transport	LOC110526506
110526506	GO:0016192	vesicle-mediated transport	LOC110526506
110526506	GO:0030117	membrane coat	LOC110526506
110526506	GO:0015031	protein transport	LOC110526506
110526506	GO:0030131	clathrin adaptor complex	LOC110526506
110526506	GO:0035615	clathrin adaptor activity	LOC110526506
110526506	GO:0072583	clathrin-dependent endocytosis	LOC110526506
110526506	GO:0030122	AP-2 adaptor complex	LOC110526506
110526506	GO:0016020	membrane	LOC110526506
110526506	GO:0005886	plasma membrane	LOC110526506
110526506	GO:0005905	clathrin-coated pit	LOC110526506
110526506	GO:0006897	endocytosis	LOC110526506
110526506	GO:0005794	Golgi apparatus	LOC110526506
110526506	GO:0030121	AP-1 adaptor complex	LOC110526506
110527557	GO:0003677	DNA binding	LOC110527557
110527557	GO:0006355	regulation of DNA-templated transcrip- tion	LOC110527557
110527557	GO:0005634	nucleus	LOC110527557
110527973	GO:0016020	membrane	LOC110527973
110527973	GO:0008021	synaptic vesicle	LOC110527973
110528552	GO:0005886	plasma membrane	LOC110528552
110528552	GO:0016020	membrane	LOC110528552

Entrez Gene ID	GO ID	GO name	Gene Name
110528552	GO:0005509	calcium ion binding	LOC110528552
110528552	GO:0007156	homophilic cell adhesion via plasma mem- brane adhesion molecules	LOC110528552
110528552	GO:0007155	cell adhesion	LOC110528552
110528552	GO:0098609	cell-cell adhesion	LOC110528552
110528552	GO:0046872	metal ion binding	LOC110528552
110528552	GO:0016477	cell migration	LOC110528552
110528552	GO:0005912	adherens junction	LOC110528552
110528552	GO:0034332	adherens junction organization	LOC110528552
110528552	GO:0045216	cell-cell junction organization	LOC110528552
110528552	GO:0045296	cadherin binding	LOC110528552
110528552	GO:0000902	cell morphogenesis	LOC110528552
110528552	GO:0007043	cell-cell junction assembly	LOC110528552
110528552	GO:0016339		LOC110528552
110528552	GO:0044331	cell-cell adhesion mediated by cadherin	LOC110528552
110528552	GO:0016342		LOC110528552
110528552	GO:0008013	beta-catenin binding	LOC110528552
110528552	GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules	LOC110528552
110528552	GO:0002009	morphogenesis of an epithelium	LOC110528552
110534571	GO:0005515	protein binding	LOC110534571
110534571	GO:0005737	cytoplasm	LOC110534571
110538772	GO:0005524	ATP binding	LOC110538772
110538772	GO:0005515	protein binding	LOC110538772

Entrez Gene ID	GO ID	GO name	Gene Name
110538772	GO:0006468	protein phosphorylation	LOC110538772
110538772	GO:0004674	protein serine/threonine kinase activity	LOC110538772
110538772	GO:0004672	protein kinase activity	LOC110538772
110538772	GO:0004540	RNA nuclease activity	LOC110538772
110538772	GO:0006397	mRNA processing	LOC110538772
110538772	GO:0004521	RNA endonuclease activity	LOC110538772
110538772	GO:0030968	endoplasmic reticulum unfolded protein response	LOC110538772
110538772	GO:0046872	metal ion binding	LOC110538772
110538772	GO:0016787	hydrolase activity	LOC110538772
110538772	GO:0006915	apoptotic process	LOC110538772
110538772	GO:0016020	membrane	LOC110538772
110538772	GO:0000166	nucleotide binding	LOC110538772
110538772	GO:0016301	kinase activity	LOC110538772
110538772	GO:0016740	transferase activity	LOC110538772
110538772	GO:0005783	endoplasmic reticulum	LOC110538772
110538772	GO:0005789	endoplasmic reticulum membrane	LOC110538772
110538772	GO:0003824	catalytic activity	LOC110538772
110538772	GO:0006986	response to unfolded protein	LOC110538772
110538772	GO:0010629	negative regulation of gene expression	LOC110538772
110538772	GO:0042803	protein homodimerization activity	LOC110538772
110538772	GO:0032991	protein-containing complex	LOC110538772
110538772	GO:0051082	unfolded protein binding	LOC110538772
110538772	GO:0070059		LOC110538772

Entrez Gene ID	GO ID	GO name	Gene Name
110538772	GO:0036498		LOC110538772
110538772	GO:1990604	IRE1-TRAF2-ASK1 complex	LOC110538772
110538772	GO:1905898	positive regulation of response to endoplasmic reticulum stress	LOC110538772
110538772	GO:0033120	positive regulation of RNA splicing	LOC110538772
110538772	GO:0034976	response to endoplasmic reticulum stress	LOC110538772

Table 2.4: Results from allele frequency shift comparisons for admixed populations in relatively high temperatures.

Population	Prop. Bear	Prop. Snake	Prop. Y.S.	Total intro- duced	Avg. shift EA SNPs	Avg shift other SNPs	T-test p- value	Adj. t-test p- value	Prop. Swamped
Big Bull Elk Creek	0.000	0.000	0.103	0.103	0.4456	0.2074	1.3E-04	1.2E-04	0.167
Dry Creek	0.000	0.714	0.000	0.714	0.4966	0.1629	2.32E-08	4.6E-08	0.074
Cassia Creek	0.014	0.099	0.000	0.114	0.5204	0.1909	7.91E-07	1.19E-06	0.129
S Beaver Creek	0.000	0.000	0.803	0.803	0.2939	0.0416	2.31E-10	6.93E-10	0.056
Cottonwood Creek	0.028	0.195	0.000	0.223	0.5276	0.1583	6.46E-11	3.87E-12	0
George Creek	0.072	0.000	0.463	0.535	0.4188	0.1911	1.80E-05	2.16E-05	0.037

Table 2.5: Genetic diversity estimates for each population with greater than 5 samples. The quantiles of Watterson’s theta (thetaw) values are calculated from the distribution of populations excluding references from other species and subspecies as well as populations from historically fishless alpine lakes.

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
AragonLake	10080002	0.001416	NA	0.002287	0.000301
ArleeBroodstock	NA	0.000852	NA	0.001151	0.000575
AtlanticCreek	10070001	0.001051	0.327	0.002025	0.000321
AuburnSRC	NA	0.001005	0.226	0.001894	5.99E-05
BaptisteCreek	10080002	0.001536	0.818	0.002359	0.000389
BaptisteLake	10080002	0.001389	NA	0.002238	0.000361
BaptisteLakeCreek	10080002	0.002086	NA	0.003558	0.000278
BarstowLake	17040103	0.001136	NA	0.001727	0.000393
BasinCreek	17040101	0.000998	0.195	0.001796	0.000418
BearCreek	10080001	0.001153	0.553	0.002119	0.000389
BearsEarsLake	10080003	0.001004	NA	0.001997	0.000337
BeaverCreek	17040101	0.000984	0.151	0.001846	0.000396
BigBullElkCreek	10080010	0.00128	0.698	0.002017	0.000479
BitchCreekTetonRiver	17040204	0.00103	0.264	0.001785	0.000377
BlackfootRiver	17040207	0.001655	0.849	0.002978	0.000264
BlackrockCk	17040101	0.001061	0.352	0.001956	0.000411
BlacktailSprings	17040101	0.001139	0.522	0.002134	0.000411
BlueberryLake	10080001	0.002126	NA	0.003709	0.000174
BrushyForkWillowCr	10070006	0.000999	0.201	0.002108	0.000159
BuffaloCreek	10070001	0.002084	0.943	0.00345	0.000858
BugCreek	10080001	0.001765	0.862	0.003718	0.001333

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
BurnsCanyon	17040104	0.00105	0.308	0.001895	0.00023
BurwellCreek	10080001	0.001112	0.465	0.001757	0.000301
CaldwellCreek	10080001	0.001036	0.277	0.001886	0.000332
CassiaCreekRaftRiver	17040210	0.001405	0.786	0.002123	0.000612
CedarCreek	10080010	0.00099	0.170	0.001944	0.000415
CliffCk	17040103	0.001117	0.478	0.001819	0.000315
CliffLake	10080003	0.001503	NA	0.002255	0.000446
CorralCreek	17040214	0.000591	0.006	0.00128	8.95E-05
CorwinSpringsYellowstoneRiver	10070002	0.001072	0.390	0.001908	0.000181
CottonwoodCk	17040101	0.001092	0.434	0.002033	0.000398
CottonwoodCkGV	17040102	0.001322	0.730	0.002745	0.000165
CottonwoodCreek	10080007	0.000879	0.082	0.001687	0.000318
CottonwoodCreekClydeCreek	17040210	0.001408	0.792	0.002486	0.00028
CRCbroodstock	NA	0.000761	0.038	0.001651	0.000145
CrookedCr	10080010	0.001724	0.855	0.003425	0.000149
CrystalCk	17040102	0.001159	0.572	0.002131	0.000358
CubLake	10080001	0.00138	NA	0.002449	0.000196
DanielhatcheryGDT	NA	0.000655	NA	0.001222	0.000417
DarrenLake	10080002	0.001033	NA	0.001821	0.000263
DeadmanCreek	17040103	0.00133	0.748	0.002381	0.000376
DeadmanLake	10080001	0.001206	NA	0.001918	0.000345
DeadmansBarSpring	17040101	0.001219	0.629	0.002472	0.000406
DeathCanyon	17040101	0.000944	0.113	0.001957	0.000253
DeerCreek	10080010	0.000673	0.013	0.001403	0.000175

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
DiamondCreekBlackfootRiver	17040207	0.001565	0.836	0.00219	0.000514
DickCreek	10080009	0.001106	0.440	0.002122	0.000344
DinwoodyCreek	10080001	0.002313	0.969	0.003994	0.000178
DogCk	17040103	0.001109	0.459	0.002051	0.00042
DonsLake	10080001	0.002286	NA	0.003455	0.001234
DownMeadowsDinwoodyCreek	10080001	0.001901	0.893	0.003428	0.000214
DryCreekSnakeRiver	17040212	0.002316	0.975	0.003293	0.00022
DryMedicineLodgeCreek	10080008	0.000681	0.019	0.001424	0.000146
DuckCreekHenrysFork	17040202	0.001366	0.755	0.002184	0.000265
DugoutCreek	10070003	0.001174	0.579	0.002243	0.000131
EastForkGreysRiver	17040103	0.001446	0.799	0.00258	0.00026
EastForkWestRedLodgeCr	10070006	0.000989	0.164	0.0019	0.000191
EastRosebudCreek	10070005	0.00156	0.830	0.002872	0.00016
EastShoreYSL	10070001	0.001065	0.371	0.001807	0.000277
EastTensleepCreek	10080008	0.001041	0.289	0.001616	0.000216
EFkofEFkWindRiver	10080001	0.001147	0.535	0.001937	0.000332
EFkWindRiver	10080001	0.001073	0.396	0.002006	0.000323
ElkhornCreek	10080016	0.000693	0.025	0.00151	0.000142
EmeraldCreek	10080001	0.000996	0.189	0.0017	0.000286
FallCreek	17040104	0.001267	0.692	0.001809	0.000103
FlatCk	17040103	0.001011	0.239	0.001925	0.000373
FrontierCreek	10080001	0.001249	0.654	0.001967	0.000351
GeorgeCreek	17040210	0.001251	0.660	0.002059	0.00062
GibsonJackCreekPortneufRiver	17040208	0.002028	0.925	0.003353	0.000275

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
GlenwoodFishHatchery	NA	0.001206	0.623	0.002518	4.90E-05
GooseCreekSnakeRiver	17040211	0.002165	0.956	0.003464	0.000122
GraniteCk	17040103	0.001178	0.585	0.002322	0.000206
GrapeCreekRaftRiver	17040210	0.00098	0.145	0.001617	0.00022
GraveCreek	10080002	0.001284	0.704	0.00218	0.000432
GraveLake	10080002	0.001568	NA	0.002315	0.000424
GravelCreek	17040101	0.002337	0.987	0.003455	0.000399
GreybullRiver	10080009	0.001185	0.604	0.002058	0.00023
GrizzlyBearLake	17040101	0.001153	NA	0.002457	0.000185
GrosVentrelakes	17040102	0.001294	0.711	0.00269	0.000375
GrouseCreek	17040101	0.001559	0.824	0.003083	0.001467
HarknessCreekPortneufRiver	17040208	0.00212	0.950	0.003944	0.000191
HatchetLake	10080001	0.001805	NA	0.002131	0.000731
HeartRiver	17040101	0.000991	0.176	0.001872	0.000398
HenrysLakeEggCollectionFacility	NA	0.001133	NA	0.002027	NA
HiddenLake	10080001	0.001761	NA	0.003286	0.000241
HighMeadowCreek	10080003	0.001379	0.761	0.00191	0.000527
HighMeadowLake	10080003	0.001106	NA	0.002197	0.000271
HoneymoonLake	10080001	0.001306	NA	0.002736	0.000125
HorseshoeLake	10080001	0.002751	NA	0.004099	0.00139
HunterLake	10080012	0.001055	NA	0.00189	0.000356
IshawooaCreek	10080013	0.002083	0.937	0.003417	0.000197
IslandLake	10080003	0.001216	NA	0.00208	0.000471
JacksonNationalFishHatchery	NA	0.001928	0.899	0.003257	6.05E-05

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
KagevahLake	10080002	0.001166	NA	0.001795	0.000314
Lake24	10080002	0.001681	NA	0.002484	0.000364
Lake26	10080001	0.001546	NA	0.002485	0.000313
Lake47a	10080002	0.001613	NA	0.00245	0.000413
Lake74	10080001	0.001244	NA	0.001987	0.000339
Lake84	10080001	0.001328	NA	0.002198	0.000252
Lake85	10080001	0.001381	NA	0.002144	0.000266
Lake87a	10080001	0.001636	NA	0.002343	0.000396
Lake87	10080001	0.001581	NA	0.002237	0.000504
Lake99	10080001	0.001403	NA	0.002144	0.000394
LakeSolitude	17040101	0.00104	NA	0.002057	0.000312
LamarRiver	10070001	0.001106	0.453	0.001782	0.000282
LeidyCreek	17040101	0.001084	0.428	0.002333	0.000332
LittleLamar	10070001	0.000738	0.031	0.001494	0.000221
LittleRockyCreek	10070005	0.000964	0.126	0.002036	0.000172
LittleWarmSpringsCreek	10080001	0.001051	0.321	0.002183	0.00034
LonesomeLake	10080003	0.001509	NA	0.002622	0.000681
LowerBarBC	17040102	0.001155	0.560	0.002252	0.000335
LowerBlackJoeLake	10080003	0.001074	NA	0.002157	0.000454
LowerFishCk	17040103	0.001183	0.591	0.002222	0.000392
LowerFrontierCreek	10080001	0.000964	0.132	0.001673	0.0003
LowerGlacierLake	10080001	0.000971	NA	0.001429	0.000628
LowerIceLake	10080002	0.001837	NA	0.002954	0.000303
LowerLakeCreek	10080008	0.001306	NA	0.002175	0.000258

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
LowerWigginsForkCreek	10080001	0.001072	0.384	0.001929	0.000363
LydleLake	10080001	0.001244	NA	0.002071	0.000305
MainBasinYSL	10070001	0.001141	0.528	0.001955	0.000275
MarkedTreeLake	10080001	0.002042	NA	0.00358	0.000181
MarquetteCreek	10080013	0.001325	0.742	0.00242	0.000407
MartenCreek	17040103	0.001236	0.642	0.002301	0.00045
MeeteetseCreek	10080009	0.001074	0.403	0.001492	0.000282
MiddleDeepCreek	10080003	0.002018	0.918	0.003121	0.000303
MiddleForkDryCreek	10080001	0.001644	0.843	0.002373	0.000562
MiddleWigginsForkCreek	10080001	0.001019	0.245	0.001678	0.000286
MillCreek	10080008	0.000933	0.107	0.00187	0.000234
MinerCr	10070002	0.001185	0.597	0.002168	0.000273
MooseLake	10080001	0.001871	NA	0.002853	0.00047
MuddyCreek	10070006	0.001257	0.679	0.002263	0.000273
NBeaverCreek	10080010	0.001024	0.252	0.002041	0.000167
NBuffaloFk	17040101	0.001049	0.302	0.001923	0.000415
NFSMiddleCreek	10080012	0.002055	0.931	0.003324	0.000589
NFSTroutCreek	10080012	0.001395	0.780	0.002336	0.000361
NorthYSL	10070001	0.001063	0.358	0.001849	0.000271
NorthForkPopoAgieRiver	10080003	0.001818	0.881	0.002238	0.000673
NorthLeighCreek	17040204	0.001254	0.673	0.002073	0.000358
OnemileCreek	17040210	0.000889	0.088	0.001724	0.000397
PacificCreekaboveMink	17040101	0.001066	0.377	0.001724	0.000434
PalisadesCreek	17040104	0.002215	0.962	0.003713	0.000159

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
PalisadesCreekTributaries	17040104	0.001314	0.723	0.002661	0.000427
PickettCreek	10080009	0.001039	0.283	0.001848	0.000276
PilotPeakHatcheryLCT	NA	0.000734	NA	0.00137	0.000212
PineCreek	17040104	0.001031	0.270	0.001926	0.000348
PineyCreek	10080009	0.001055	0.346	0.001886	0.000265
RaftCreek	10080002	0.001198	0.616	0.002156	0.000333
RaineyCreek	17040104	0.001129	0.503	0.002	0.000376
RawhideCreek	10080009	0.000831	0.057	0.001727	0.000143
RedGulchCreek	10080016	0.000809	0.050	0.001644	0.000151
RobertsLake	10080002	0.001089	NA	0.001978	0.000314
RobertsLakeDrainage	10080002	0.001117	NA	0.00223	0.000373
RockCreek	17040101	0.001148	0.541	0.002381	0.000338
RossFork	17040206	0.001324	0.736	0.002366	0.000553
SaltRiver	17040105	0.001124	0.497	0.002123	0.000355
SBeaverCreek	10080010	0.000948	0.119	0.002033	0.00017
SFHobackRiver	17040103	0.001065	0.365	0.002126	0.000423
SForkLittleTongueRiver	10090101	0.001002	0.214	0.001983	0.000221
SFOwlCreek	10080007	0.001838	0.887	0.003072	0.000458
ShangralaDinwoodyCreek	10080001	0.001783	0.868	0.002639	0.000555
SheepCreek	17040103	0.001223	0.635	0.002116	0.000385
SheridanCreek	10080001	0.000919	0.094	0.001837	0.00025
SickleCreek	17040101	0.000986	0.157	0.001768	0.000418
SlateCk	17040102	0.001027	0.258	0.001925	0.000373
SloughCreek	10070001	0.001806	0.874	0.002325	0.000352

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
SodaButteCreek	10070001	0.00258	1.000	0.003899	0.000157
SodaForkBuffaloRiver	17040101	0.001967	0.912	0.002788	0.000439
SoldierCreek	10080008	0.001238	0.648	0.002634	0.000296
SouthArmYSL	10070001	0.001079	0.415	0.001913	0.000257
SoutheastArmYSL	10070001	0.001152	0.547	0.002041	0.000236
SouthForkWoodRiver	10080009	0.00087	0.069	0.001399	0.000232
SouthLeighCreekTetonRiver	17040204	0.001081	0.421	0.001901	0.000367
SouthPaintRockCreek	10080008	0.001003	0.220	0.002024	0.000405
SpiderLake	10080002	0.000959	NA	0.001787	0.000215
SpreadCreekditchbymouth	17040101	0.001459	0.805	0.002391	0.000431
SpringCreek	17040105	0.001119	0.491	0.002053	0.000399
SRsidechannel	17040101	0.000995	0.182	0.001871	0.000395
StoneflyCreek	10080001	0.000867	0.063	0.001817	0.000156
StoryHatcheryRBT	NA	0.002268	NA	0.003162	0.00169
StoughCreek	10080003	0.002335	0.981	0.003437	0.000367
SurpriseCreek	17040101	0.000975	0.138	0.001748	0.000392
TensleepYCT	NA	0.001009	0.233	0.001989	8.66E-05
ThorofareCreek	10070001	0.002545	0.994	0.004275	0.000106
ThreeChannelSpring	17040101	0.001002	0.208	0.001853	0.000343
TimberCreek	10080009	0.000927	0.101	0.001757	0.000261
TrapperCreek	10080010	0.001055	0.340	0.002288	0.000166
TroutCreekBighornLake	10080010	0.001219	NA	0.001776	0.00034
TroutCreek	10080010	0.000763	0.044	0.001173	0.000168
TwinLake	10080001	0.001517	NA	0.002307	0.000485

population	huc8	thetaw	thetaw quantile	thetapi	obs_het
UnnamedGravesLakes	10080002	0.001418	NA	0.002348	0.000392
UpperBarBC	17040101	0.001119	0.484	0.002093	0.000406
UpperFishCk	17040103	0.001053	0.333	0.001974	0.000338
UpperFrontierCreek	10080001	0.00113	0.509	0.001755	0.000302
UpperRossLake	10080001	0.001626	NA	0.001981	0.000455
UpperSFShoshoneRiver	10080013	0.001481	0.811	0.002496	0.000207
UpperTincupCreek	17040105	0.001252	0.667	0.002306	0.000423
UpperWigginsForkCreek	10080001	0.001043	0.296	0.001731	0.000287
UpperWoodRiver	10080009	0.001381	0.767	0.002435	0.000254
WallaceCreek	10070002	0.000876	0.075	0.001576	0.000304
WashakieReservoir	10080002	0.001157	0.566	0.002103	0.000328
WashoeHatchery	NA	0.001303	NA	0.002203	0.000468
WatkinsLake	10080001	0.000885	NA	0.001728	0.000281
WestForkGreysRiver	17040103	0.001076	0.409	0.001805	0.00031
WestThumb	10070001	0.001132	0.516	0.001989	0.000237
WhetstoneCreek	17040101	0.001265	0.686	0.001903	0.000405
WigwamBRC	NA	0.00105	0.314	0.00193	5.66E-05
WillowCreek	17040205	0.001114	0.472	0.00198	0.000304
WittombonaLake	10080001	0.002353	NA	0.002367	0.000752
WoodRiver	10080009	0.001106	0.447	0.00205	0.000252
WyomingCreek	17040202	0.001189	0.610	0.001929	0.000646

2.7 Appendix

Discussion of specific populations

Admixture from Westslope Cutthroat Trout in Lodgegrass Creek (Big Horn basin) and Lahontan Cutthroat Trout in Pumpkin Creek (Big Horn basin) and Hope Creek (Upper Snake basin) were surprising results. Yellowstone River ancestry in the Snake Headwaters basin was almost entirely confined to several small populations in Grand Teton National Park (Death Canyon, Grizzly Bear Lake, and Lake Solitude) that had previously been identified as having anomalously large spots and Snake River mitochondrial haplotypes (M. A. Novak, Kershner, and Mock 2005).

Within the North Fork Popo Agie drainage, which includes Lonesome Lake, the mainstem North Fork Popo Agie River, and Papoose Lake (also referred to as Dog Tooth Lake or Two Dog Tooth Lake in some historical documents), some uncovered ancestry was expected whereas others was surprising. The presences of Snake River (stocked in 1986) and Yellowstone River ancestry, were expected. However, the presence of Colorado River Cutthroat Trout ancestry was surprising. The absence of native ancestry within the upper North Fork Popo Agie drainage indicates that a hypothesized waterfall barrier within North Fork Canyon is likely a true barrier to upstream fish movement from lower portions of the North Fork Popo Agie and mainstem Popo Agie rivers.

Colorado River ancestry within the North Fork Popo Agie drainage could be present due to several factors. A stocking record from D.C. Booth National Fish Hatchery indicates stocking of “black-spotted trout” in the “Big Popo Agie headwaters” in 1913; however, it is unknown if the record refers to the Middle Fork or North Fork Popo Agie. These fish were also most likely Yellowstone Cutthroat Trout. An unnamed cutthroat trout lineage from Jackson National Fish Hatchery was stocked in the upper drainage in 1976, but these were likely Snake River Cutthroat Trout (Sean Henderson, Jackson National Fish Hatchery, personal communication). A more likely source of Colorado River Cutthroat Trout within

the North Fork Popo Agie drainage is stocking from early outfitters. For example, outfitter Finis Mitchell transplanted Colorado River cutthroat trout from Big Sandy Lake to multiple fishless lakes in 1930 (Udall 1986). It is possible that Lonesome Lake, which is across the continental divide but only 2.5 trail miles from Big Sandy Lake, was one of the lakes stocked. Additionally, Mitchell stated that he transplanted cutthroat trout of unknown lineage from Lonesome Lake to other North Fork Popo Agie waters, including Bear, Lizard Head (likely East Lizard Head), and Two Dog Tooth (likely Papoose) lakes in 1931 (Wyoming Game and Fish Department, unpublished interview, 1972).

Interestingly, the ancestry of cutthroat trout within Black Joe Lake was remarkably similar to those within North Fork Popo Agie drainage waters, with Yellowstone, Snake River, and Colorado River ancestry present. Black Joe Lake lies within the Colorado River basin. According to Finis Mitchell, Colorado River Cutthroat Trout existed in Black Joe Lake in 1930 (Mitchell 1985). Despite being across the continental divide from the Popo Agie watershed, Black Joe Lake is only 4 trail miles from Lonesome Lake, making it relatively easy to transplant fish from one basin to another.

Within the Wiggins Fork, genetic ancestry closely matches historical stocking records. In Emerald and Burwell (a tributary to Emerald) creeks, there are no historical stocking records: one unpublished Wyoming Game and Fish Department card file shows a note that simply states “no stocking at present” for Emerald Creek in 1953. However, moving downstream to the Upper Wiggins Fork site, which lies near the confluence of Emerald Creek and the Wiggins Fork, stocking records exist as far back as 1933. These stocking records do not specify which lineages of cutthroat trout were stocked, though they start to differentiate between Snake River Cutthroat Trout and Yellowstone Cutthroat Trout in 1959, partially implying that either (or both) of these strains were being stocked prior to that year. This could explain the appearance of limited Snake River ancestry here and in Emerald Creek. The same is true for the Middle Wiggins Fork site, which lies just above the confluence with Frontier Creek, and was considered the same site as the Upper Wiggins

Fork in historical stocking records. Interestingly, records indicate that stocking at these upstream sites largely ceased in the late 1940s, when managers transitioned to stocking fish primarily downstream of Double Cabin. Between 1962 and 1990 Snake River cutthroat trout were stocked across 22 different years at multiple locations in the Wiggins Fork downstream of Double Cabin. The increased downstream stocking pressure is reflected in the ancestry of the Lower Wiggins Fork and nearby tributaries: Caldwell, Bug, and Frontier creeks exhibit higher proportions of Snake River ancestry.

Populations in the Portneuf River showed ancestry from both Bear River and Bonneville Basin sources. This is consistent with previous hypotheses regarding temporary connectivity between the Bonneville Basin and the Snake River during the Bonneville flood (Matthew R. Campbell et al. [2011](#)), hence why these ancestries are denoted as “native”.

2.8 Additional figures



Figure 2.7: Results from entropy models to identify admixture from other species. Only populations with ancestry from other species are shown, but all individuals from each of those populations are included.

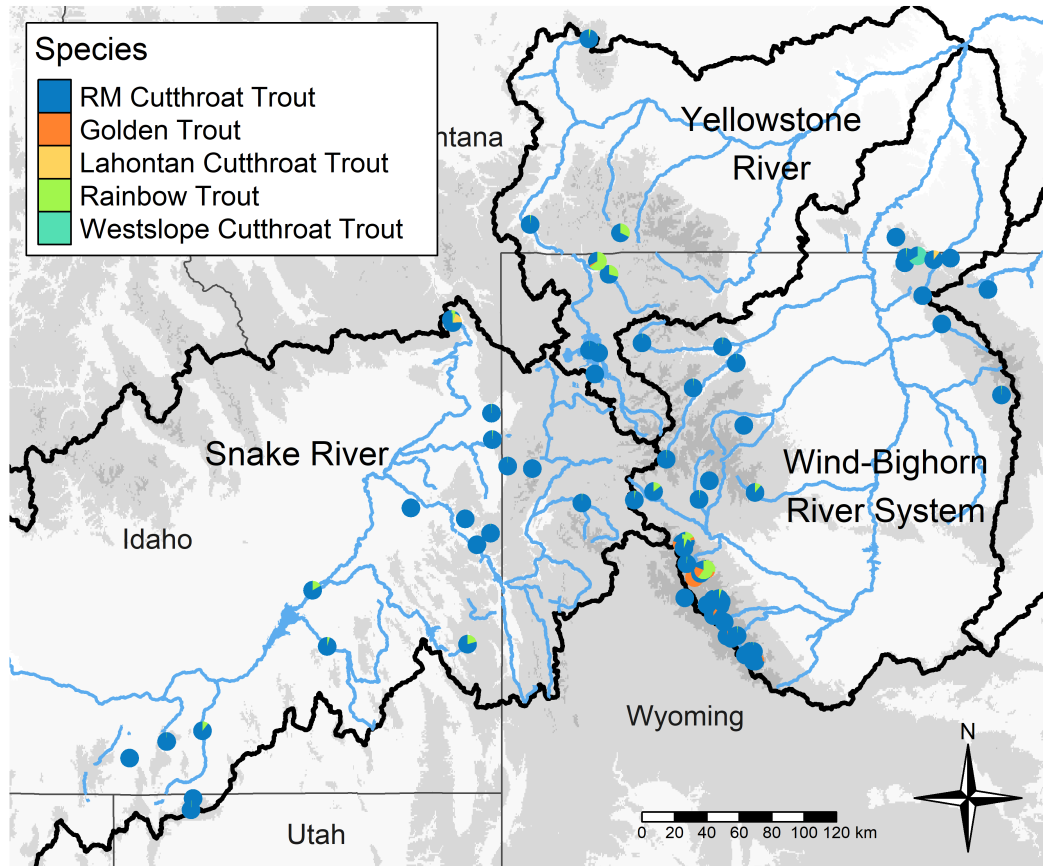


Figure 2.8: Results from the entropy models to identify interspecific hybridization, shown on a map with sample locations. Each pie chart shows the average ancestry proportions for all individuals in that population. Only populations with individuals showing evidence for interspecific ancestry ($\geq 1.5\%$) are shown.

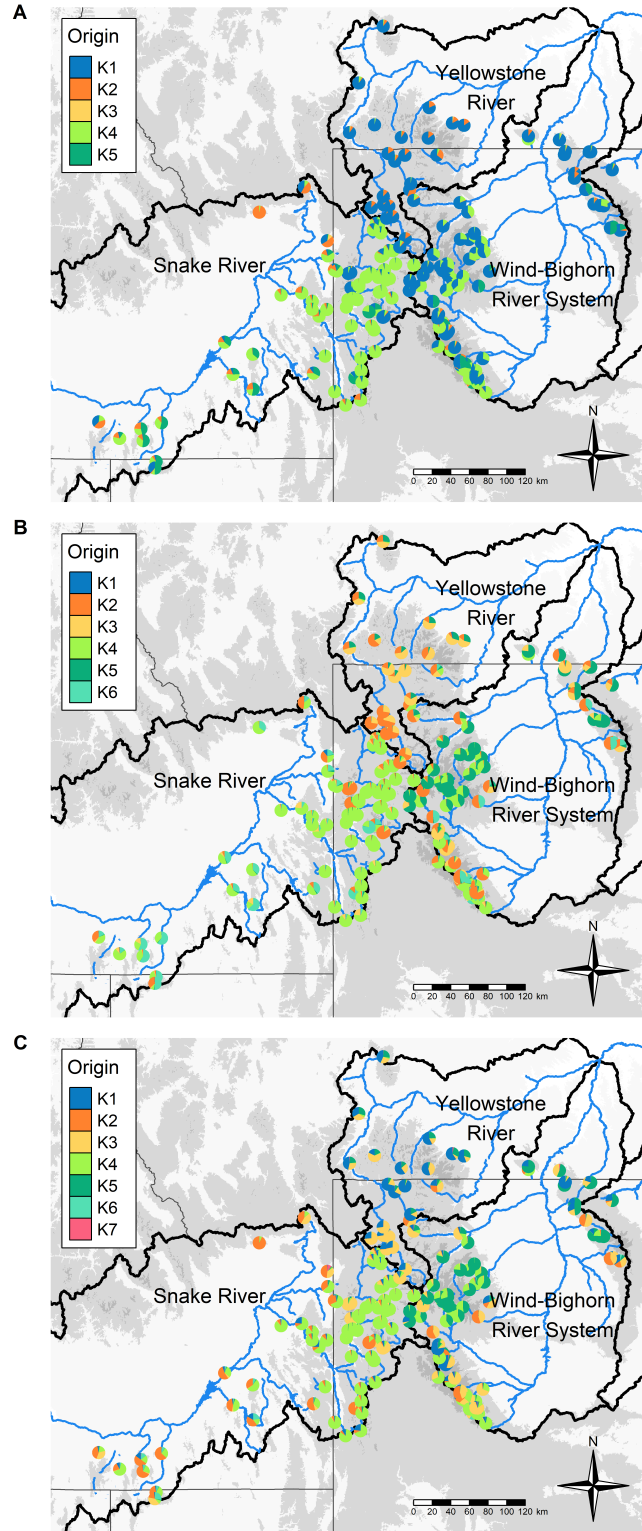


Figure 2.9: Results from admixture models with $K=5$ (A), $K=6$ (B), and $K=7$ (C), shown on a map with sample locations. Each pie chart shows the average ancestry proportions for all individuals in that population.

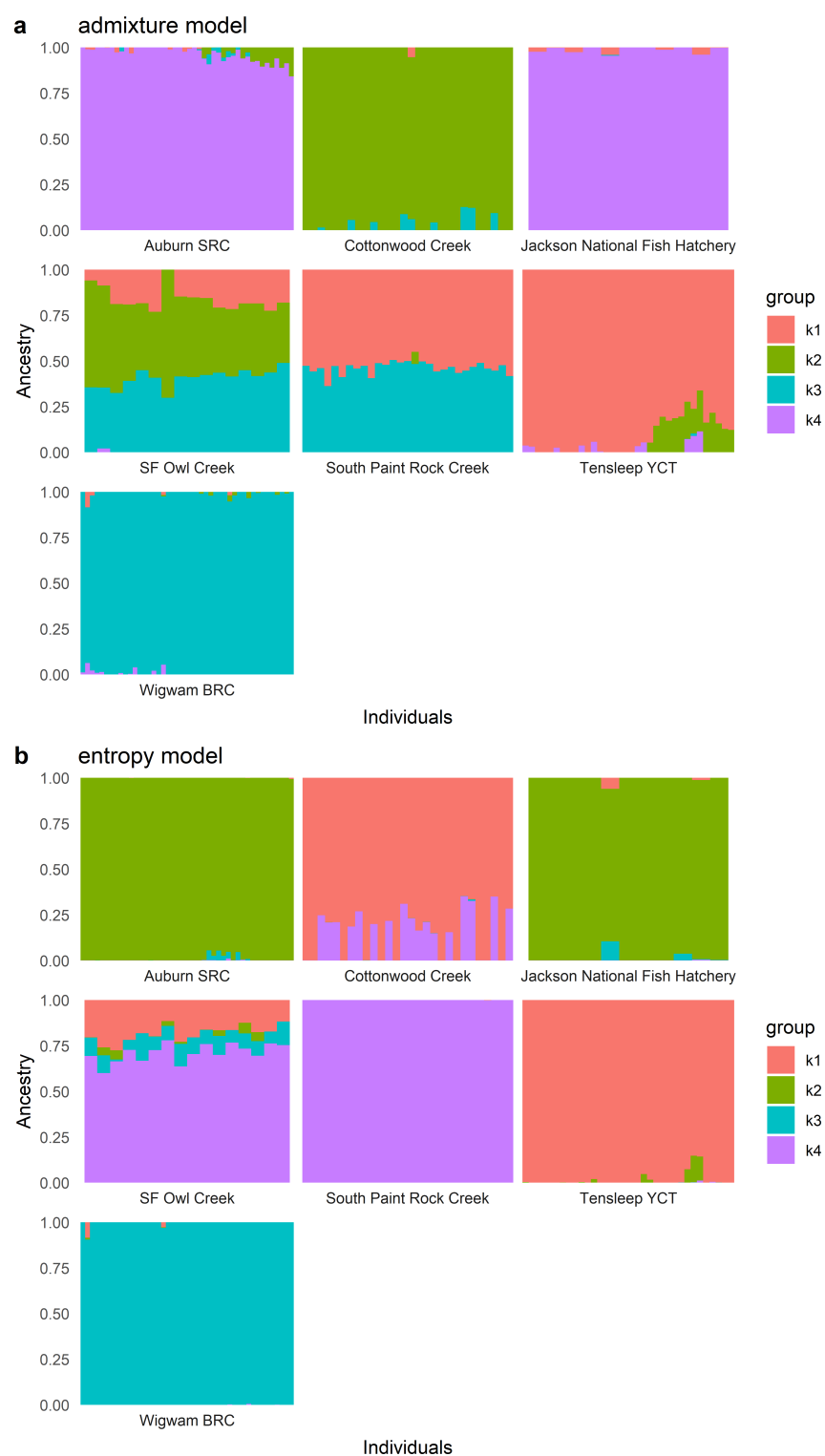
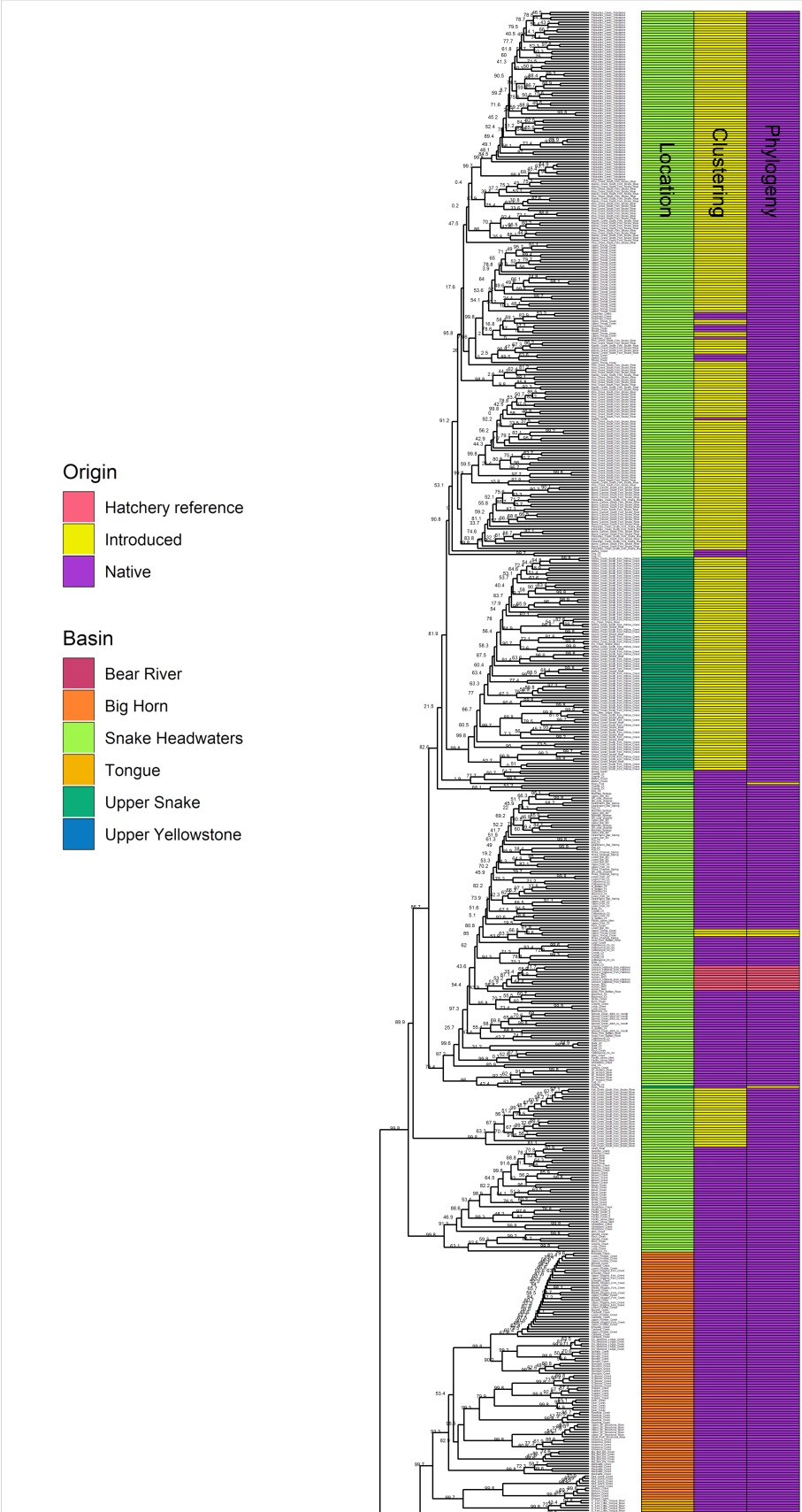


Figure 2.10: Results for the admixture and entropy models that did not agree about cluster assignment. A) admixture model results and b) entropy model results.



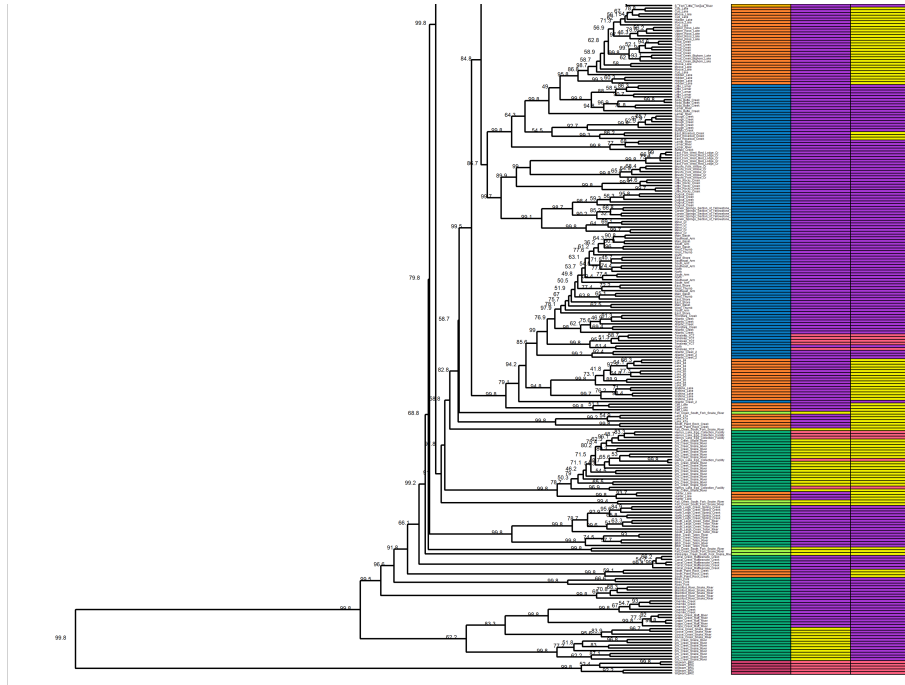


Figure 2.10: A ultrametric tree of *O. vigrinalis* populations using the berst-fit IQTree topology. Tips are labeled by the population of origin and colored by the basin (HUC 6) for that population. Plot is split across 2 pages for ease of visualization.

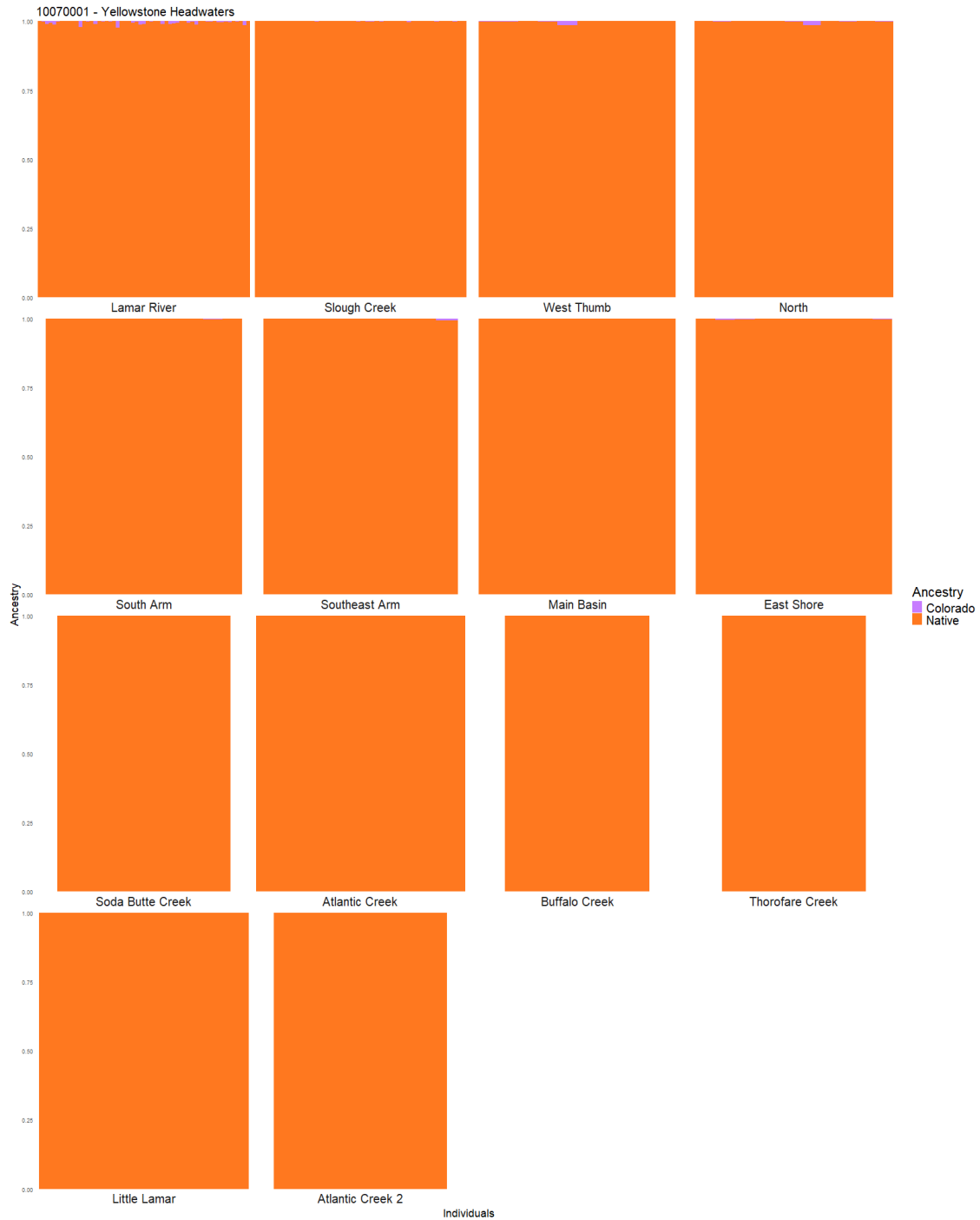


Figure 2.11: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10070001 (Yellowstone Headwaters).

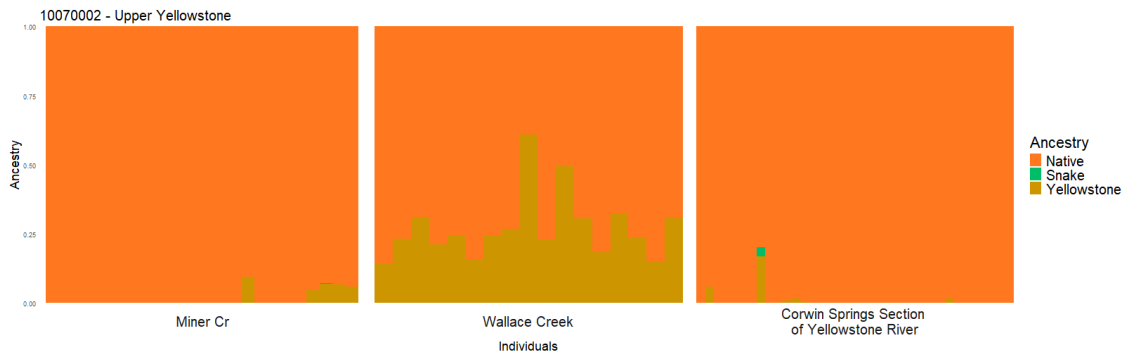


Figure 2.12: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10070002 (Upper Yellowstone).

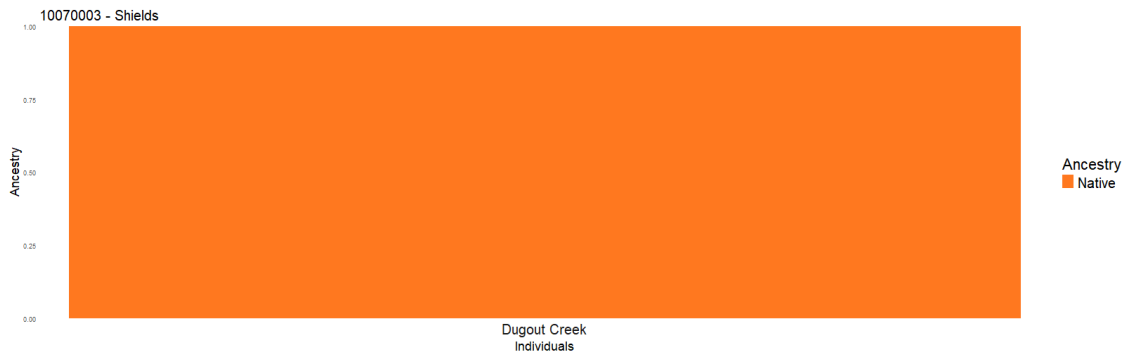


Figure 2.13: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10070003 (Shields).

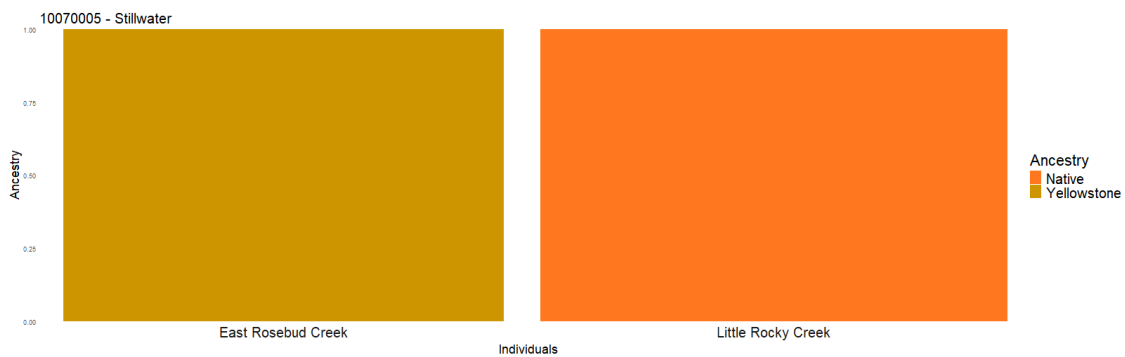


Figure 2.14: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10070005 (Stillwater).

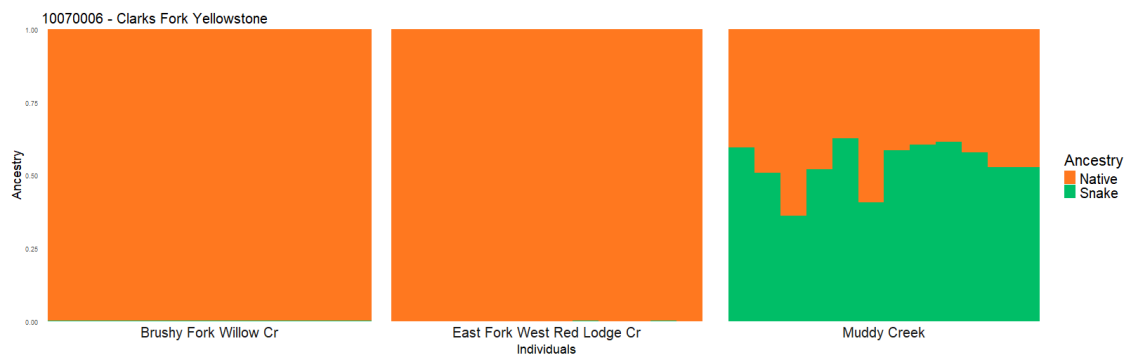


Figure 2.15: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10070006 (Clarks Fork Yellowstone).

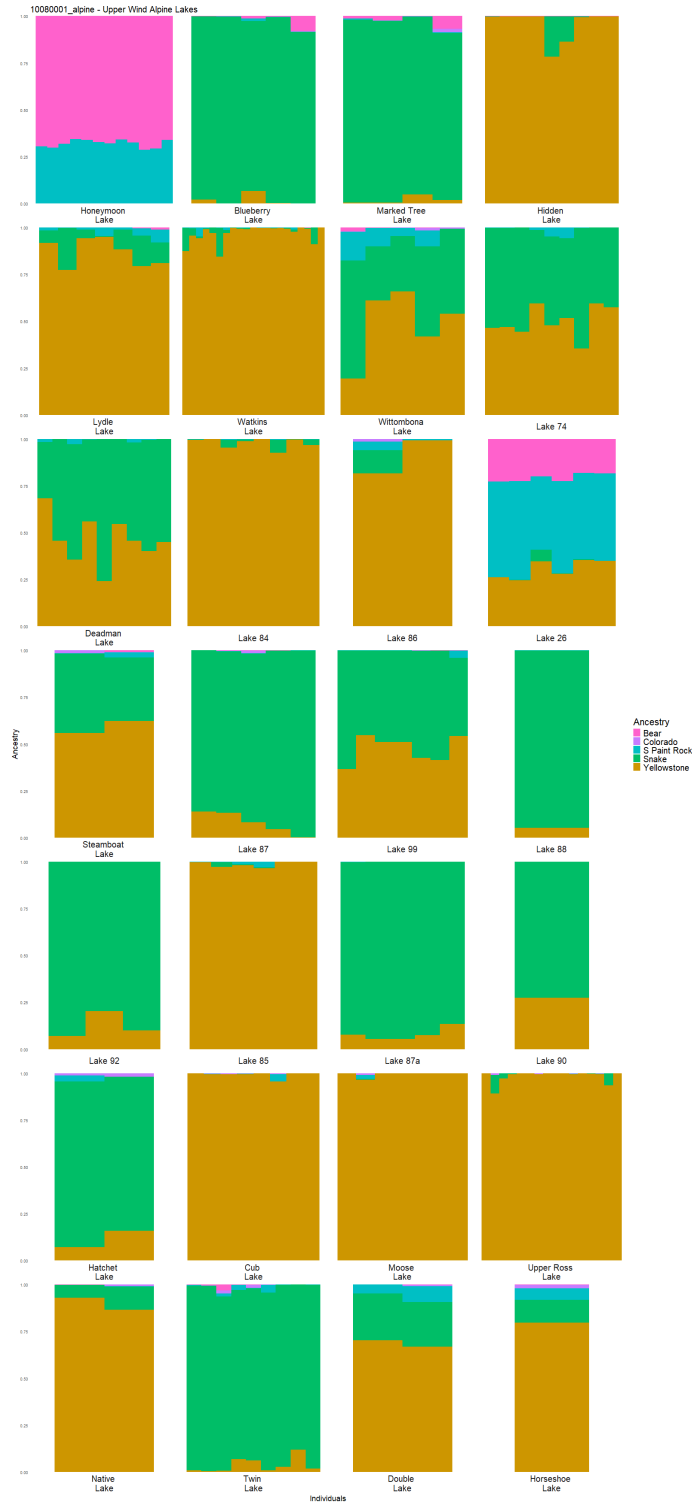


Figure 2.16: Proportion native ancestry per sample, grouped by population, for all alpine lake *O. virginalis* populations within HUC 8 10080001 (Upper Wind).



Figure 2.17: Proportion native ancestry per sample, grouped by population, for all non-alpine-lake *O. virginalis* populations within HUC 8 10080001 (Upper Wind).

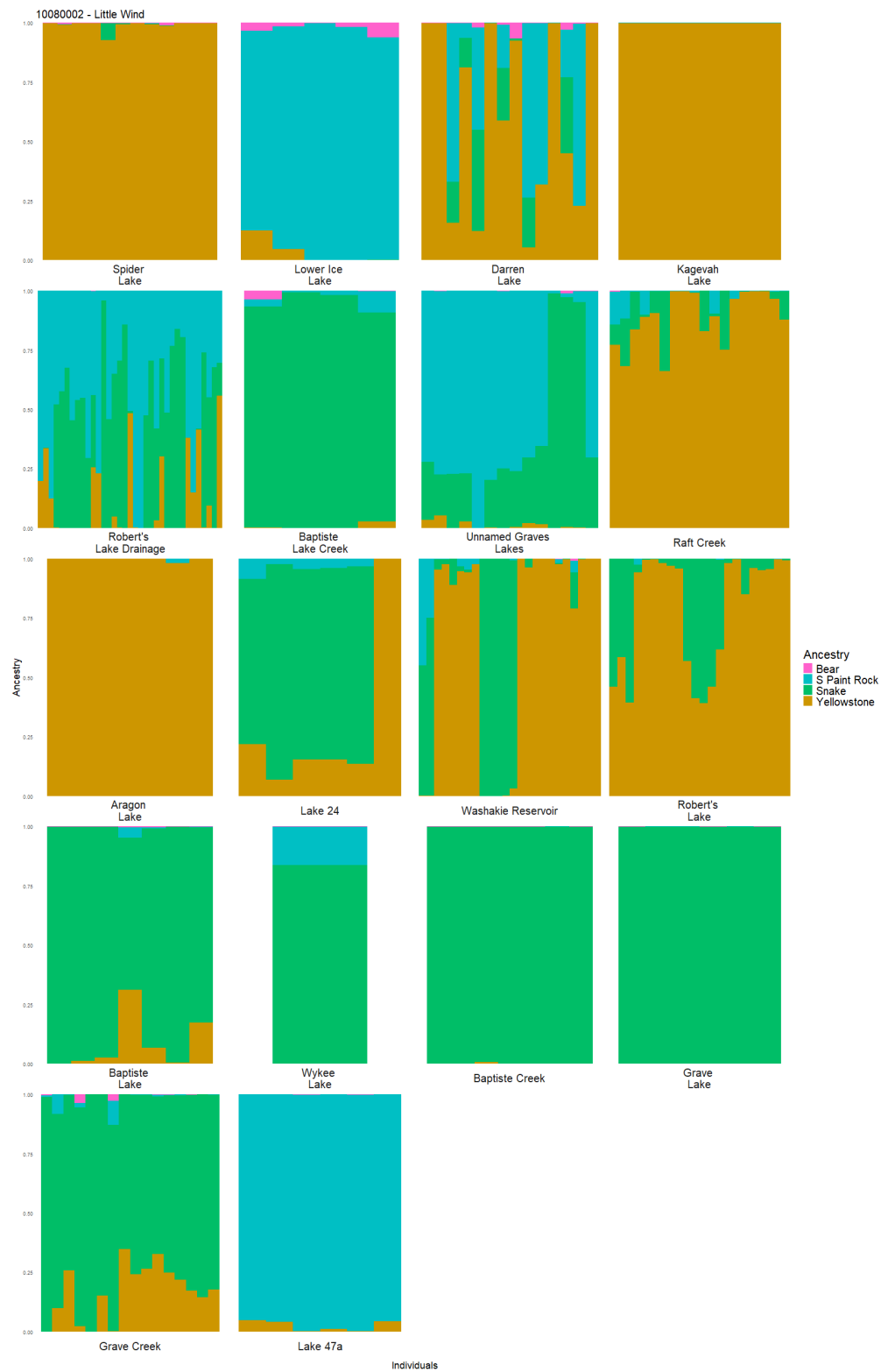


Figure 2.18: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080002 (Little Wind).

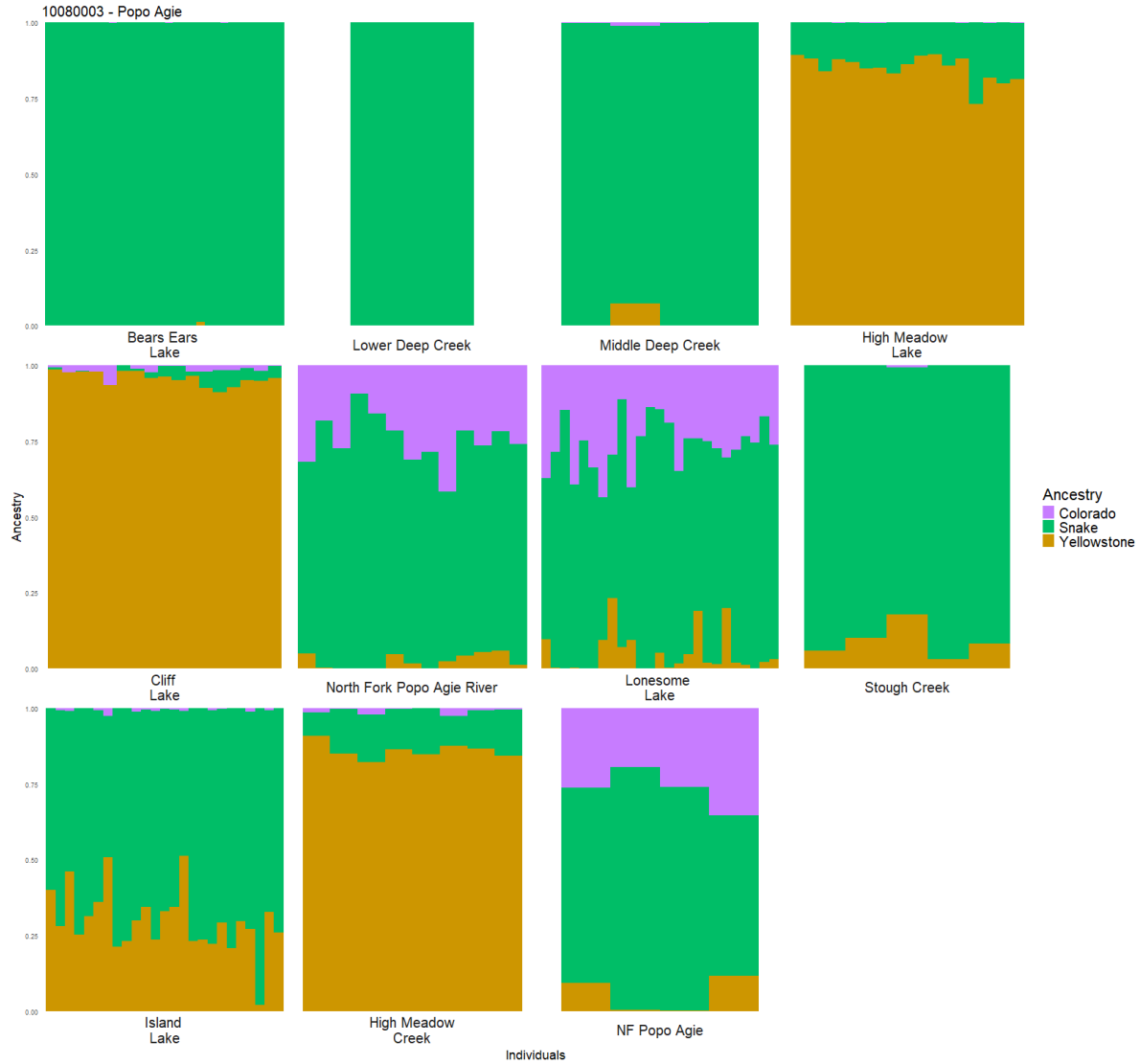


Figure 2.19: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080003 (Popo Agie).

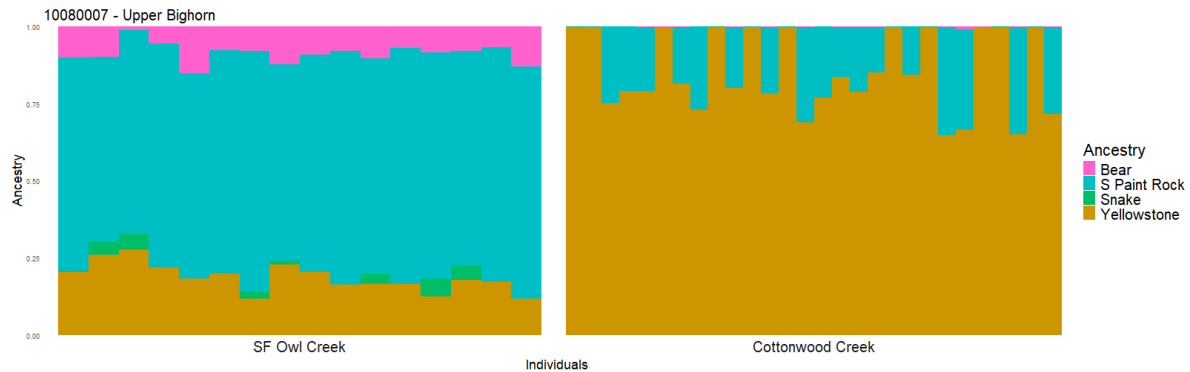


Figure 2.20: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080007 (Upper Bighorn).

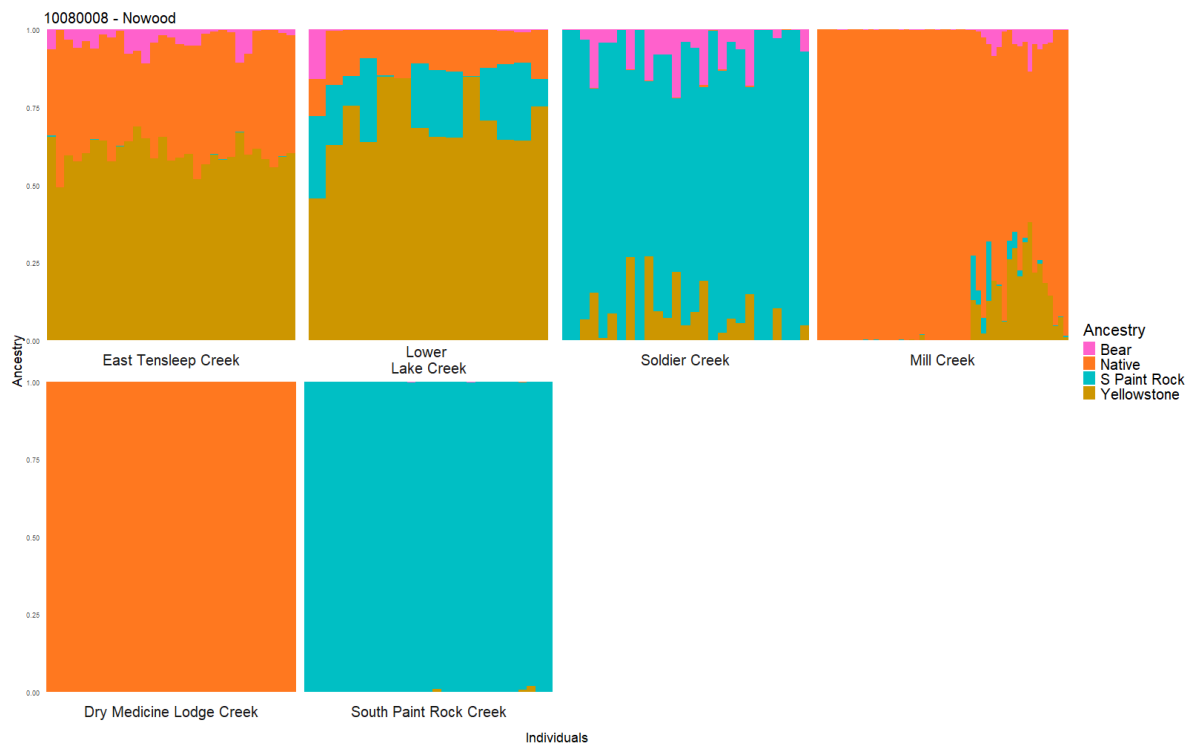


Figure 2.21: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080008 (Nowood).

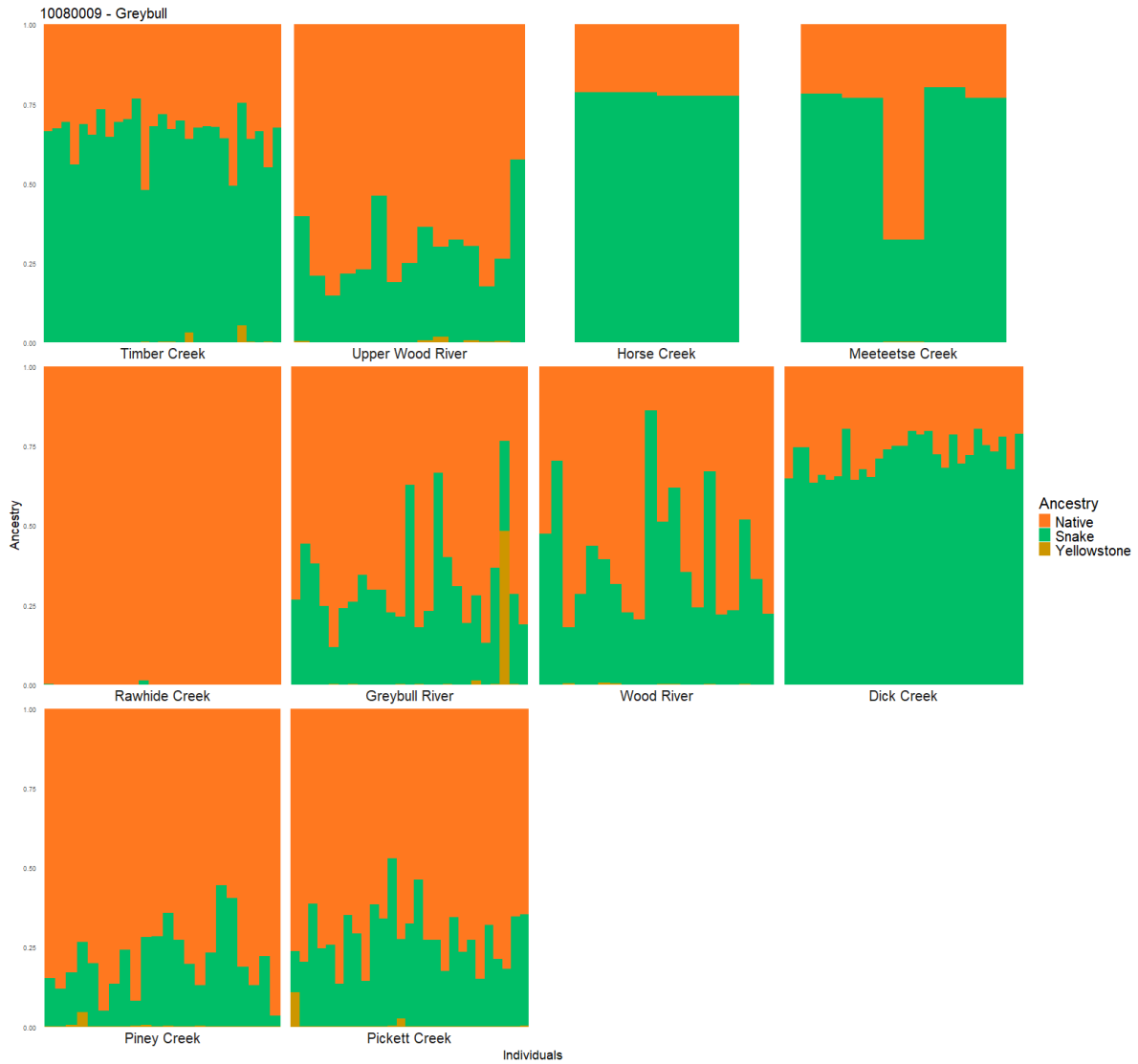


Figure 2.22: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080009 (Greybull).

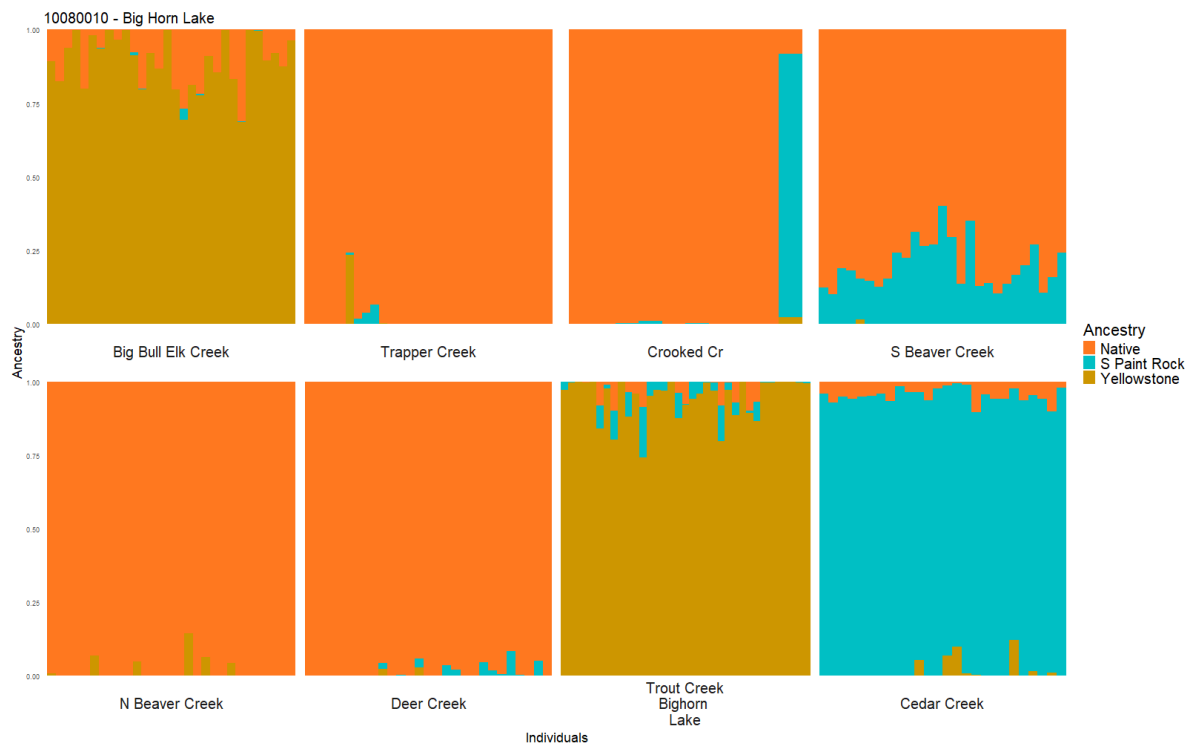


Figure 2.23: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080010 (Bighorn Lake).

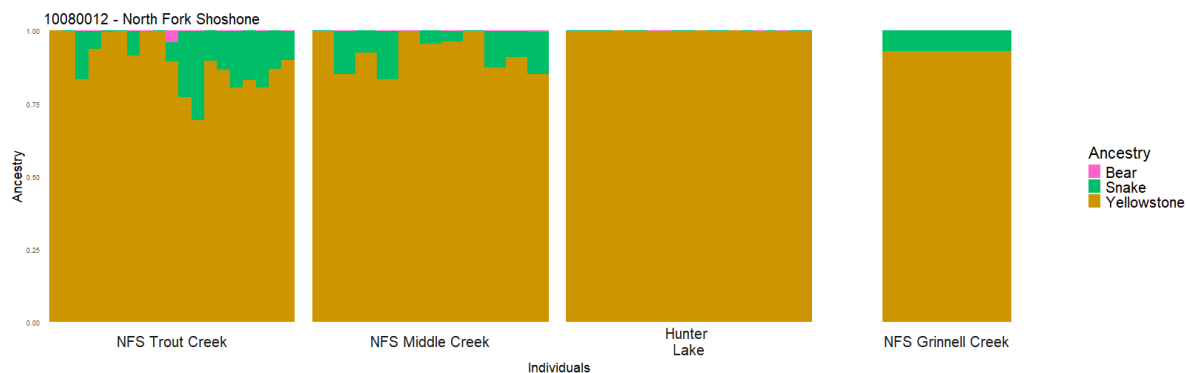


Figure 2.24: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080012 (North Fork Shoshone).

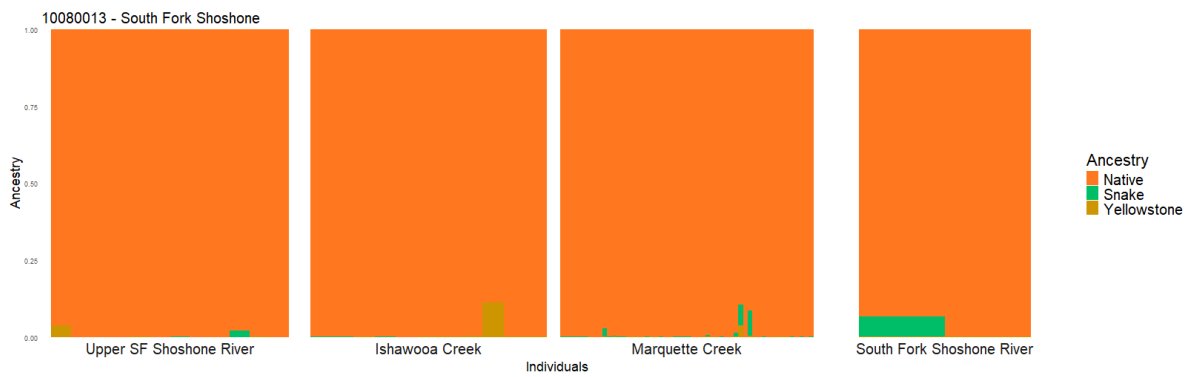


Figure 2.25: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080013 (South Fork Shoshone).



Figure 2.26: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10080016 (Little Bighorn).

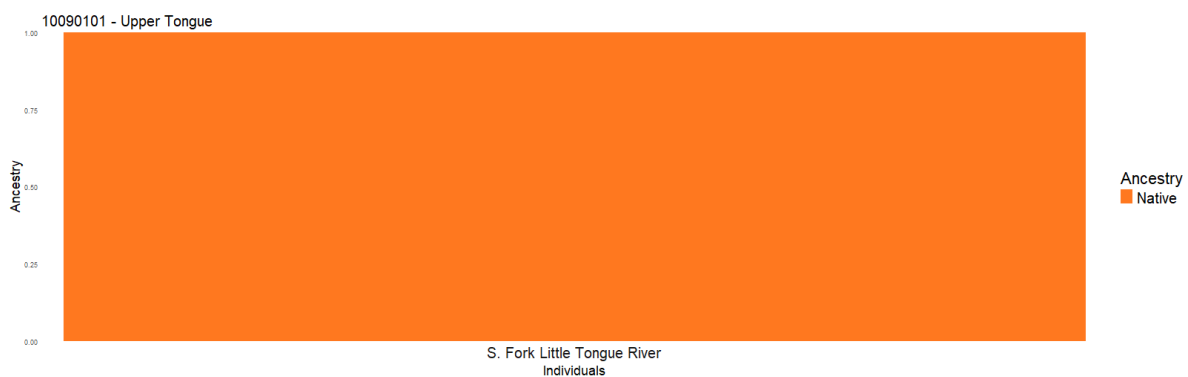


Figure 2.27: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 10090101 (Upper Tongue).

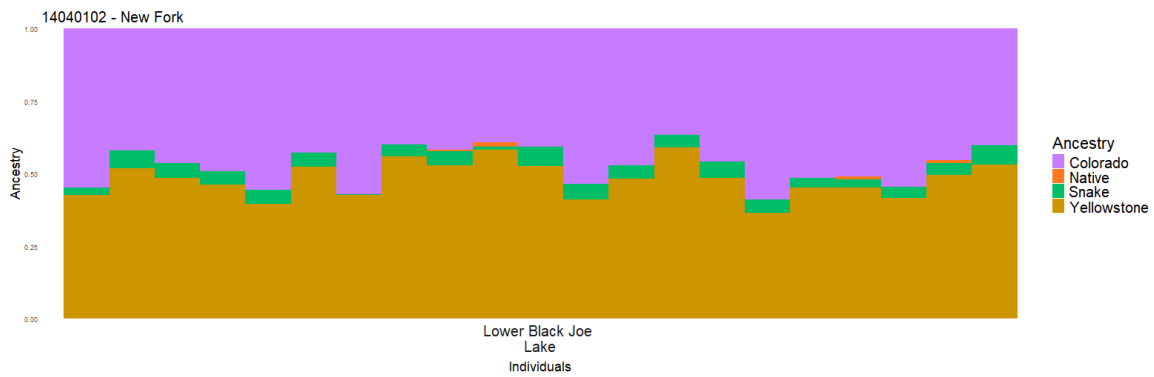


Figure 2.28: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 14040102 (New Fork).



Figure 2.29: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040101 (Snake Headwaters).

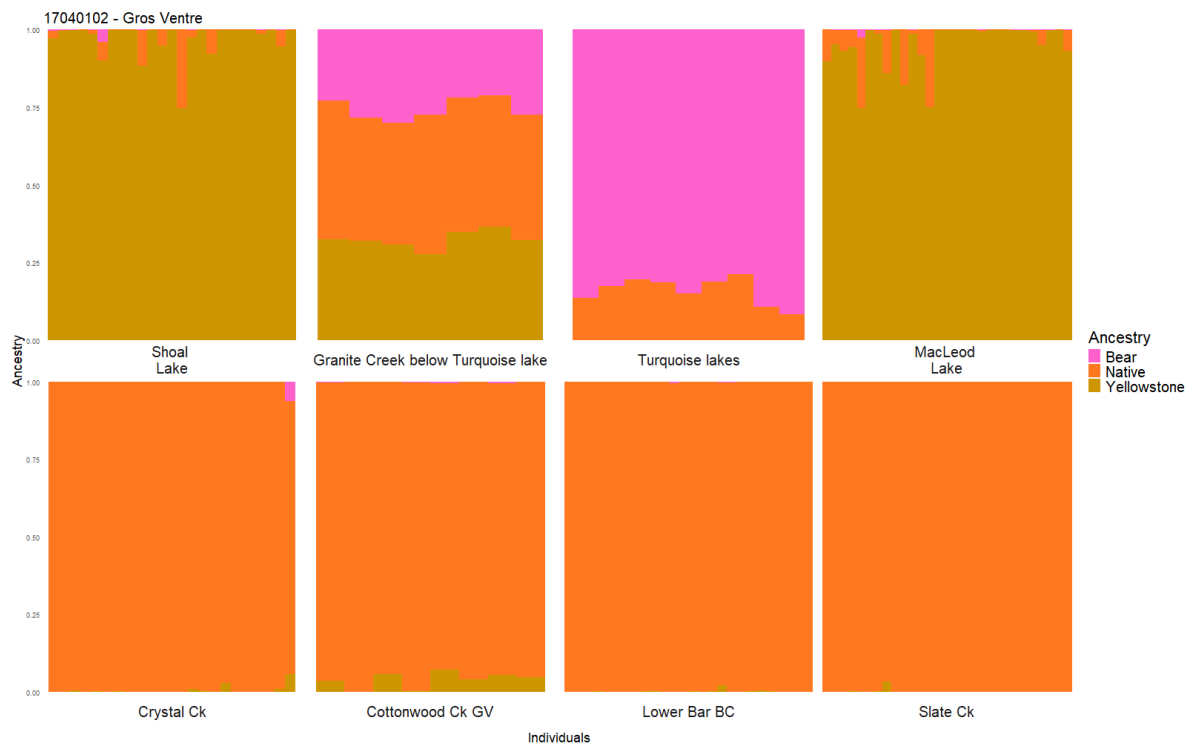


Figure 2.30: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040102 (Gros Ventre).



Figure 2.31: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040103 (Greys-Hoback).

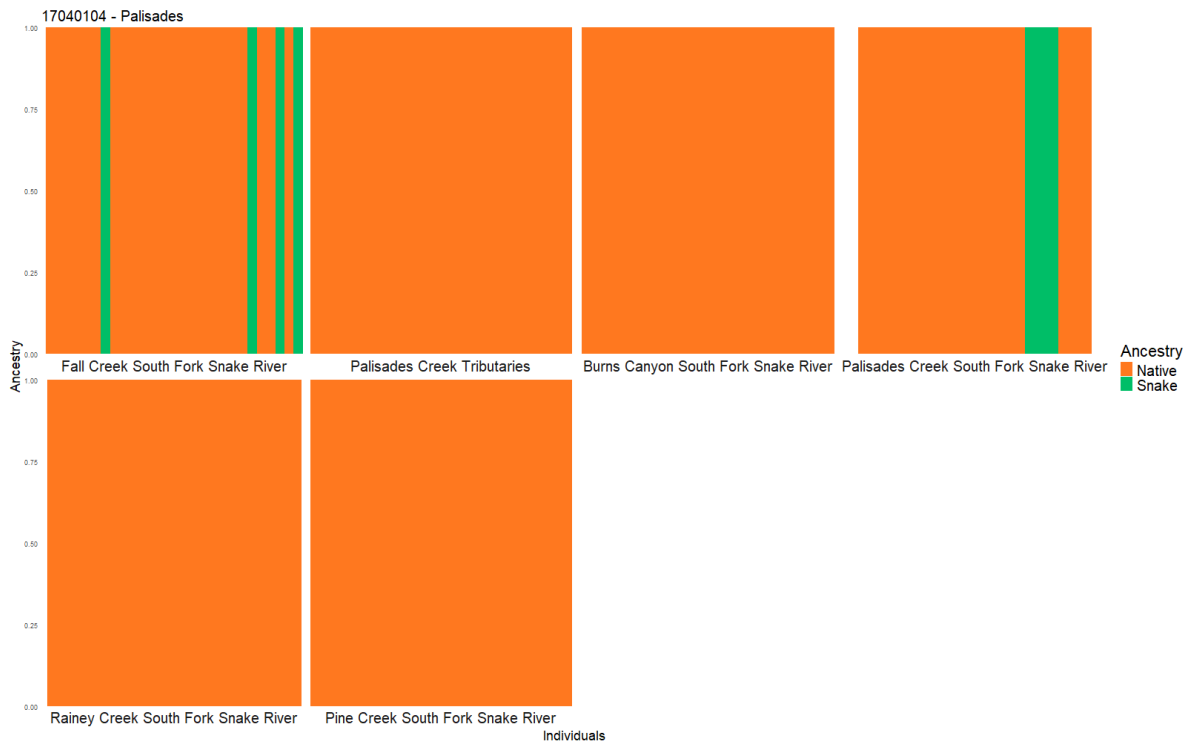


Figure 2.32: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040104 (Palisades).



Figure 2.33: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040105 (Salt).

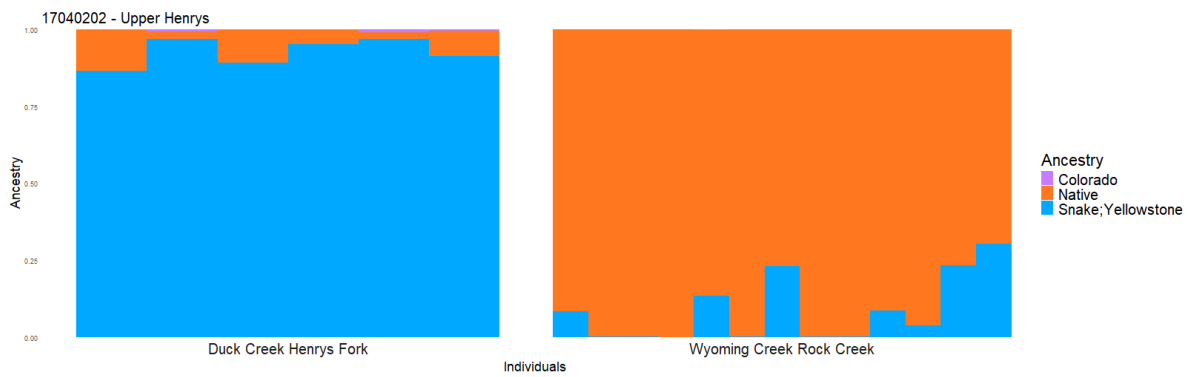


Figure 2.34: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040202 (Upper Henrys).

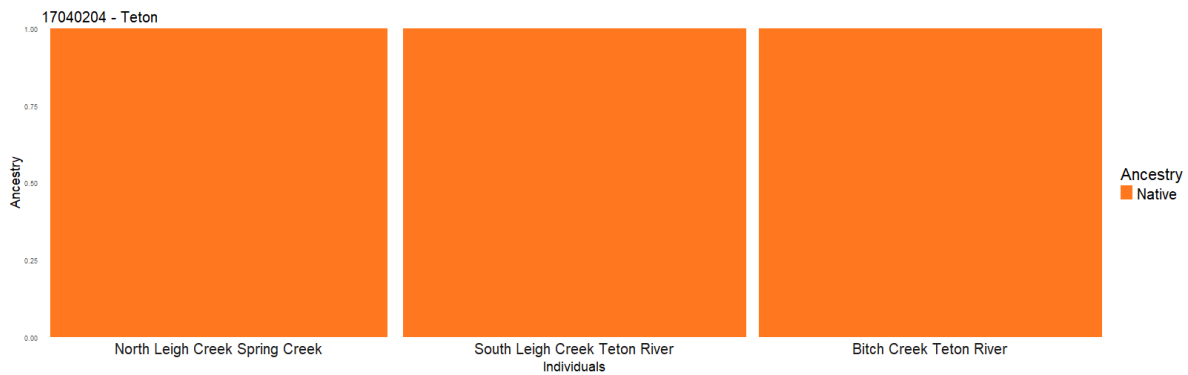


Figure 2.35: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040204 (Teton).

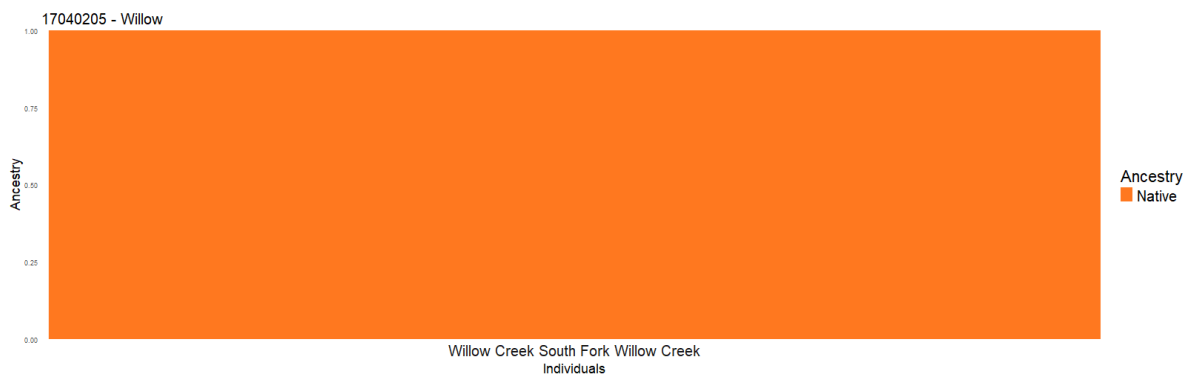


Figure 2.36: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040205 (Willow).

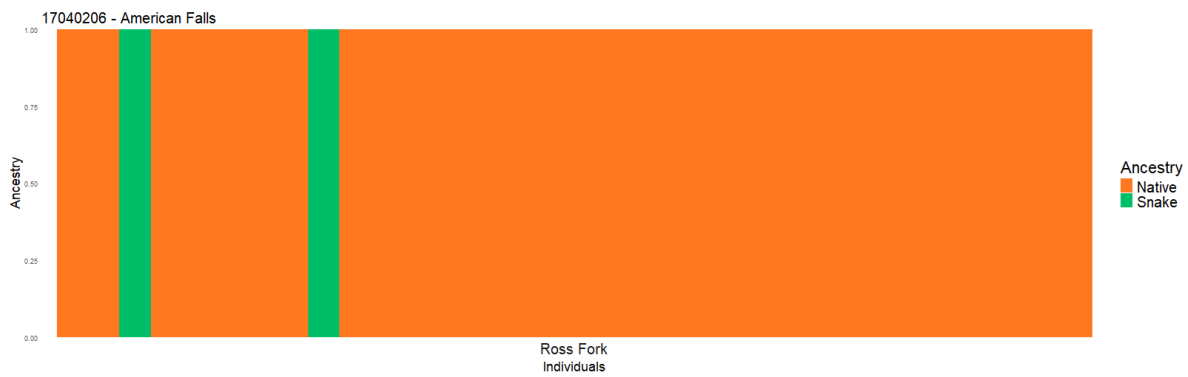


Figure 2.37: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040206 (American Falls).

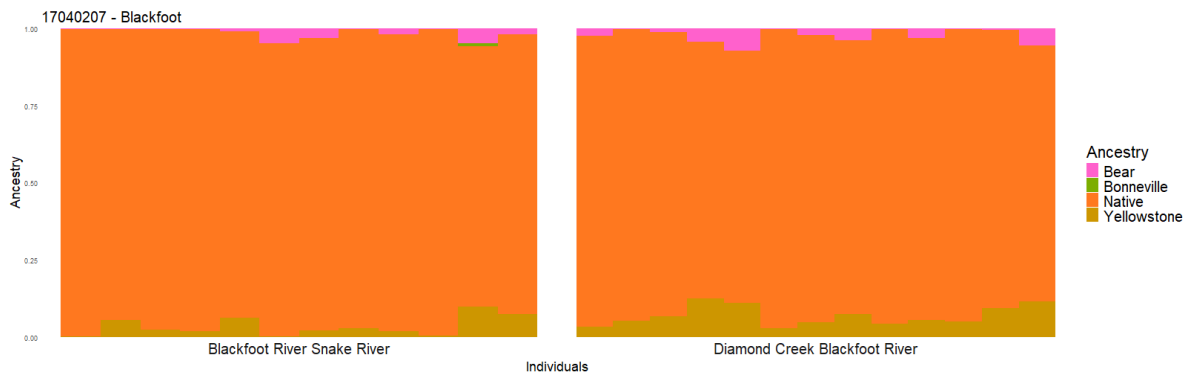


Figure 2.38: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040207 (Blackfoot).

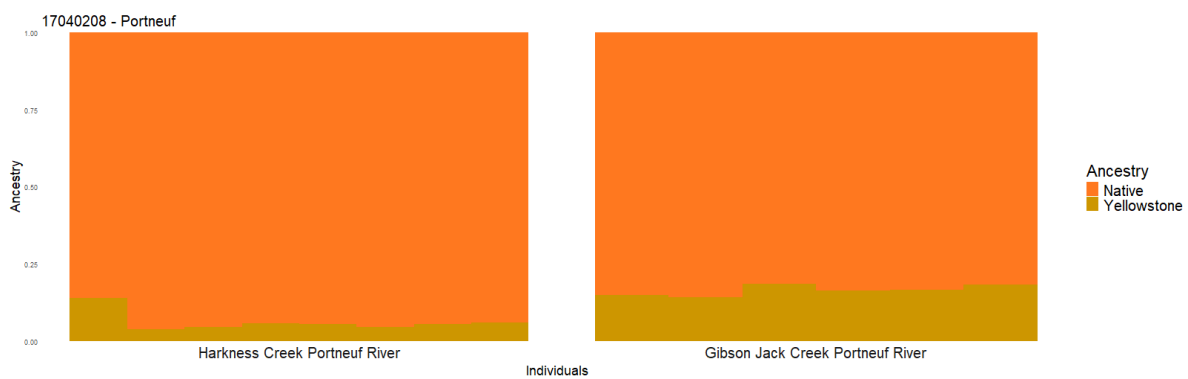


Figure 2.39: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040208 (Portneuf).



Figure 2.40: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040210 (Raft).



Figure 2.41: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040211 (Goose).

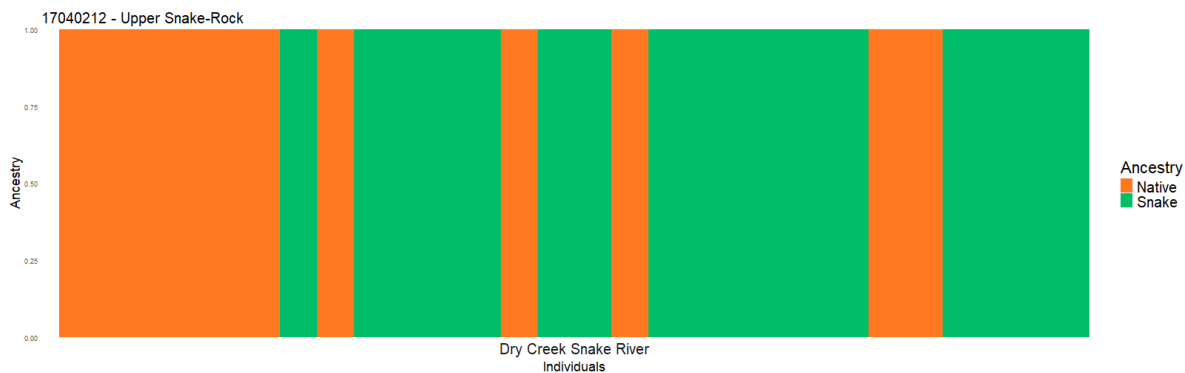


Figure 2.42: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040212 (Upper Snake-Rock).

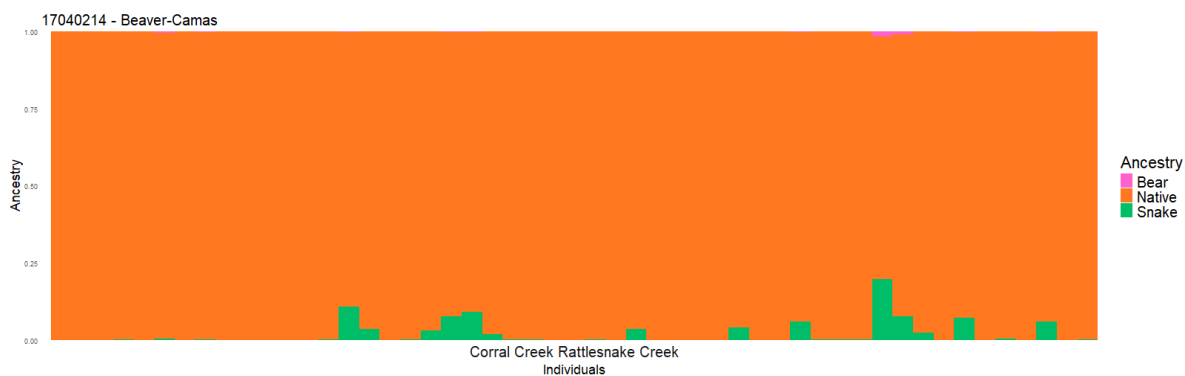


Figure 2.43: Proportion native ancestry per sample, grouped by population, for all *O. virginalis* populations within HUC 8 17040214 (Beaver-Camas).

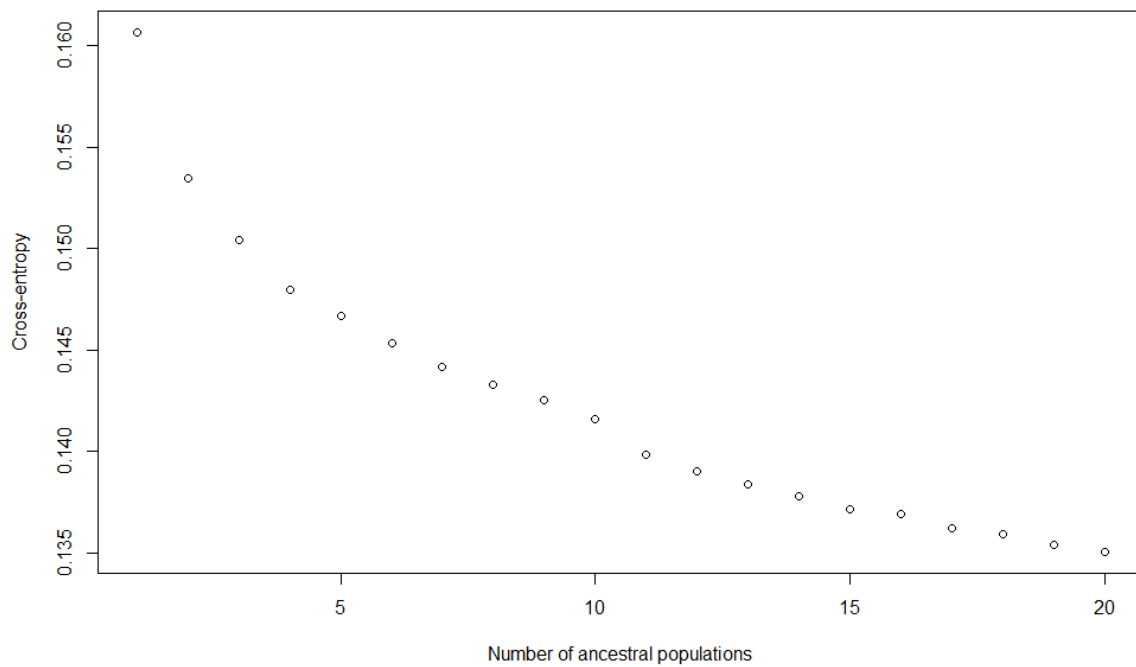


Figure 2.44: Cross-entropy values for snmf analyses run at 1-20 ancestral populations

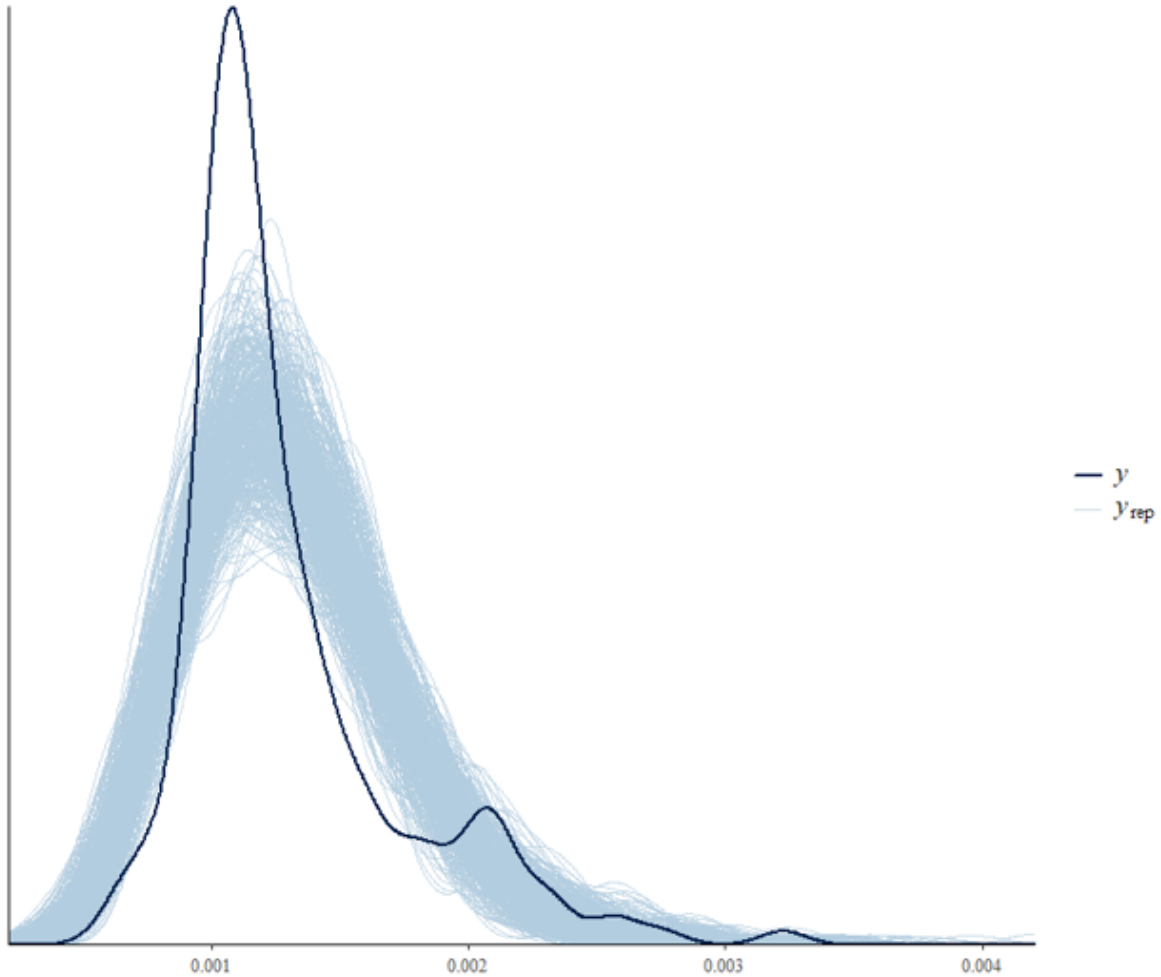


Figure 2.45: Results of the posterior predictive check for the Bayesian generalized linear mixed model of the effects of proportion introduced ancestry on Watterson's theta. The black line ("y") displays the observed data, and the blue lines ("yrep") show 500 draws from the model's posterior probability distribution.

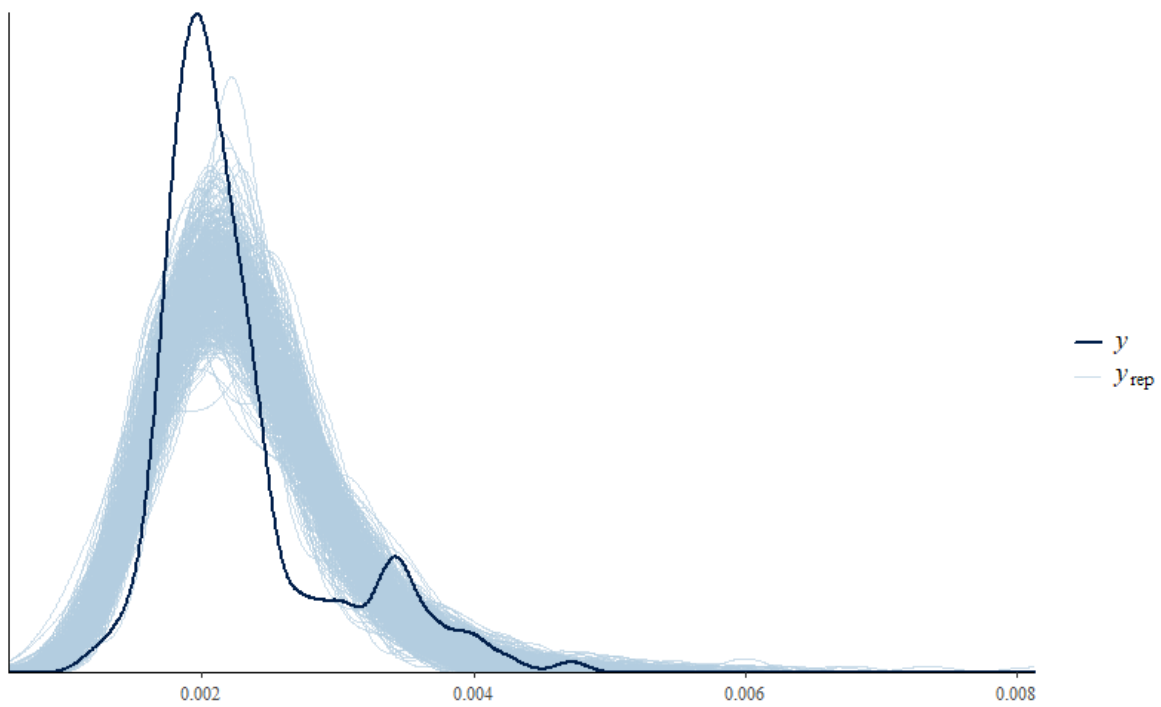


Figure 2.46: Results of the posterior predictive check for the Bayesian generalized linear mixed model of the effects of proportion nonnative ancestry on Tajima's pi. The black line ("y") displays the observed data, and the blue lines ("yrep") show 500 draws from the model's posterior probability distribution.

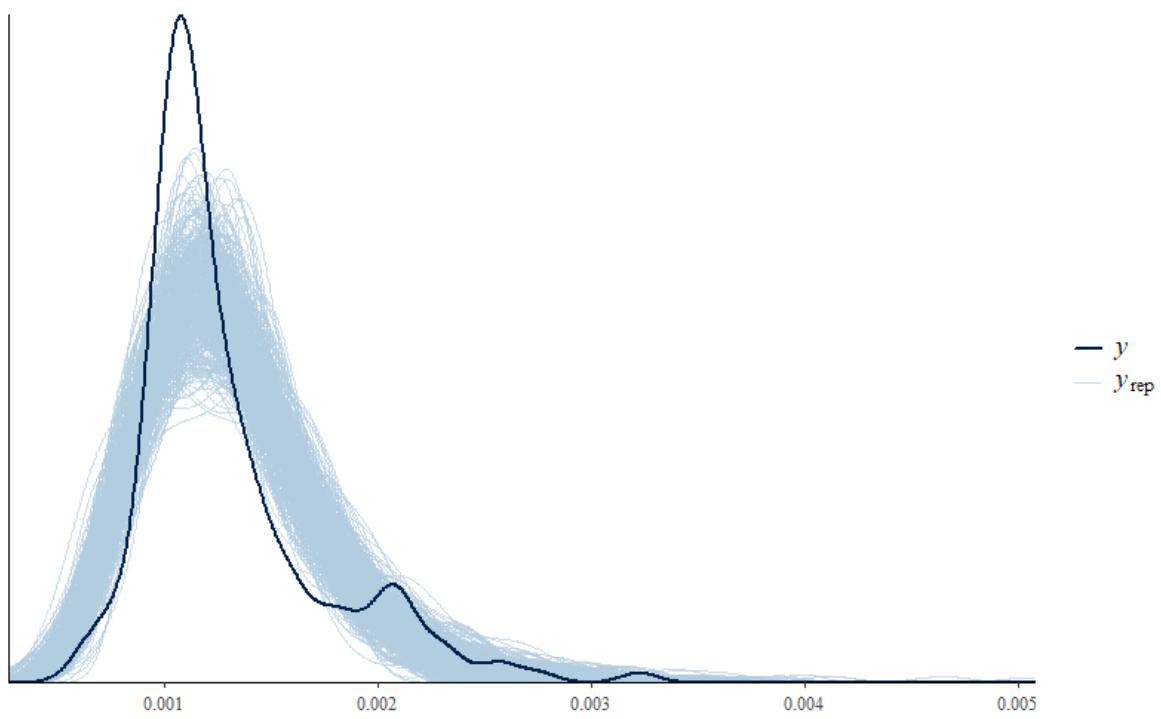


Figure 2.47: Results of the posterior predictive check for the Bayesian generalized linear mixed model of the effects of proportion introduced Snake River ancestry on Watterson’s theta. The black line (“y”) displays the observed data, and the blue lines (“yrep”) show 500 draws from the model’s posterior probability distribution.

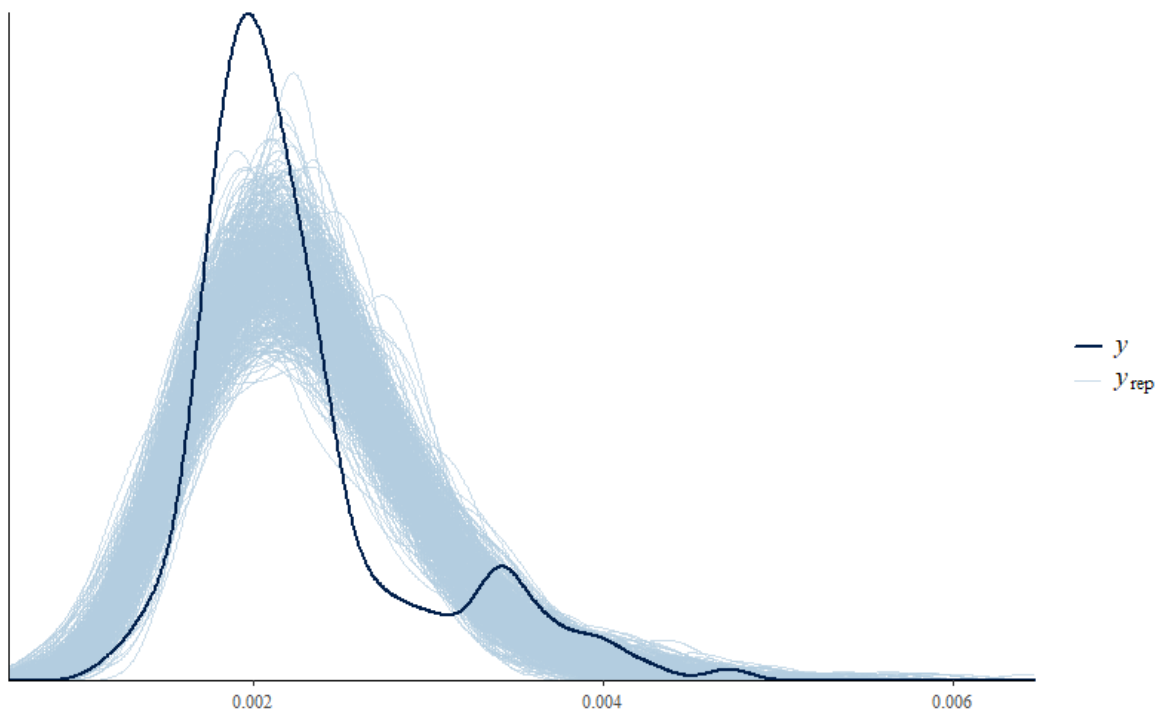


Figure 2.48: Results of the posterior predictive check for the Bayesian generalized linear mixed model of the effects of proportion introduced Snake River ancestry on Tajima's π . The black line (“y”) displays the observed data, and the blue lines (“yrep”) show 500 draws from the model's posterior probability distribution.

Chapter 3

History and inter-basin dispersal of Yellowstone cutthroat trout

Abstract

Freshwater fishes are limited in their distribution by the bounds of hydrologic basins: this makes the instances where they are able to disperse between basins an especially important for understanding their biogeography. Cutthroat trouts (*Oncorhynchus clarkii*, *O. lewisi*, *O. henshawi*, and *O. virginalis*) are known to exemplify this principle, as their diversity is almost completely described by hydrologic organization, and the geologic history of their range has played a major role in shaping their distribution. We used reduced-representation genomic data to estimate a phylogenetic tree for all cutthroat trout species, with a focus on Yellowstone Cutthroat Trout (*O. virginalis bouvieri*). We utilized demographic models in conjunction with our phylogenetic analysis to verify the suspected timing and location at which Yellowstone Cutthroat Trout crossed the Continental Divide of the Americas to access the Missouri River basin. We also corroborated previously inferred admixture between the Bonneville and Raft River basins, identified the presence of a distinct isolated lineage, and characterized the history of two important hatchery broods. Lastly, we examined the evolutionary distinctiveness of the fine-spotting phenotype that characterizes

the putative lineage of Snake River Finespotted Cutthroat Trout (*O. v. behnkei*) found in the Snake River headwaters. Phylogenetics and discrete character models showed that the phenotype was not limited to any monophyletic group and that its presence was highly variable. These results add to our understanding of cutthroat trout history and underscore the importance of geologic activity in shaping freshwater biogeography.

3.1 Introduction

Colonization is a fundamental component of our understanding of biogeography, and has long been considered a primary determinant of biodiversity distribution, especially in previously glaciated areas (Nekola 1999). Species' dispersal capabilities are known to exert control over their colonization potential (Stoddart 1981), and better understanding of how they translate to broader biogeographic patterns has importance for understanding biodiversity responses to previous and future climate change (Travis et al. 2013; Santini et al. 2016). For freshwater fishes in particular, whose dispersal capabilities are extremely limited by hydrologic basins, historical connectivity is a major determinant of species distributions and species richness (Bermingham and A. P. Martin 1998; Dias, Oberdorff, et al. 2014). When they do occur, colonization events can therefore have dramatic effects that are observable across multiple levels of biological organization (Cavin 2017; Cortés-Guzmán et al. 2024).

Both geology and lineage-specific traits interact to determine the ability for freshwater fish to move between drainage basins and colonize new habitat. Geology presents a primary control on colonization opportunity; geologic processes can limit the distribution of freshwater fish within basins (Dias, Cornu, et al. 2013; Stokes et al. 2023; D. Kim et al. 2023), alter basin connectivity and flow direction through river network reorganization (Barreto et al. 2022), or facilitate dispersal between previously disconnected freshwater habitats during periods of sea-level change (Thomaz et al. 2015). In parallel, the dispersal

abilities of a lineage, which have been shown to be heritable and predictable based physical and ecological characteristics (Louis Bernatchez and Dodson 1987; Comte and Olden 2018b; Comte and Olden 2018a), control which colonization opportunities a lineage is able to exploit. Lineages with high dispersal abilities inhabiting regions with an active geologic history might therefore be expected to colonize new basins most readily (Griffiths 2015).

Cutthroat trouts in the genus *Oncorhynchus* comprise one such group; they exhibit large body sizes, can undergo long-distance migrations to reach headwater spawning habitats (Schmetterling 2001), and evolved in western North America as it underwent several hydrology-altering geologic events. Specifically, the Yellowstone hotspot migration (17 Ma through 2 Ma; R. B. Smith and Siegel (2000)) and Quaternary glacial periods caused significant reorganization of hydrologic basins hypothesized to have enabled colonization by cutthroat trout through headwater capture events and overtopping basin divides (Link and Ernest R Keeley 2018). Prolific colonization and subsequent isolation created a number of cutthroat trout lineages, all geographically circumscribed by hydrologic basins and generally morphologically similar (though still distinguishable through meristic traits; R. J. Behnke (1992) and Janet L Loxterman and Ernest R Keeley (2012)). Despite thorough geologic investigation of dispersal routes and molecular studies, relationships between cutthroat trout lineages remain incompletely understood (Shiozawa, Evans, Houston, et al. 2018).

For one cutthroat trout lineage in particular, the Yellowstone cutthroat trout (*O. virginalis bouvieri*), diversity remains largely undocumented. Yellowstone cutthroat trout inhabit both the Snake River and Missouri River basins, which are separated by the Continental Divide of the Americas. They are thought to have accessed the Missouri River basin from the Snake River basin via Two Ocean Pass (R. J. Behnke 1992), a relatively low portion of the divide that features a continuous stretch of ponds, though this hypothesis has not been thoroughly examined with genomic data from fish on either side of Two Ocean Pass. It is also unknown if Two Ocean Pass still facilitates inter-basin movement of

Yellowstone cutthroat trout or other aquatic taxa in one or both directions. Additionally, within the Snake River basin upstream of Palisades Reservoir, there exists a putative sympatric lineage known as the Snake River Fine-spotted cutthroat trout (*O. v. behnkei*; R. J. Behnke (1992)). Mitochondrial DNA and meristic examinations have failed to resolve this variety as as a distinct evolutionary unit (M. A. Novak, Kershner, and Mock 2005), but again genomic data has not yet been applied to addressing this question and may be necessary given the relatively shallow divergence of the lineages involved. Here we utilize reduced-representation genomic data from collections of Yellowstone cutthroat trout and other lineages within *Oncorhynchus* to provide more information regarding relationships between cutthroat trout species and subspecies, characterize the ancestry and origin of historic stocking sources, identify the location and timing of Yellowstone cutthroat trout colonization of the Missouri River basin, and investigate the status of Snake River Fine-spotted cutthroat trout as an evolutionarily distinct unit.

3.2 Materials and Methods

Sampling, DNA extraction, and sequencing

We utilized previously generated sequence data from Rosenthal et al. in prep., which included newly collected tissue samples and archived DNA (Figure 3.1, Figure 3.2). Sampling was conducted via backpack electrofishing, rod and reel angling, raft electrofishing, gillnets, and weirs. Samples acquired specifically for this project were collected under University of Wyoming IACUC permit number 20220810CW00569-01 and Yellowstone National Park sampling permit number 8264. Archived DNA was obtained from the Montana Parks & Wildlife Conservation Genetics Lab at the University of Montana, the Idaho Department of Fish & Game’s Eagle Genetics Lab, and from Pisces Molecular, LLC (Boulder, Colorado, USA). In addition to samples from many wild Yellowstone cutthroat trout populations, this sample set also included individuals from

other species of trout: Lahontan cutthroat trout (*O. henshawi*; from United States Fish & Wildlife Service’s Lahontan National Fish Hatchery), Westslope cutthroat trout (*O. lewisi*; from Montana Department of Fish, Wildlife & Parks’ Washoe Hatchery), and Rainbow Trout (*O. mykiss*; from the Arlee strain). Individuals from hatchery lines within Rocky Mountain cutthroat trout (*O. virginalis*) were included as well. Three lines of Yellowstone Cutthroat trout were represented: the LeHardy Rapids strain (from the Wyoming Game & Fish Department’s Tensleep Hatchery), the Henrys Lake strain (from the Idaho Department of Fish & Game’s Henrys Lake Egg Collection Facility), and the Bar BC strain (from the Wyoming Game & Fish Department’s Auburn Hatchery and U.S. Fish & Wildlife Service’s Jackson National Fish Hatchery). Colorado River cutthroat trout were represented by fish from the Wyoming Game & Fish Department’s Daniel Fish Hatchery, Bonneville cutthroat trout were represented by fish from Utah Division of Wildlife Resources’ Glenwood hatchery (Southern Bonneville basin) and the Wyoming Game & Fish Department’s Wigwam Hatchery (Bear River basin). In addition to the samples from Rosenthal et al. in prep, we included samples from Willow and Whitehorse Creeks in Oregon, obtained from the Oregon Department of Fish & Wildlife, samples from Kelso Creek and Nanita Lake from the Colorado Department of Parks & Wildlife, and whole-genome sequences of four coastal cutthroat trout from the U.S. Forest Service.

Sequence data from Rosenthal et al. in prep and data for all other non-coastal cutthroat trout samples was prepared in twenty-two reduced representation genomic libraries for high throughput DNA sequencing at the University of Wyoming Genome Technologies Laboratory using the genotyping-by-sequencing protocol in (Thomas L. Parchman et al. 2012), with 288 or 384 individuals multiplexed per library. We digested DNA using restriction enzymes EcoRI and MseI before ligating unique 8-10 base pair nucleotide barcodes to the resulting fragments. We then multiplexed and amplified samples in two PCR replicates. We size-selected each library to retain only fragments 325-400 base pairs in length via BluePippin (Sage Science). We sequenced each library

multiplexed to 288 individuals on one Illumina NextSeq200 P3 flowcell to produce approximately 1.2 billion, 100 base pair, single-end reads. We sequenced each library multiplexed to 384 individuals on one lane of an Illumina NovaSeq 1.5B flowcell to produce approximately 1.5 billion, 150 base pair, single-end reads. Sequencing on the Illumina NextSeq platform was performed at the University of Wyoming Integrated Microscopy Core (Laramie, Wyoming, USA) and sequencing on the Illumina NovaSeq platform was performed at Admera Health (South Plainfield, New Jersey, USA).

Sequence assembly and variant calling

We assigned reads to each individual sample by parsing barcodes and then aligned each sample to the Rainbow Trout reference genome (Gao et al. 2021) (accession GCA_013265735.3) using the bwa-mem algorithm (Li 2013). For libraries with paired-end reads, we only retained the forward read as our library preparation protocol does not attach reverse adapters. To ameliorate library effects from the inconsistencies inherent in size selection across so many libraries and from different sequencing effort and read lengths, we identified variable loci on each library separately using `bcftools` (Danecek, Bonfield, et al. 2021), removing indels. Each resultant vcf was then filtered to keep only biallelic sites (SNPs), sites with at least 3 reads per individual, and sites with missing data less than 65%. We then used `vcftools` (Danecek, Auton, et al. 2011) to identify which sites were found across all filtered library vcfs and create a new vcf containing all individuals and only those sites. We filtered this vcf to retain only sites with a minor allele count of 3 or greater, sites with maximum 20% missing data, sites with quality scores greater than 19, sites with genotype quality scores greater than 9, average depth less than 50, and individuals with less than 75% missing data. Finally, we filtered our dataset to include only those individuals identified by Rosenthal et al. in prep as having less than 1.5% non-native ancestry, a selection of high-alpine lakes (Death Canyon, Grizzly Bear Lake, Lake Alice, and Lake Solitude), and the Arlee strain Rainbow Trout to be used as an outgroup. All

sequence processing was performed on the University of Wyoming’s Medicine Bow Computing Cluster.

Phylogenetic inference and demographic modeling

We constructed a phylogeny using both **CASTER** (Zhang, Nielsen, and Mirarab 2025) and **IQTree** (Minh et al. 2020) using only the individual with the lowest missing data across our filtered set of SNPs from each population. We first used **CASTER-site** to infer the topology of our tree, as it can infer the species tree using a model consistent with the multispecies coalescent that does not require concatenating all SNPs into one locus, which in turn makes our inferred topology more resilient to the effects of incomplete lineage sorting. We also inferred support for this topology by generating 1000 block bootstrap replicates. Using the best-fit topology from **CASTER**, we fit a sequence evolution model to our data using **ModelFinder Plus** (Kalyaanamoorthy et al. 2017) and then estimated branch lengths for our tree using the best-fitting model as inferred by AICc. For ease of visualization, we created ultrametric trees using the **chronos** function in the R package **ape** (Paradis and Schliep 2019): we calibrated the node splitting rainbow trout from all cutthroat trout at 10.2 Ma (Stearley and Gerald R Smith 2016) and fit a molecular clock model.

To investigate the timing of movement across Two Ocean Pass, we used $\delta a\delta i$ (Gutenkunst et al. 2009) to fit demographic models. We re-called variants for samples only from our collections on either side of Two Oceans Pass, Pacific Creek and Atlantic Creek, filtered only to remove sites and individuals with greater than 50% missing data, sites with observed heterozygosity over 0.8, and sites with average depth over 50. We then generated site frequency spectra for each population using **easySFS** (<https://github.com/isaacovercast/easySFS>). We constructed a demographic model that featured the following steps: one population splitting into two before growing exponentially over a period of time. Then, the two populations remained at that constant size for another period of time. To test for the timing of gene flow between between populations,

we allowed asymmetrical migration between populations at varying time points: during the population growth phase, during the constant population size phase, and during both phases. Before continuing to model selection, we examined model residuals to ensure that they appeared normally distributed and small in magnitude. Because our models differed only in the timing of migration and not in the number of parameters fit, we compared fitted log-likelihoods directly and chose the model with the highest log-likelihood. We conducted 20 rounds of model fitting to verify that our models identified similar optima in the likelihood surface. We estimated uncertainty in the parameter estimates using the Godambe Information Matrix methods implemented in *δaδi* (Coffman et al. 2016). To estimate the date of Yellowstone cutthroat trout colonizing the Missouri River basin, we converted the number of *δaδi* units since population divergence to years assuming a generation time of three years and using mutation rates from Atlantic Salmon (*Salmo salar*; $1.25 * 10^{-8}$; Rougemont et al. (2023)) and an estimated average mutation rate for all fish ($5.97 * 10^{-9}$; Bergeron et al. (2023)). Converting the parameter estimates to real-time also required an estimate of the sequence length used to generate the site frequency spectrum. We estimated this length by retaining all sites (variant and invariant, no filter for missing data) and multiplying this by the proportion of sites removed from our variants-only vcf during subsequent filtering to account for the effects of our post-variant calling filters.

Genetic clustering and D-statistics

We wanted to understand the history and origins of two fish hatchery stocking sources of Yellowstone cutthroat trout with uncertain genetic status, Henrys Lake and South Paint Rock Creek, and therefore estimated ancestry coefficients using the program **admixture** (Alexander, Novembre, and Lange 2009). In addition to the two focal lines, we included samples from nearby populations: Dry Medicine Lodge Creek being nearby South Paint Rock Creek and Wyoming and Corral Creeks being near Henrys Lake. These wild populations were paired with samples from the Wyoming Game & Fish Department's

hatchery system representing cutthroat trout from the Bear (Wigwam BRC), Snake (Auburn SRC), and Yellowstone (Tensleep YCT) river basins. We ran **admixture** with 1 through 10 clusters, and selected a best-fit cluster by examining where cross-validation errors stopped decreasing with increasing clusters.

To understand the history of gene flow between *O. virginalis* lineages, we utilized D-statistics. Specifically, we tested for the presence of gene flow between South Paint Rock Creek and the Bear River basin as well as between the Bonneville and Snake River basins in the Portneuf River and Raft River drainages, as this has been indicated in previous research using mitochondrial haplotypes (Matthew R. Campbell et al. 2011). We calculated D-statistics using **Dsuite** (Malinsky 2019). To assess the potential for gene flow from the Bear River basin into South Paint Rock Creek, we retained only samples from South Paint Rock Creek, the Tensleep YCT hatchery line, and the Wigwam BRC hatchery line and designated each sample as its own “species”. For the analysis of gene flow between the Bonneville and Snake basins, we designated all Raft River populations as one “species”, all Portneuf River populations as another “species”, retained the Glenwood Fish Hatchery samples (our representative for non-Bear River, Bonneville Basin ancestry) as another “species”, and designated samples from all collections from Yellowstone Lake as another “species”. We removed all other samples from the analysis. We conducted a second **Dsuite** analysis to confirm. We corrected the resultant p-values for multiple testing bias using a Bonferroni-Hochberg correction.

Character evolution modeling

To investigate the lability of the fine-spotted phenotype that characterizes the Snake River Fine-spotted cutthroat trout, we fit discrete trait evolution models onto a subset of our **CASTER** tree. For this, we retained all non-alpine lake, Snake River basin populations. Data on spotting phenotypes for each population was obtained from the Yellowstone cutthroat trout geodatabase obtained from StreamNet (Alger, Ryan et al. 2018), which categorized

each population as having “large”, “fine”, or “both large and fine” spotting. Because these data use a polymorphic categorization, we utilized the polymorphic trait models implemented in `phytools` (Revell 2024), which model polymorphic traits with the assumption that a clade must transition through the polymorphic state before transitioning to the other state (e.g., fine spotting to both fine and large spotting before, finally, large spotting). We fitted three different models of trait evolution: all rates different (“ARD”), equal rates (“ER”), symmetrical rates (“SYM”), or the transient model. We selected the best-fitting model via AIC. We then used the parameters from the model to reconstruct ancestral states via 100 replicates of stochastic character mapping (Huelsenbeck, Nielsen, and Bollback 2003), again implemented in `phytools`.

3.3 Results

DNA sequencing, sequence assembly, and variant calling

DNA sequencing generated an average of 2.41 million reads per individual; on average, 98% were successfully aligned to the reference genome. After removing 1,314 individuals with high proportions of missing data, 6,059 individuals remained: we selected individuals for subsequent analysis from these samples. Before filtering, we retained 109,362 sites across all libraries. After filtering, we retained 105 individuals and 58,157 sites for phylogenetic analysis. The same set of SNPs was used for D-statistics calculation. For demographic modeling, we retained 21 samples from Pacific Creek collections and 20 from Atlantic Creek collections. The site frequency spectra were generated from 294,311 SNPs.

Phylogenetic inference and demographic modeling

The best-fit topology inferred by `CASTER` showed strong support across many nodes, including the backbone of our tree. Average bootstrap support across all nodes was 82 (Figure 3.3, Figure 3.4). ModelFinder Plus selected TVMe+ASC+R4 as the best sequence

evolution model. Model selection in all 20 rounds of demographic model fitting identified that gene flow only during the population growth phase was supported over gene flow during the constant population size phase or continuous gene flow since divergence (Table S1, Figure S1). Gene flow from Atlantic Creek to Pacific Creek was estimated to be roughly 2.1 times gene flow in the opposite direction, though the standard deviation of this estimate was approximately 6.6 times that for gene flow in the opposite direction and larger than the magnitude of the estimate. Divergence date estimates ranged from 2.2-15.4 thousand years ago (mean 8.9) for the salmon mutation rate and 3.6-32.2 thousand years ago (mean 18.6) for the average fish mutation rate (Figure 3.5). Parameter estimate uncertainty was generally low and consistent across subsequent rounds of model fitting: the average uncertainty across all time estimates was 0.078, which corresponds to approximately 3,100 years if using the salmon mutation rate.

Genetic clustering, D-statistics, and character evolution modeling

Cross-validation errors from **admixture** suggested $K = 7$ as the best number of clusters to represent our data. At this number of clusters, each population was assigned its own cluster except for Henrys Lake (Figure 3.6), which was inferred to be mix of ancestries from Corral Creek and the upper Yellowstone River (Tensleep YCT). At $K = 6$, the reduction of clusters causes South Paint Rock Creek to now be inferred as a mix of ancestries from the Bear River (Wigwam BRC) and upper Yellowstone River (Tensleep YCT).

D-statistics supported the **admixture** results at $K = 6$ clusters by supporting gene flow between the Bear River (Wigwam BRC) and South Paint Rock Creek (p-value $< 1 * 10^{-5}$). D-statistics also supported excess allele sharing between the Southern Bonneville Basin (Glenwood Fish Hatchery) and both Raft River populations (Grape Creek and Onemile Creek). P-values of D-statistic estimates were $3.21 * 10^{-14}$ and $6.54 * 10^{-11}$, respectively.

The symmetric rates evolution model, where transitions from Fine to Fine + Large spotting are equal to those in the opposite direction, was identified as the best fitting

model by AIC. The polymorphic trait evolution model identified relatively high instantaneous transition rates of 3.02 between Fine and Fine + Large and 0.94 between Fine + Large and Large. Ancestral state reconstruction visualized this high lability via frequent transitions, including even within the same branch (Figure 3.7).

3.4 Discussion

Understanding the relative influence of fundamental processes in biodiversity maintenance and generation, including dispersal and colonization, across environments and taxa stands to improve our understanding of Earth’s history and its potential futures under anthropogenic climate change (McFadden et al. 2023). Here we use genomic data to examine the history of lineages within the cutthroat trouts (genus *Oncorhynchus*), whose diversity is primarily associated with colonization rather than in-situ speciation (Janet L Loxterman and Ernest R Keeley 2012), with a focus on Yellowstone cutthroat trout (*O. virginalis bouvieri*).

Relationships between cutthroat trout lineages

Our phylogenetic topology supports recent hypotheses regarding the history of cutthroat trout lineages and also differs in several notable aspects (Figure 3.3). We recover Rocky Mountain cutthroat trout (*O. virginalis*) as sister to the Lahontan cutthroat trout (*O. henshawi*), which supports the colonization of the Bonneville Basin (and subsequently the Colorado and Snake River basins) from the Lahontan basin. This placement is in agreement with Kokkonen et al. (2024), whereas Van Orden et al. (2026) identifies *O. virginalis* as sister to the group containing both *O. henshawi* and Westslope cutthroat trout (*O. lewisi*). We also identify Coastal cutthroat trout (*O. clarkii*) as sister to all other cutthroat trouts, in agreement with Van Orden et al. (2026), whereas Kokkonen et al. (2024) finds Westslope cutthroat trout (*O. lewisi*) as sister to all other cutthroat trout

species. Our inferred relationships between lineages within *O. virginalis* lineages are identical to that of Van Orden et al. (2026) and those found in some trees presented by Kokkonen et al. (2024). Shiozawa, Evans, Houston, et al. (2018) recovers Southern Bonneville Basin as nested within the lineage containing both Colorado River cutthroat trout lineages (*O. v. pleuriticus*), Greenback cutthroat trout (*O. v. stomias*), and Rio Grande cutthroat trout (*O. v. virginalis*): our analysis identifies Southern Bonneville Basin fish as sister to this lineage (though we lack data from *O. v. stomias* and *O. v. virginalis*). Lastly, our results confirm the results of many studies that find cutthroat trout from the Bear River basin as more closely related Yellowstone cutthroat trout rather than nested within other Great Basin lineages (G. Smith et al. 2002; Shiozawa, Evans, Unmack, et al. 2010; Janet L Loxterman and Ernest R Keeley 2012; Shiozawa, Evans, Houston, et al. 2018; Kokkonen et al. 2024; Van Orden et al. 2026), which indicates that their current taxonomic designation within Bonneville cutthroat trout (*O. v. utah*) is in need of revision. In fact, our results do not indicate that cutthroat from the Bear River basin form a monophyletic group and instead show them as a paraphyletic group within Yellowstone cutthroat trout.

The differences in the genetic data used in our analysis and the studies discussed above likely explains the discrepancies between them: Kokkonen et al. (2024) used a transcriptomic dataset from hundreds of nuclear genes, whereas Van Orden et al. (2026) utilized mitochondrial genomes and Shiozawa, Evans, Houston, et al. (2018) used mitochondrial gene sequences. We observe higher concordance between our results and those of Van Orden et al. (2026), which might be expected given the shared use of larger amounts of data when compared to previous cutthroat trout phylogenetic investigations (Shiozawa, Evans, Houston, et al. 2018). The estimated divergence dates of Van Orden et al. (2026) differ from ours in that they estimate a much older divergence for many *O. virginalis* lineages compared to ours; we estimate that all splits within *O. virginalis* occurred within the last 250 thousand years, whereas Van Orden et al. (2026) estimate this time window as within the last 1-2 million years. This discrepancy could be due to the

differences inherent to nuclear versus mitochondrial lineage signals, or it could be due to the methodological differences imposed by the use of SNP versus sequence data.

Regardless, the propensity for cutthroat trout to disperse across basins and ability to hybridize with one another increases the likelihood of natural reticulations within their history, which further muddies our ability to accurately infer their history.

Within Yellowstone cutthroat trout, patterns of differentiation also reflect divisions between hydrologic basins (Figure 3.4). A few exceptions to this pattern are noticeable, however. First, Willow Creek is within the Upper Snake basin but we place them phylogenetically within the Snake Headwaters basin. We attribute this to the sampling location's close proximity to the boundary between the Upper Snake and Snake Headwaters, which is just between Willow Creek and our next upstream populations. Willow Creek's placement in the tree is adjacent to those geographically-proximate, Snake Headwaters basin neighbors, and so its similarity to those populations could be attributable to geographic proximity. Alternatively, this could be attributable to trace amounts of admixture from translocations of Snake Headwaters ancestry fish not detected in Rosenthal et al. in prep.

Henrys Lake shows a departure from our expectation based on its geographic location; our genetic clustering analysis also infers admixture from the upper Yellowstone River basin. Altogether this suggests that fish from the Yellowstone River were introduced by human activity into Henrys Lake, which may not be surprising given Henrys Lake's close proximity to a lake that historically housed a large and prolific fish hatchery (Franke 1997). South Paint Rock Creek, which was the source for a historic captive line of Yellowstone cutthroat trout, is indicated by D-statistics and genetic clustering to contain ancestry from the Bear River basin. Again human activity is the likely causative agent for this admixture. The presence of Yellowstone River ancestry in South Paint Rock Creek rather than ancestry from the nearby Dry Medicine Lodge Creek also suggests that the population received admixture from that region as well; in fact, South Paint Rock Creek

possesses a large and likely impassible waterfall barrier in its downstream reaches (Alger, Ryan et al. 2018), and may therefore have been naturally fishless before anthropogenic introductions. Lastly, Lake Solitude, Grizzly Bear Lake, and Death Canyon all show ancestry from the Yellowstone Lake region despite being located within the Snake Headwaters basin. These alpine populations have been previously noted to have anomalously large spots and were found to possess Snake River mitochondrial haplotypes (M. A. Novak, Kershner, and Mock 2005), though Rosenthal et al. in prep found that their nuclear genomes consisted of predominantly Yellowstone River ancestry. Although stocking records are not known from these locations, the most likely explanation for this distribution would be human introduction.

Vicariance and inter-basin dispersal

D-statistics were significantly non-zero for trios of populations that indicate excess allele sharing between the Southern Bonneville Basin and populations in the Raft River basin. Gene flow between these basins has been previously suggested by mitochondrial haplotype data (Matthew R. Campbell et al. 2011), and is thought to have been facilitated by stream capture events. Stream capture events, a mode of vicariance in freshwater biogeography, have shaped other facets of cutthroat trout history, including the capture of the Bear River by the Bonneville Basin (Pederson et al. 2016), and have been implicated in driving patterns of biodiversity in galaxiid fishes as well (Waters, Burrige, and Craw 2020). For cutthroat trouts, geology-induced vicariance and colonization appear to be the primary driver of diversification, and discrepancies between patterns of relatedness and hydrologic organization have historically been able to be attributed to geologic events (Janet L Loxterman and Ernest R Keeley 2012).

Wyoming Creek is estimated to be sister to all Yellowstone cutthroat trout and cutthroat trout from the Bear River, which we attribute to basalt flows that occurred periodically from over 800 kya until as recently as 28 kya that may have isolated the

Henrys Fork from other Snake River populations (Alt and Hyndman 1989; Moore et al. 2022): signs of this divergence were first noticed by (Matthew R. Campbell et al. 2011). Our phylogenetic tree estimates that Wyoming Creek diverged from all other Yellowstone cutthroat trout populations approximately 112 kya, which somewhat corresponds with the age of the lava flows that formed the Hatchery Butte Bench approximately 81 kya and may have blocked the flow of the Henrys Fork and Warm River (Moore et al. 2022). Geology-induced vicariance is therefore also likely responsible for creating the most evolutionarily distinct lineage of cutthroat trout across the upper Snake and Bear River basins.

The sequence of splits within our tree follows the expected colonization pathway across Two-Ocean Pass, into the Yellowstone River basin, and then into the Bighorn and Tongue basins (Figure 3.4). Our demographic models inferred a split time between samples from Atlantic and Pacific Creeks between approximately 20 and 10 thousand years ago (Figure 3.5), which roughly coincides with the retreat of the Pinedale Glaciation from the Yellowstone Plateau (Licciardi and Kenneth L. Pierce 2018). The results of both our phylogenetic and demographic modeling analyses suggest that Yellowstone cutthroat trout only colonized the Missouri River basin after the Pinedale Glaciation despite existing as an evolutionarily distinct lineage for at least 340,000 years before then (Shiozawa, Evans, Houston, et al. 2018). Ecological niche models predict that suitable habitat for cutthroat trout existed within the Missouri River basin during the last glacial maximum (Jessica Lynne Metcalf 2007), and so fish able to access the drainage before the Pinedale Glaciation could have potentially survived. However, the Pinedale Glaciation is likely to have entirely re-shaped the topography (and therefore also hydrology) of Two-Ocean Pass, as it also erased much of the evidence of the effects from the previous glaciation (Kenneth L Pierce 1982). These lines of evidence point to a lack of opportunity for Yellowstone cutthroat trout to cross the Continental Divide into the Missouri River basin before the end of the Pinedale Glaciation, and that this is likely because Two Ocean Pass

was not traversable by fish until after this point.

Our best-fitting demographic model was one that featured no gene flow between populations up to the present, but did have gene flow during the period of population growth before the present. We do not believe that this constitutes evidence for an inability for fish to disperse across Two-Ocean Pass in the present, and instead view it as evidence for little realized gene flow compared to the initial colonization period of the Yellowstone River. Anecdotally, while our surveys via backpack electrofishing in Pacific Creek did find fish at most every location visited, the density of fish appeared low relative to that in Atlantic Creek. Now that the Yellowstone River upstream of Yellowstone Lake is home to a robust, migratory population of Yellowstone cutthroat trout and not a largely or entirely fishless ecosystem, any potential dispersers from the small, resident population of Pacific Creek trout might have little hope of introgression.

Spotting phenotype evolution

Within the Snake River Headwaters, spotting pattern was highly variable across space and across our phylogeny (Figure 3.7); though presence of fine spotted individuals within a population was geographically concentrated, populations with fine spotting did not strictly form a monophyletic clade. Polymorphic character models confirmed a high transition rate: a sample of our stochastic character mapping shows that multiple transitions per individual population branch could occur (Figure 3.7). These results suggest that the fine-spotted trait is a highly variable trait maintained in polymorphism within the Snake Headwaters metapopulation and not a particularly reliable trait for identifying if an individual belongs to a particular lineage. One caveat is that the phenotype data we used here is not directly linked to the individuals used in genetic analysis, and instead is population-level assessments from (M. A. Novak, Kershner, and Mock 2005). There are several possibilities that may explain this striking phenotypic variation in the Snake Headwaters that does not rely in a distinct, monophyletic group possessing this phenotype. First, fine-spotting may

be maintained in the population due to selective advantage, either for the spotting pattern itself or due to the effects of linked loci. Second, the fine-spot phenotype may have arisen during a period of glacial refuge and subsequently spread beyond its origin, with no impacts on gene flow based on phenotype. A more in-depth survey of Snake River Yellowstone cutthroat trout spotting phenotypes as they relate to environment, as well as examination of the genes underlying the fine-spotting phenotype, would clarify the trait's genetic underpinnings and ecological role.

Conclusions

Genomic data can provide enormous and multi-faceted utility for understanding the history of populations and species. We generated a phylogenetic tree for Yellowstone cutthroat trout populations and other lineages of cutthroat trout, and find results that work towards a scientific consensus for the history of this group. Our phylogeny also identified that diversification within Yellowstone cutthroat trout was primarily organized around drainage basins, as is common within the species. Within wild populations, we found evidence for admixture, isolation, and dispersal all likely driven by geologic events: geologic activity may therefore be an underappreciated determinant of biodiversity for dispersal-constrained taxa like freshwater fish. Along similar lines, some of the phenotypic basis for the identification of a sympatric lineage of cutthroat trout within the Snake River headwaters appears to be highly variable within and between populations. Overall, our findings shed light onto relatively under-studied questions in cutthroat trout biogeography and history, and further emphasize the role geologic activity has had on the history of the group.

3.5 Figures

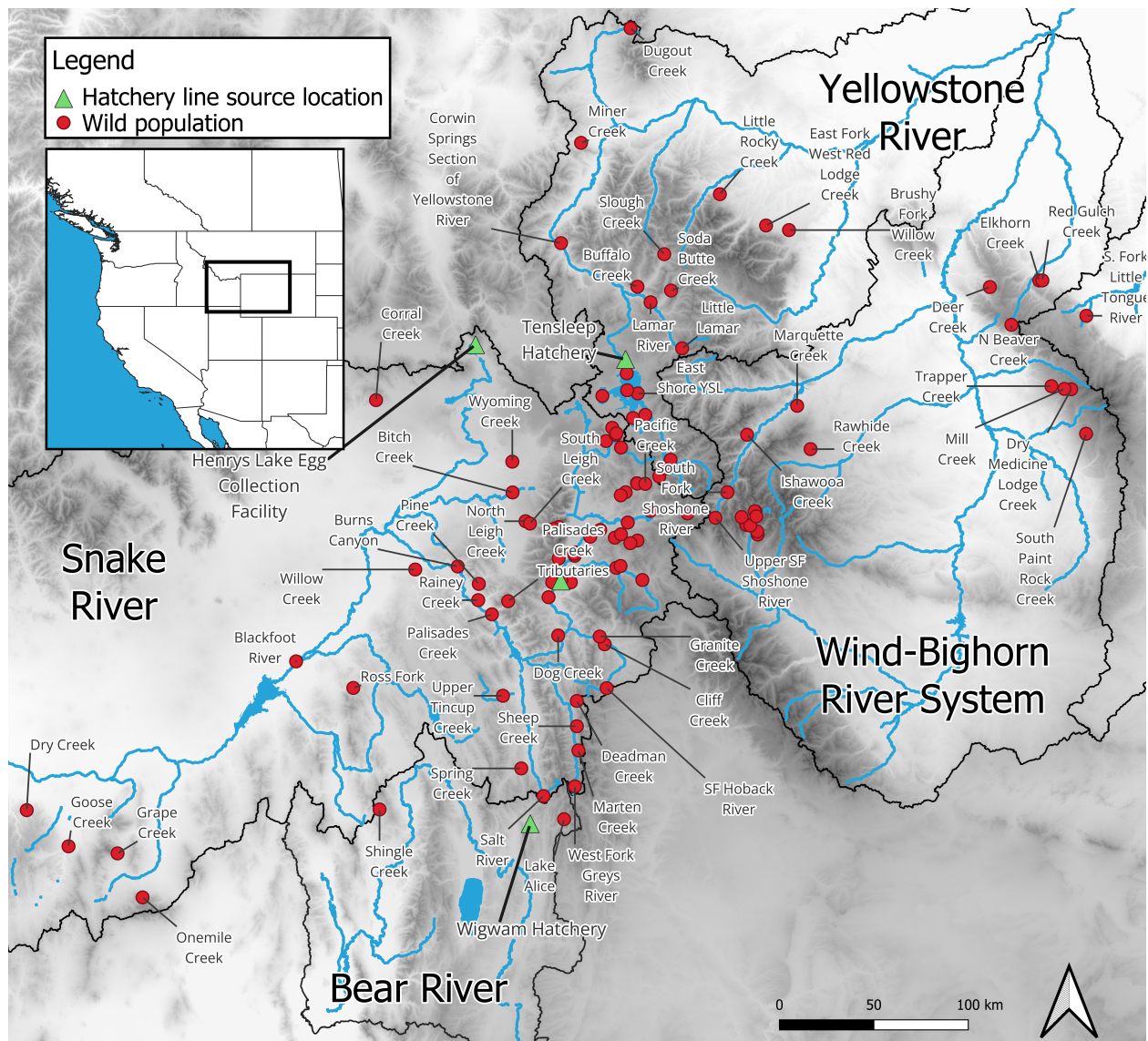


Figure 3.1: Map of sample locations and source locations for hatchery lines used in this study. Points not labeled in this figure are labeled in the next figure.

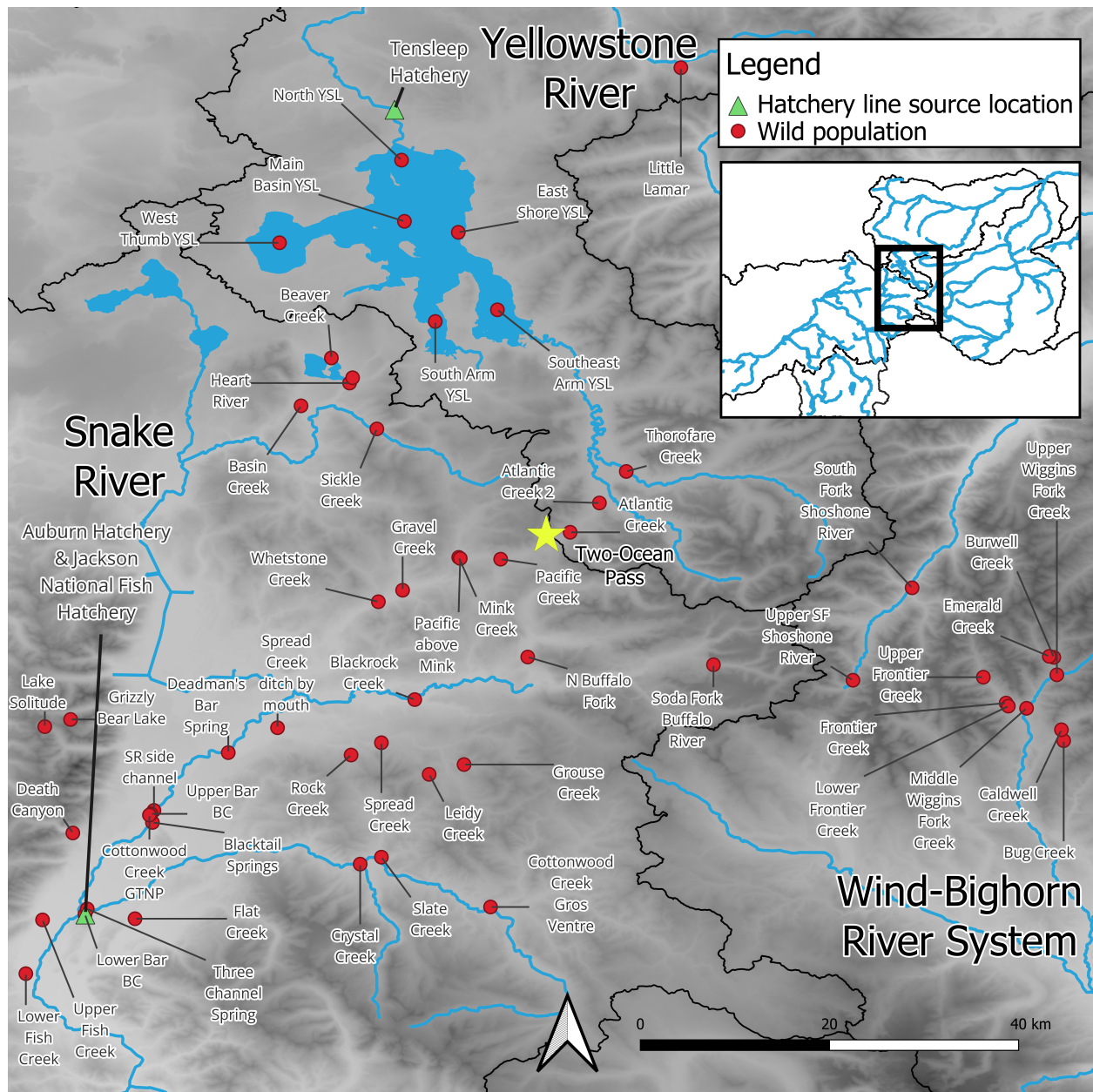


Figure 3.2: Map of sample locations and source locations for hatchery lines used in this study from the northern Snake Headwaters and Yellowstone Lake regions. The star denotes the location of Two-Ocean Pass.

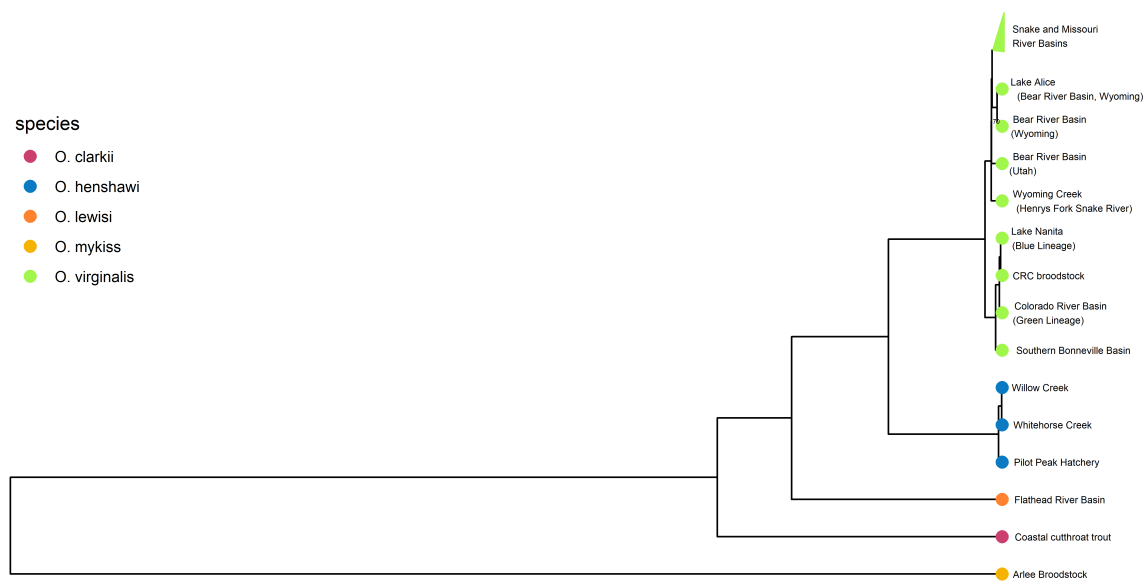


Figure 3.3: Phylogenetic tree of cutthroat trout lineages from nuclear SNP data, with topology inferred by **CASTER**. Bootstrap support is over 95 for each node shown, and so none are displayed. Tip point colors refer to the species designation of that tip, and tip labels provide additional information about the location of origin.

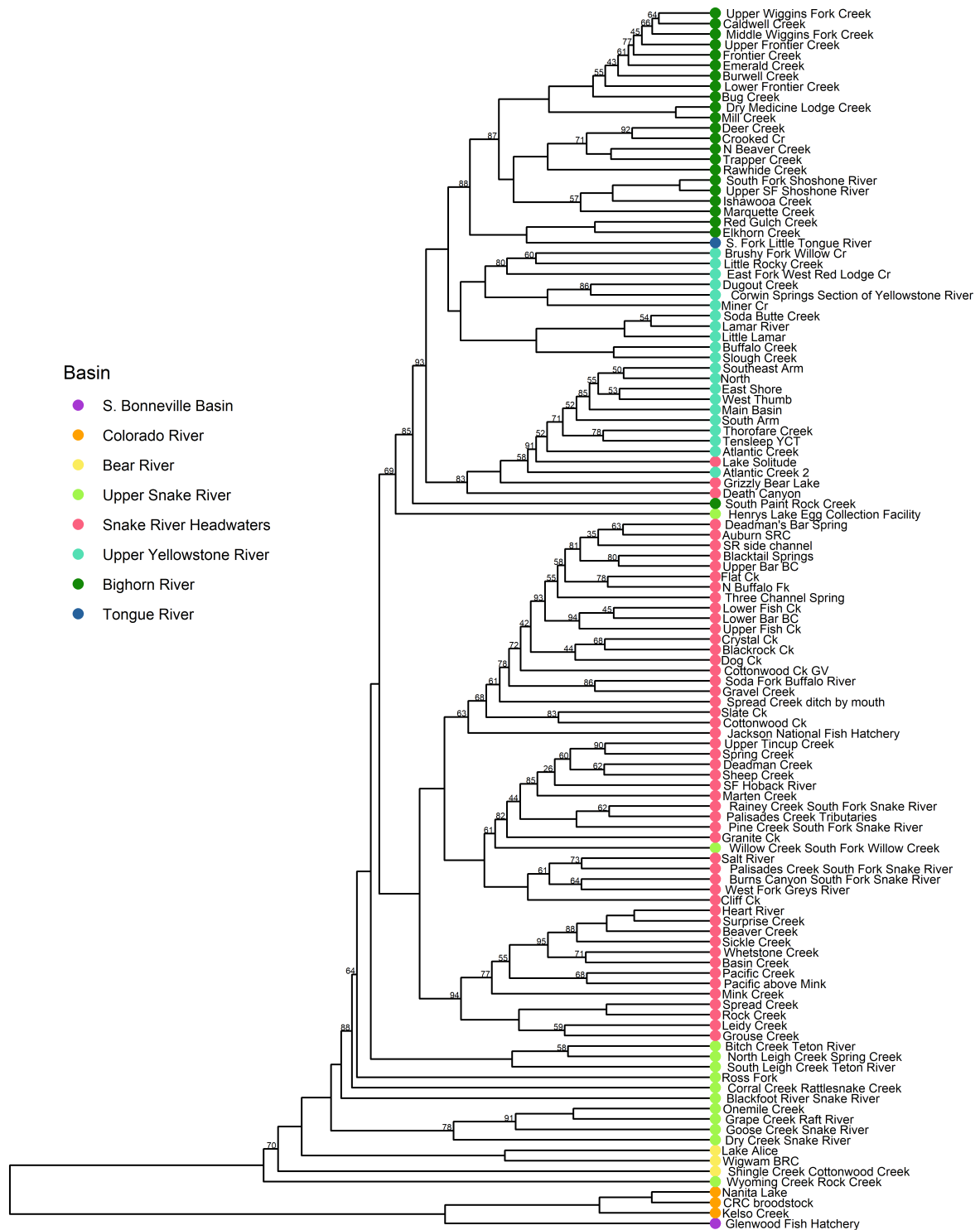


Figure 3.4: Phylogenetic tree of lineages within *Oncorhynchus virginalis* from nuclear SNP data, with topology inferred by **CASTER**. Bootstrap support is shown for each node where the support is below 95. Tip point colors refer to the basin of origin for that tip. Tips with placement outside of the expectation from their basin of origin are discussed in the discussion section.

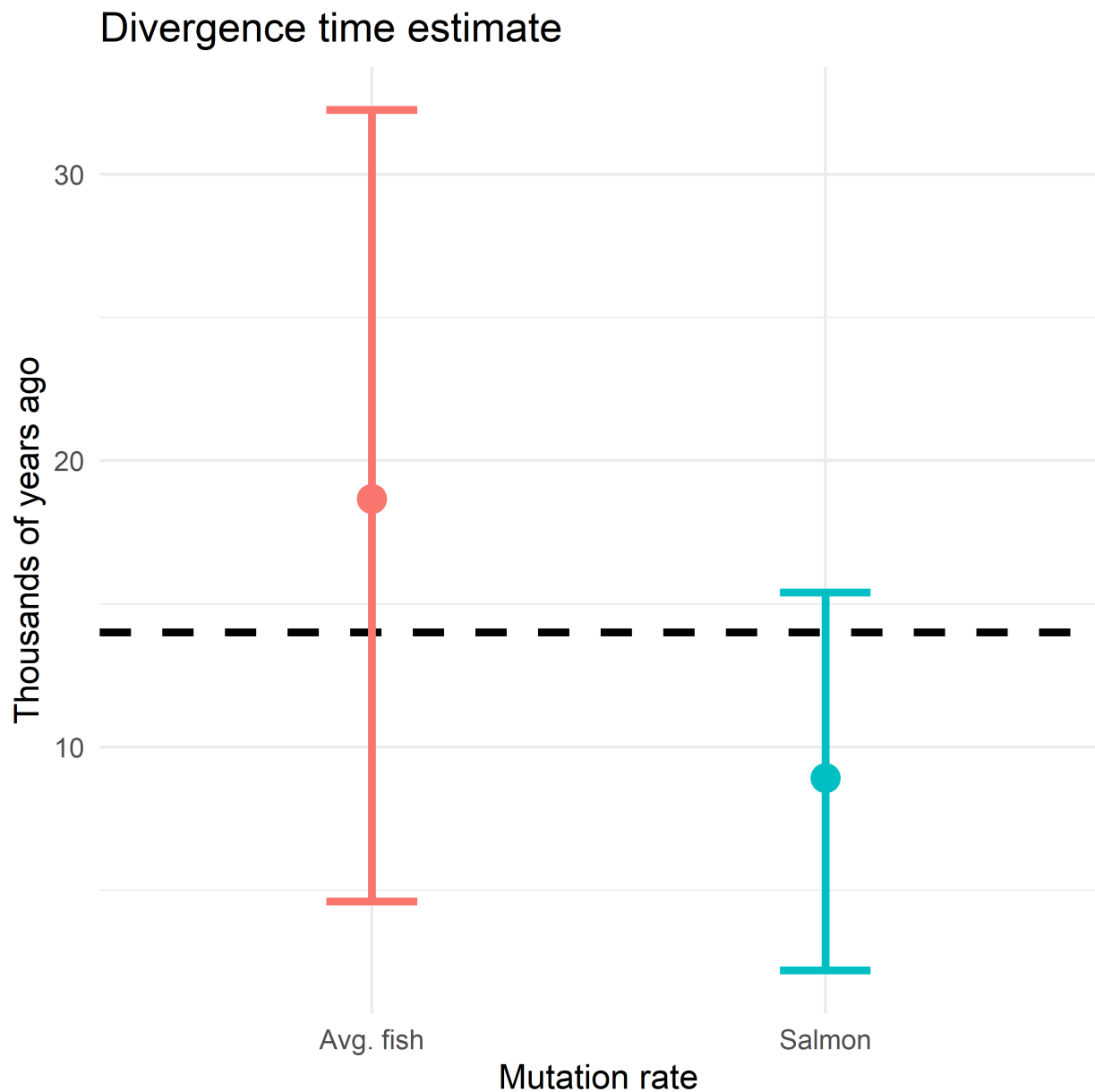


Figure 3.5: Estimates and their accompanying uncertainty of divergence time between Atlantic Creek and Pacific Creek, which are connected by Two-Ocean Pass. Two mutation rates are used to convert coalescent units to real years, with an assumed generation time of three years per generation. The dashed line shows the estimated date at which the Pinedale Glaciation left the Yellowstone Plateau, from (Licciardi and Kenneth L. Pierce [2018](#)).

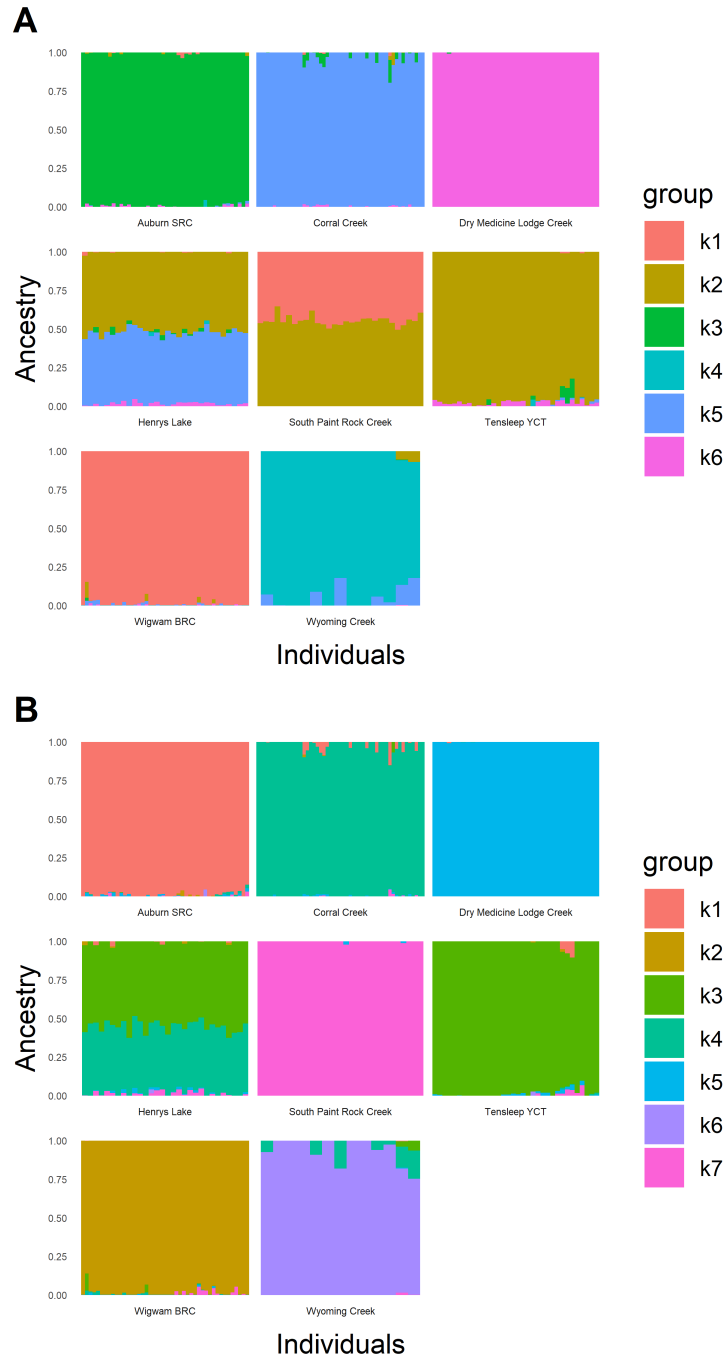


Figure 3.6: Results of genomic clustering performed in **admixture** to understand the history of the Henrys Lake and South Paint Rock Creek hatchery lines at $K=6$ (A) and $K=7$ (B). Included are nearby populations – Corral Creek and Wyoming Creek for Henrys Lake, Dry Medicine Lodge Creek for South Paint Rock Creek – and hatchery-line references for Yellowstone (YCT), Snake (SRC), and Bear River (BRC) ancestries.

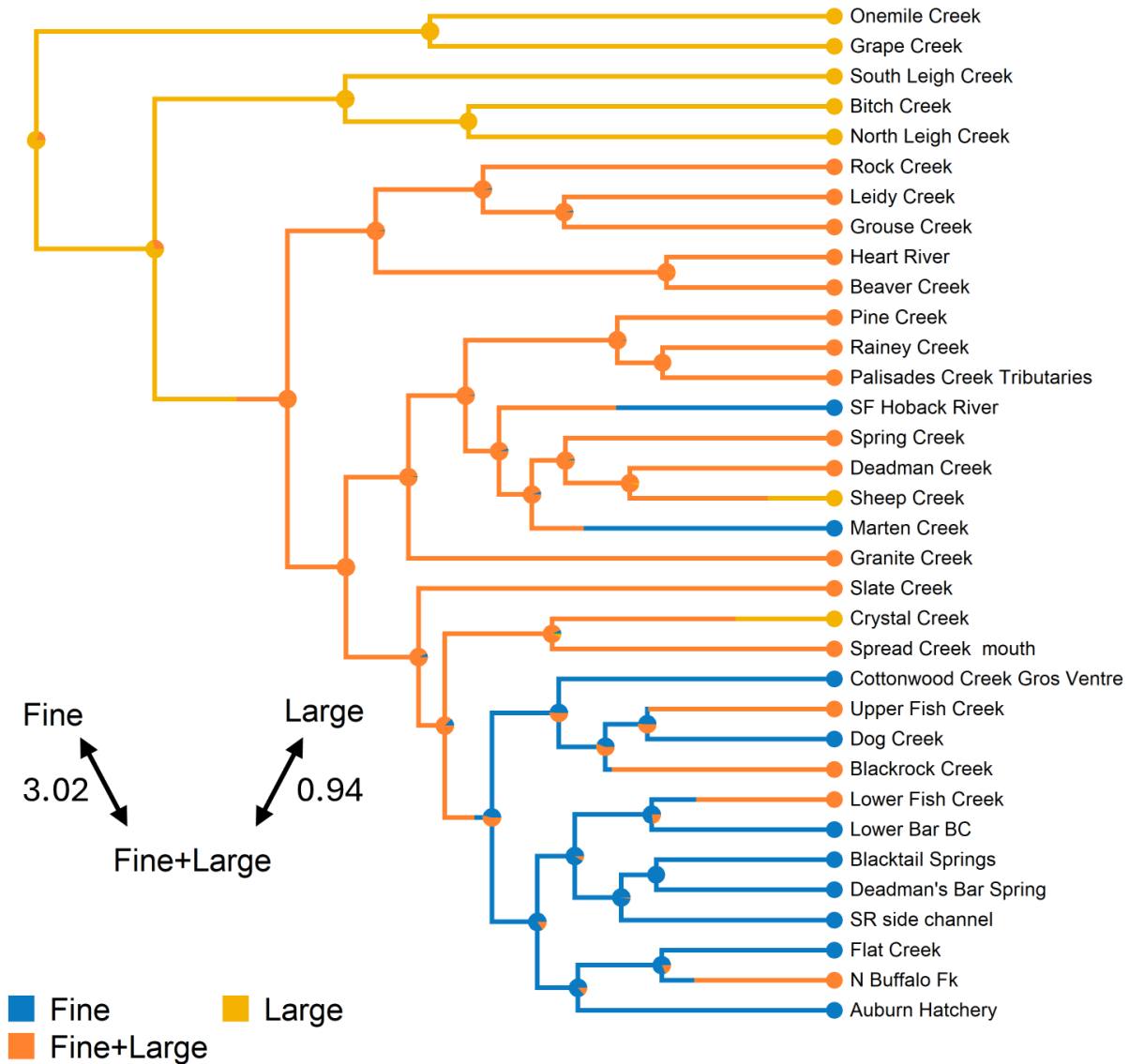


Figure 3.7: Phylogenetic tree of lineages within *Oncorhynchus virginalis bouvieri* within the Snake River basin from nuclear SNP data, with topology inferred by **CASTER**. Tip point colors refer to the spotting phenotype observed in individuals from the location represented by that tip. Colors at nodes shows the reconstructed ancestral state at that node based on 100 stochastic character maps and a polymorphic discrete trait model. Colors along branches show the results from a single stochastic character map. Model structure and instantaneous transition rates are shown in lower left corner, above the legend.

Chapter 4

Roles of history and environment on population genetic diversity in Yellowstone cutthroat trout

Abstract

Identifying associations between local environmental conditions and population genetic diversity can help bridge the gap between ecological and evolutionary outcomes while also enabling landscape-scale predictions of adaptive potential for conservation. However, previous attempts to predict species genetic diversity have had mixed success, potentially driven by variation in species-specific characteristics such as recombination rate or life history. We use population genetic diversity and phylogenetic data from 165 populations of Yellowstone Cutthroat Trout in conjunction with environmental data test for factors determining population genetic diversity. Statistical support for environmental associations was dependent on the diversity metric used, but associations between summer air temperature and position within the watershed (mainstem vs. headwaters) were identified in models that explained roughly 25% of the variation in the data. To understand whether genetic diversity is predicted by evolutionary history, we fit models of trait evolution,

assuming genetic diversity acts as an evolving trait. After testing several diversity metrics, we find the strongest phylogenetic signal for observed heterozygosity, which also shows significant differences across populations on either side of a major geographic barrier. Although predicting genetic diversity poses many challenges regardless of taxonomic scope, we are able to identify environmental predictors of population genetic diversity that can facilitate conservation and restoration prioritization.

4.1 Introduction

Genetic diversity is the source from which all novel, heritable adaptations derive, and as such is key to life's diversity. The degree of genetic diversity present within a lineage could therefore be a meaningful predictor of their potential to adaptively respond to future conditions, and is considered an important component of modern conservation science (DeWoody et al. 2021; Kardos et al. 2021). Harnessing the information contained within biodiversity below the species level could be crucial for creating long-lasting conservation successes that enable species to survive our current era of unprecedented environmental change and thrive after it ends. However, the data required to assess genetic diversity within a species or sub-specific lineage is not easy to generate, as it requires sampling at the population level across the range of the taxon. This often presents logistical hurdles, as well as financial ones. However, if preserving genetic diversity before it is irrevocably lost is a goal of conservation science, there is clear need to understand where we might expect to find more or less diversity to prioritize conservation efforts.

Predicting the amount of genetic diversity within a lineage as it relates to species-specific or environmental factors presents a potential solution. Models have been successful in predicting conservation status and extinction risk in both animal and plant taxa (Lee and Jetz 2011; Bland et al. 2015; Pelletier et al. 2018). For genetic diversity, models have had mixed success: genetic diversity within species of *Silene* was found to be

well-predicted by range size (Muyle et al. 2021), and a meta-analysis of terrestrial animals and plants found that population genetic diversity was well predicted by species-specific characteristics and contemporary climate (as opposed to historic climate; De Kort et al. (2021)), though these examples form the exception (Glémin, Bazin, and Charlesworth 2006; Romiguier et al. 2014; Buffalo 2021). Species-specific characteristics, primarily life history traits and recombination rates, have been the focus of several studies investigating the determinants of interspecific population genetic diversity (Turner et al. 2006; R. Frankham 2012; Romiguier et al. 2014; Ellegren and Galtier 2016; Hague and Routman 2016; Mackintosh et al. 2019; De Kort et al. 2021). However, within a species, we might expect these traits to show little variation and, in turn, little predictive utility.

Census populations size is theoretically linked to population genetic diversity because of its association with long-term demographic trends. Empirical studies have verified this connection in some cases (Hague and Routman 2016; Grundler et al. 2019; Buffalo 2021). One study in butterflies did not recover this association, but found that body size better correlated with long-term demographic trends than contemporary census population size (Mackintosh et al. 2019). (Buffalo 2021) found strong phylogenetic signal in the correlation between census size and population genetic diversity, suggesting that lineage- or species-specific traits can alter the strength of the association. Habitat fragmentation has also been shown to affect population genetic diversity across a variety of taxa (Dixo et al. 2009; Osborne et al. 2014; Gouskov and Vorburger 2016; Zhai et al. 2019; Ferreira et al. 2022) and in meta-analyses (Lino et al. 2019; González et al. 2020) likely through the reduction of inter-population dispersal and reduction in overall habitat size and/or quality. Altogether, these studies suggest that predicting genetic diversity from environmental parameters is likely best accomplished within a relatively narrow taxonomic scope (where variation in lineage-specific determinants is low) and by examining aspects of the environment likely to affect long-term demographic trends and dispersal.

For obligate freshwater taxa, the constraints of environment on demography and

dispersal are often highly evident and easily measured, as they correspond to characteristics of lake and river networks. For riverine taxa, the effects of river network structure on communities has been extensively documented, where species richness is highest in mainstem reaches and species turnover is highest in headwater reaches (Altermatt, Seymour, and Martinez 2013; Muneeppeerakul et al. 2008; Hughes, Schmidt, and Finn 2009; Finn et al. 2011; Carrara et al. 2012) due to the relative isolation of headwater reaches from one another and the relative connectedness of mainstem reaches. Higher genetic diversity in fish populations in mainstem reaches than those in headwater reaches has also been observed, suggesting that some of the processes influencing community diversity may also act to affect diversity within species (Paz-Vinas and S. Blanchet 2015; Paz-Vinas, Loot, et al. 2015). The degree of branching, or complexity, of the network has also been shown to positively affect community diversity (Terui, S. Kim, et al. 2021). The influence of branching probability on diversity below the species level has not been documented, but could be expected due to the increased metapopulation connectivity and stability it creates (Fagan 2002; B. L. Brown and Swan 2010; Terui, Ishiyama, et al. 2018).

Here we use data on population genetic diversity from 165 populations of Yellowstone Cutthroat Trout (*Oncorhynchus virginalis bouvieri*), and publicly accessible environmental data, to assess the role of local environment in shaping the diversity of populations. Much of the Yellowstone Cutthroat Trout range was severely affected by the last glacial period (Licciardi and Kenneth L. Pierce 2018), and so the differentiation between and within the subspecies' major lineages reflects recent divergence (within the last 14 thousand years, Rosenthal et al. in prep). Because of the young age of this taxon and previous evidence for phylogenetic signal in species-level genetic diversity (Buffalo 2021), we included a phylogenetic mixed-effects modeling framework to account for the potential for shared history to confound the effects of environment on population genetic diversity. We also model the evolution of population genetic diversity across a phylogeny to better characterize its variation.

4.2 Materials and Methods

Population genetic diversity, admixture, and phylogenetic relationships

We utilized previously generated estimates of population genetic diversity from Rosenthal et al. in prep for 231 populations of Rocky Mountain Cutthroat Trout (*O. virginalis*), the vast majority of which were from wild Yellowstone Cutthroat Trout (*O. v. bowvieri*) populations. Genetic diversity was estimated by Watterson’s theta (θ_W), Tajima’s pi (π ; nucleotide diversity) and observed heterozygosity. For each population, we previously calculated estimated admixture with known hatchery sources and the genetic diversity estimates for these sources (Rosenthal et al in prep), which were generated from genotyping-by-sequencing data using the program **ANGSD** (Korneliussen, Albrechtsen, and Nielsen 2014). Lastly, we utilized the phylogeny presented Rosenthal et al. in prep, but retained only individuals from within Yellowstone Cutthroat Trout. The phylogeny was generated from genotyping-by-sequencing data using **CASTER** (Zhang, Nielsen, and Mirarab 2025) and **IQtree** (Minh et al. 2020). To summarize similarity in different diversity metrics, we calculated Pearson’s correlation coefficients for each pair of diversity metrics.

Environmental data

We gathered environmental data from two publicly-available sources. First, we downloaded 30 year historic average air temperature data from PRISM (PRISM Group, Oregon State University 2023) for our entire study area. We used the average temperatures for June, July, and August as our measure of average summer air temperature. We included summer temperature because high temperatures might impose limitations on a cold-water fish like Yellowstone Cutthroat Trout. We next downloaded the National Hydrography Dataset Plus High Resolution (U.S. Geological Survey 2018) to collect data on average flowline slope and network branching probability. Flowline slope was included as high water

velocity in steep drainages may impose limits on trout populations. We included branching probability as a metric of stream network complexity, and calculated it by fitting an exponential distribution to the histogram of branch lengths, as per (Terui, S. Kim, et al. 2021). Network complexity has been associated with fish species richness (Terui, S. Kim, et al. 2021), and the effects of this might be seen at lower levels of biological organization as well. Data was aggregated at the HUC10 and HUC12 hydrologic basin scale, such that all represented populations within a given HUC would have the same environmental data in our subsequent models.

Position in the watershed has also been demonstrated to affect population genetic diversity (Paz-Vinas and S. Blanchet 2015), so we calculated the drainage area for population sampling points using a digital elevation map and functions from the `whiteboxtools` R package (Lindsay 2016; Wu and A. Brown 2025).

Models of population genetic diversity

First, we explored environmental determinants of genetic diversity for those Yellowstone Cutthroat Trout populations inferred to have no admixture and represented in the Rosenthal et al. in prep phylogeny. We used phylogenetic mixed-effects models as implemented in `brms` to construct Bayesian mixed-effects models (Bürkner 2017). We set prior probability distributions for all model-estimated coefficients as normal distributions with mean zero and standard deviation 10, distributions for all intercepts as normal distributions with mean zero and standard deviation 50, and distributions for random effect standard deviations as a Student's t distribution with mean zero and standard deviation 20. For all models, we considered the model converged if it experienced no divergent transitions, had $\hat{R}$ values of 1.0 for each parameter, and has effective sample sizes over 2,000 for each parameter. We fit models across two spatial scales of environmental data (HUC10 and HUC12) and three genetic diversity metrics (θ_W , π , and observed heterozygosity), and used a gamma distribution for the response variable. We explored

several model structures, as some combinations of predictor variables led to model identifiability issues. Variables explored included average summer temperature, branching probability, and upstream drainage area. All models included a random intercept for each population. We scaled and centered the data for winter temperature and log-transformed and scaled the data for upstream drainage area before model fitting. We generated a variance-covariance matrix for the random effect from the phylogeny of Rosenthal et al. in prep using the `vcv` function in `ape` (Paradis and Schliep 2019).

Next, we constructed models utilizing data from both admixed and unadmixed populations, though we did remove populations with less than 10% native ancestry. To account for the effects of admixture on genetic diversity, we built a model that accounted for these effects directly, structured as follows. Note that we only fit sets of environmental variables that showed no convergence issues for the models with no admixed populations.

$$Gendiv \sim exp(Gendiv_{orig}) * Prop_{native} + \sum_{i=1}^5 (Gendiv_{nonnative_i} * Prop_{nonnative_i}) \quad (4.1)$$

$$Gendiv_{orig} \sim EnvironmentalVariables + HUC2 + Intercept + (1|HUC8) \quad (4.2)$$

Here we incorporate the proportion non-native ancestry assigned to each of five non-native lineages, as inferred by Rosenthal et al. in prep, as well as the genetic diversity of those lineages, as a modification to the “original” genetic diversity of the population. Modification is modeled as a mixing of genetic material with differing diversities, such that the final diversity represented a weighted mean diversity between the original diversity and all admixture sources, with weights determined by admixture proportions ($Gendiv_{nonnative}$ in equation 1). We then estimate the “original” genetic diversity of the population as a function of environmental variables and include terms to reflect the effects of history/relatedness (through HUC2 and HUC8 location). We scaled and centered the data for flowline density and winter temperature prior to model fitting. We estimated an effect

for the HUC2 drainage that each population is within, as in this case it signifies the side of the Continental Divide of the Americas that the population is from. Dispersal across this divide has been associated with genetic diversity decreases in other cutthroat trout (Kovach et al. 2022). Lastly, we included a random intercept associated with the HUC8 hydrologic basin for each population to account for the effects of shared history and geographic proximity. We fit these models with all three genetic diversity metrics as response variables and at HUC10 and HUC12 scales for environmental data. We assessed model performance via posterior predictive checks. We compared model fit using Bayesian R^2 as implemented in `brms` (Bürkner 2017).

We also fit models for unadmixed populations in a phylogenetic generalized least squares (PGLS) framework using the R package `nlme` (Pinheiro, Bates, Team, et al. 2025). Because gamma regression is not supported by `nlme`, we instead centered and scaled the response variable and used a normally-distributed response. These models included no random effects but included summer air temperature, upstream drainage area, and branching probability.

Trait evolution models, phylogenetic signal, and diversity across major basins

We used the phylogenetic tree from Rosenthal et al. in prep and accompanying genetic diversity estimates to model the evolution of each genetic diversity metric on the tree. Though genetic diversity is not typically viewed as a trait that can evolve, as it is more of an emergent property of a lineage, it can change through time and the variants that constitute it are transmitted through generations. We fit under Brownian Motion, single optimum Ornstein-Uhlenbeck, and Early Burst modes of evolution using the R package `geiger` (Pennell et al. 2014). We compared model fit via AICc, and visualized the best fitting model on our tree using `phytools` (Revell 2024).

We calculated the phylogenetic signal of genetic diversity metrics using `phytools`. We

calculated Pagel’s λ (Pagel 1999) and Blomberg’s K (Blomberg, Garland JR., and Ives 2003) for each metric, and tested for significant differences from null expectations using the methods implemented in `phytools`. For Pagel’s λ , the null expectation is that $\lambda = 0$; for Blomberg’s K, the null expectation is that the estimated value of K is equal to the distribution of K estimates from 1,000 randomizations of the tree tips. Lastly, we tested for differences in genetic diversity across the Continental Divide of the Americas in both admixed and unadmixed populations using t-tests for all three genetic diversity metrics.

4.3 Results

Models of genetic diversity

Environmental data for average summer temperatures and branching probability were partially correlated (Figures 4.1 - 4.2, correlation coefficients of -0.53 for HUC 10 and -0.35 for HUC 12). Population genetic diversity estimates all showed relatively left-skewed distributions (Figure 4.3). Observed heterozygosity was not highly correlated with either θ_w or π (correlation coefficient estimates of 0.251 and 0.155, respectively), but θ_w and π were highly correlated with one another (correlation coefficient estimate of 0.916). We recovered effects of upstream drainage area (positive) and summer temperature (negative) on genetic diversity in models for θ_w and π at both the HUC10 and HUC12 levels, but only when branching probability was excluded from the model (Table 4.1). We attribute this to the partial correlation between branching probability and summer temperature. Models fit using a PGLS framework showed almost no statistically significant results (Table 4.2). Models for observed heterozygosity did not recover statistically significant (or nearly statistically significant) effects of any environmental variable. Bayesian R^2 values were similar for both θ_w models (0.466 for HUC10, 0.512 for HUC12), and for both π models (0.442 for HUC10, 0.502 for HUC12). Models including admixed populations were able to retain significantly more populations (122 versus 84 for the unadmixed populations

model) but did not recover significant effects of environmental predictors for any model.

Trait evolution models, phylogenetic signal, and diversity across major basins

Single optimum Ornstein-Uhlenbeck models were best supported by AICc for all three genetic diversity metrics (Figure 4.4). Models for θ_W and π had relatively high values of α , 4.440 and 6.386, respectively, whereas α was estimated at only 0.0185 for observed heterozygosity. A high value of α indicates that the attraction of values towards the central value through time is relatively strong (Brownian Motion would have $\alpha = 0$). Values for Blomberg's K were below 1.0 for each genetic diversity metric examined (0.2611 for θ_W , 0.2483 for π , and 0.5998 for observed heterozygosity) though only observed heterozygosity significantly differed from the null hypothesis (p-value = 0.031). For Pagel's λ , observed heterozygosity again showed different patterns from θ_W and π . The estimate of λ for observed heterozygosity was 1.0035 and significantly different from zero (p-value = 2.141×10^{-7}), whereas the estimates for θ_W and π were both low (7.506×10^{-5}) and not significantly different from zero. T-tests showed significant differences in genetic diversity between populations on either side of the Continental Divide of the Americas only for observed heterozygosity when all populations were included and when only populations with < 10% non-native ancestry were included, though the difference in θ_W was nearly significant (p-value = 0.08) when only unadmixed populations were compared (Figure 4.5).

4.4 Discussion

Identifying environmental predictors of population genetic diversity holds promise in enabling the rapid assessment of adaptive potential across large geographic areas.

Understanding how environmental variables contribute to population genetic diversity also stands to further our understanding of how ecological processes may translate into diversity

at higher levels of biological organization (Vellend 2005; Vellend and Geber 2005). We use population genetic diversity and phylogenetic data from populations of Yellowstone Cutthroat Trout to assess correlations between genetic diversity and local environment, and to more broadly examine the variation in genetic diversity across populations.

Genetic diversity and local environment

We identified a negative association between summer air temperatures and a positive association between upstream drainage area (a measure of watershed position) with two of our three genetic diversity metrics. Summer air temperatures likely decrease genetic diversity through demographic effects, as Yellowstone Cutthroat Trout are a cold-water species that can be negatively affected by high temperatures (Isaak and Hubert 2004). Upstream drainage area, which reflects how far from the headwaters a population was sampled, showed a positive effect on genetic diversity, as expected from previous studies observing a downstream increase in genetic diversity across a number of aquatic taxa (Paz-Vinas, Loot, et al. 2015; Simon Blanchet et al. 2020). The effects of both environmental variables on genetic diversity are rather small, but the overall limited range of θ_W and π observed in these populations will inherently limit their potential effect sizes. Predictive performance of our models was also relatively low as indicated by Bayesian R^2 values below 0.30. Given the difficulties of predicting genetic diversity at the species level (Glémin, Bazin, and Charlesworth 2006; Romiguier et al. 2014; Buffalo 2021; De Kort et al. 2021), our identification of any statistical association is notable. Yellowstone Cutthroat Trout have only inhabited the majority of their range since the Pinedale Glaciation receded approximately 14 thousand years ago (Licciardi and Kenneth L. Pierce 2018), and so population diversity might still primarily reflect the recent history of colonization rather than the influences of local environment.

We do not observe statistically significant associations with any environmental variable when admixed populations are included, using a model intended to account for the

effects of introgression on genetic diversity. Because the model with admixed populations includes over twice as many observations as the model without admixed populations yet fails to identify the same effects as the smaller model, this suggests that the model does not adequately account for the effects of admixture on population genetic diversity.

Alternatively, there may be some relationship between susceptibility to anthropogenic admixture and environmental effects of genetic diversity, but there exists no theoretical or empirical evidence to suggest this is the case.

Differences in sensitivity to demographic changes could explain why we observe correlations between θ_W and π with the local environment but not for observed heterozygosity. Heterozygosity is more susceptible to sudden population contractions and inbreeding (Wright 1931; Nonaka et al. 2019), which is how Rosenthal et al. explain the lack of high heterozygosity in captive, introduction source populations that have high θ_W and π . Previous examination of genetic diversity in mammals found that θ_W was highly correlated with both π and observed heterozygosity (Jeon et al. 2024), though we only recover strong correlations between θ_W and π . Furthermore, populations with high heterozygosity all show high θ_W values, whereas the inverse is not true (Figure 4.4, suggesting that heterozygosity is more sensitive. The relatively recent declines of many Yellowstone Cutthroat Trout populations (Al-Chokhachy et al. 2017) could explain decreased observed heterozygosity relative to θ_W and π for these wild populations.

Trait evolution models and diversity across major basins

Similar to the results for models of genetic diversity and local environment, observed heterozygosity behaved slightly differently across the phylogeny of our unadmixed populations than either θ_W or π (Figure 4.4). Phylogenetic signal suggested low variability compared to a Brownian Motion expectation for all three genetic diversity metrics, though only observed heterozygosity appeared to have variation associated with the phylogeny ($\lambda \simeq 1$, $K = 0.556$, p-value for Blomberg’s K randomization test $\ll 0.05$). All metrics of

diversity were best modeled via an Ornstein-Uhlenbeck process, which is consistent with bounded variation. The distributions for all three diversity metrics show a small number of lineages with significantly higher diversity in all three metrics (Figure 4.3, Figure 4.4), which could explain why the diversity is best modeled by bounded variation rather than unbounded variation with a low rate.

Differences in genetic diversity across the Continental Divide of the Americas were small and statistically insignificant in most cases tested. Only observed heterozygosity showed the expected pattern of lower diversity in populations on the East side of the divide, and showed this pattern across all populations and for only unadmixed populations (Figure 4.5). The stronger phylogenetic signal for observed heterozygosity is likely associated with this result, as populations on either side of the divide form two of the subspecies' major clades. The genetic diversity of Yellowstone Cutthroat Trout appears less dramatically affected by their colonization of the Missouri River drainage than their Westslope Cutthroat Trout (*O. lewisi*) counterparts, though extreme fragmentation of Missouri River drainage Westslope Cutthroat Trout likely exacerbate the bottleneck effect of colonization (Kovach et al. 2022). Alternatively, the mode of dispersal across the continental divide could also play a role: Yellowstone Cutthroat Trout were able to utilize Two-Ocean Pass, which is still passable by fish today and may have been for thousands of years, whereas Westslope Cutthroat Trout are hypothesized to have used Glacial Lake Missoula, a long-gone glacial lake that existed in the Pleistocene (Drinan et al. 2011; Young et al. 2018).

Conclusions

We identified statistically significant associations between genetic diversity metrics and environmental variables, specifically the amount of upstream drainage area and summer air temperatures. Identifying these associations has implications for conservation and restoration prioritization, as areas predicted to have higher potential genetic diversity could

be prioritized. We only observed significant associations for models including only populations without admixture, suggesting that our understanding of how admixture affects genetic diversity is incomplete. Estimates of phylogenetic signal differed between the three genetic diversity metrics examined, where observed heterozygosity showed the strongest, but still weak, phylogenetic signal. Lastly, differences in genetic diversity between populations on the East and West sides of the Continental Divide of the Americas were only statistically significant for observed heterozygosity.

4.5 Figures

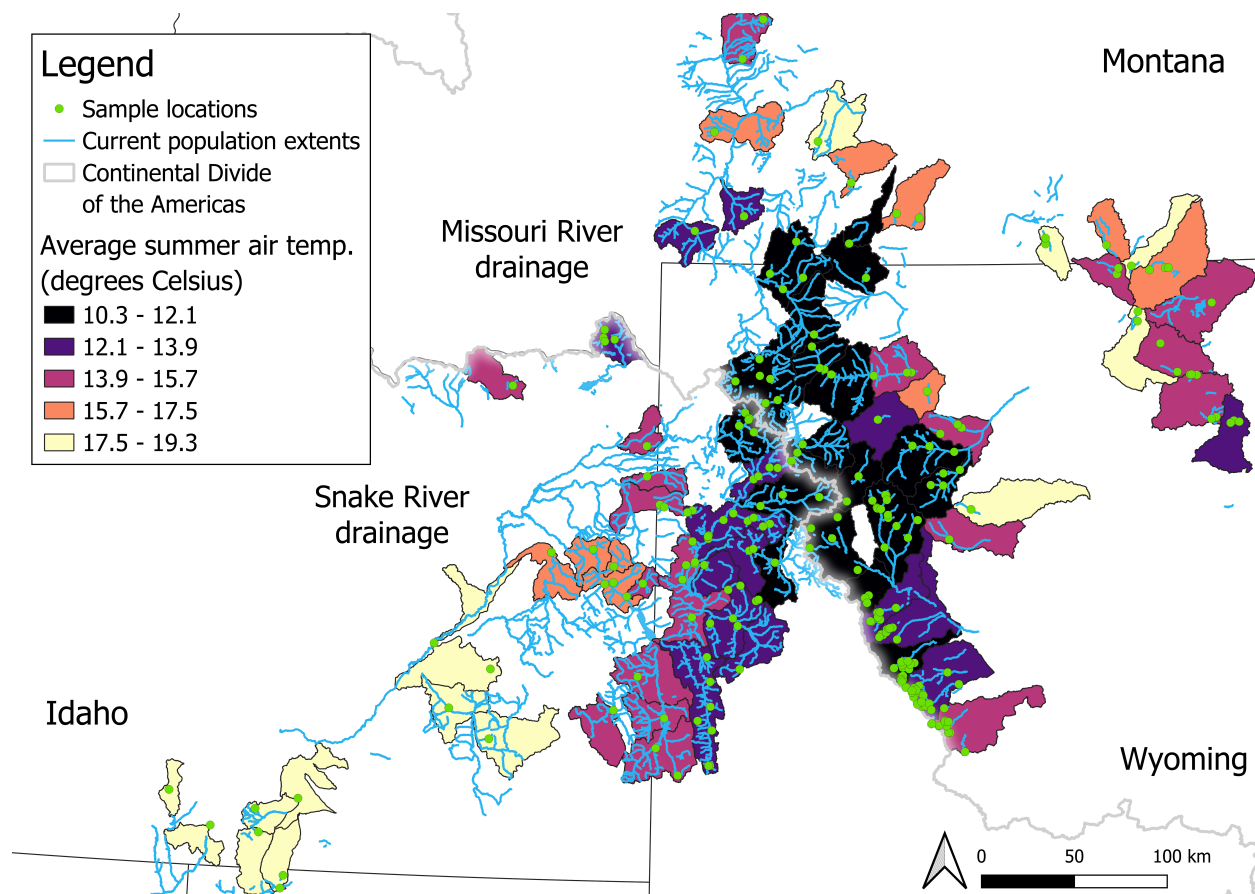


Figure 4.1: Map of sample locations and data on average historic summer air temperature aggregated at the HUC10 scale. The gray line denotes the Continental Divide of the Americas.

4.6 Tables

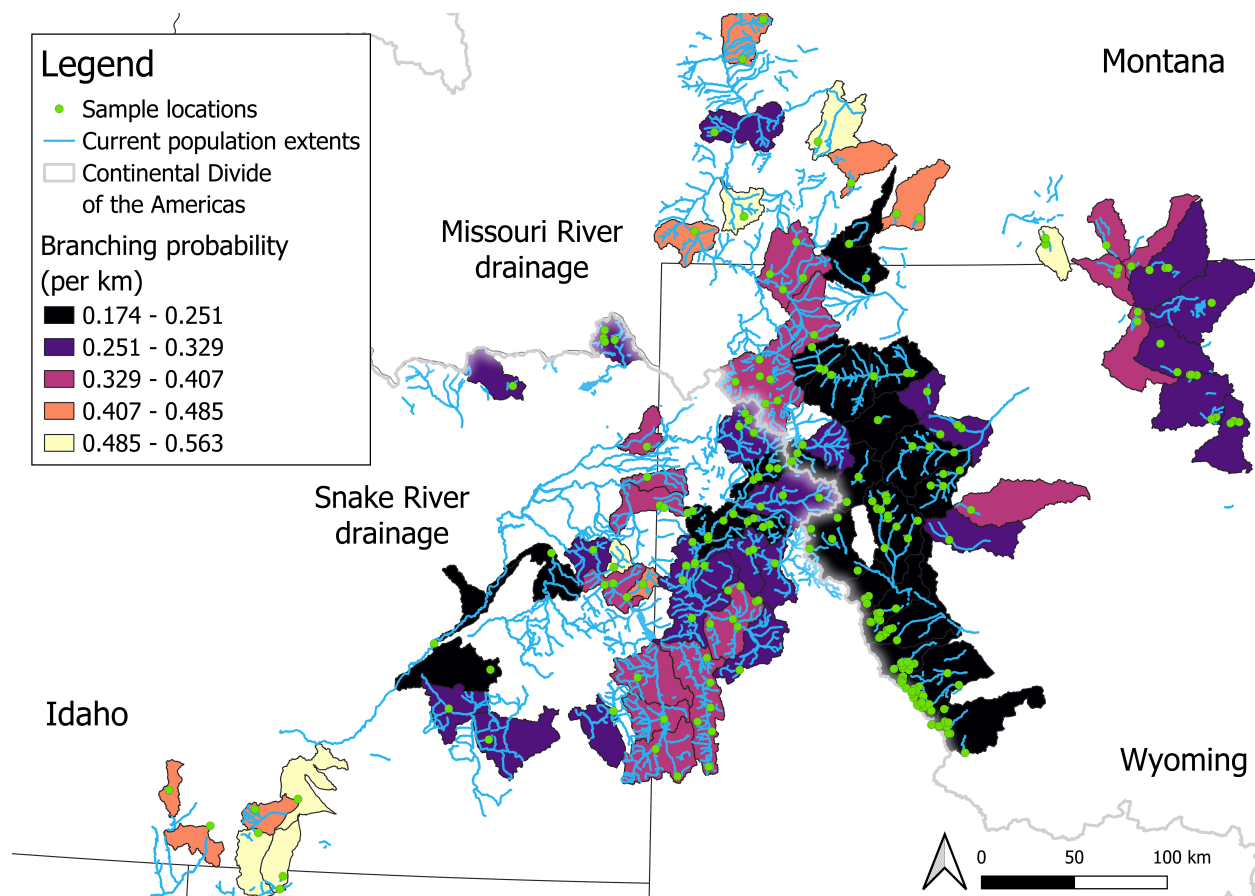


Figure 4.2: Map of sample locations and data on stream network branching probability aggregated at the HUC10 scale. The gray line denotes the Continental Divide of the Americas.

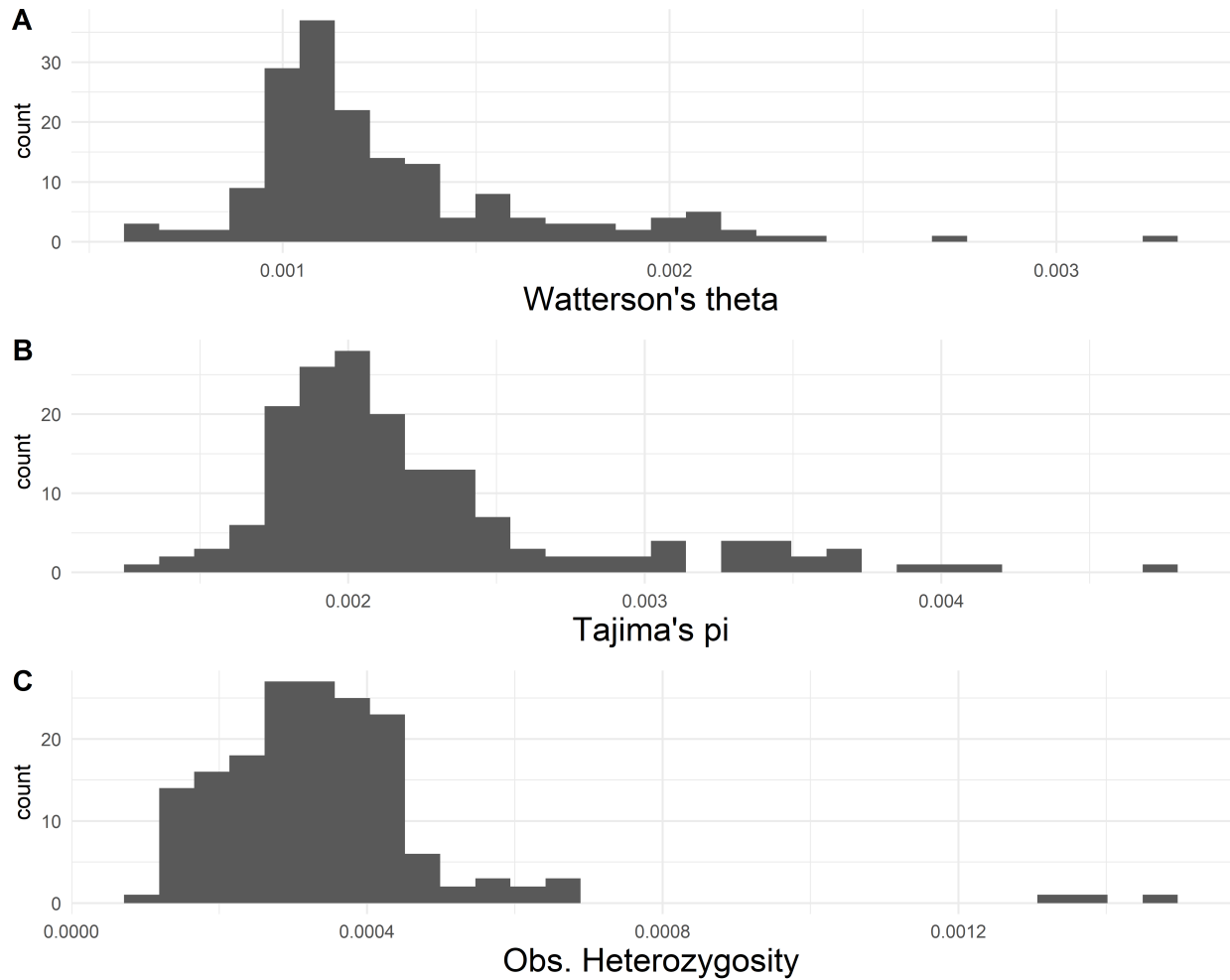


Figure 4.3: Distributions of population genetic diversity estimates for the populations included in this study (without removal of admixed populations). Panel A shows Watterson's theta (θ_W), panel B shows Tajima's pi (π , nucleotide diversity), and panel C shows observed heterozygosity.

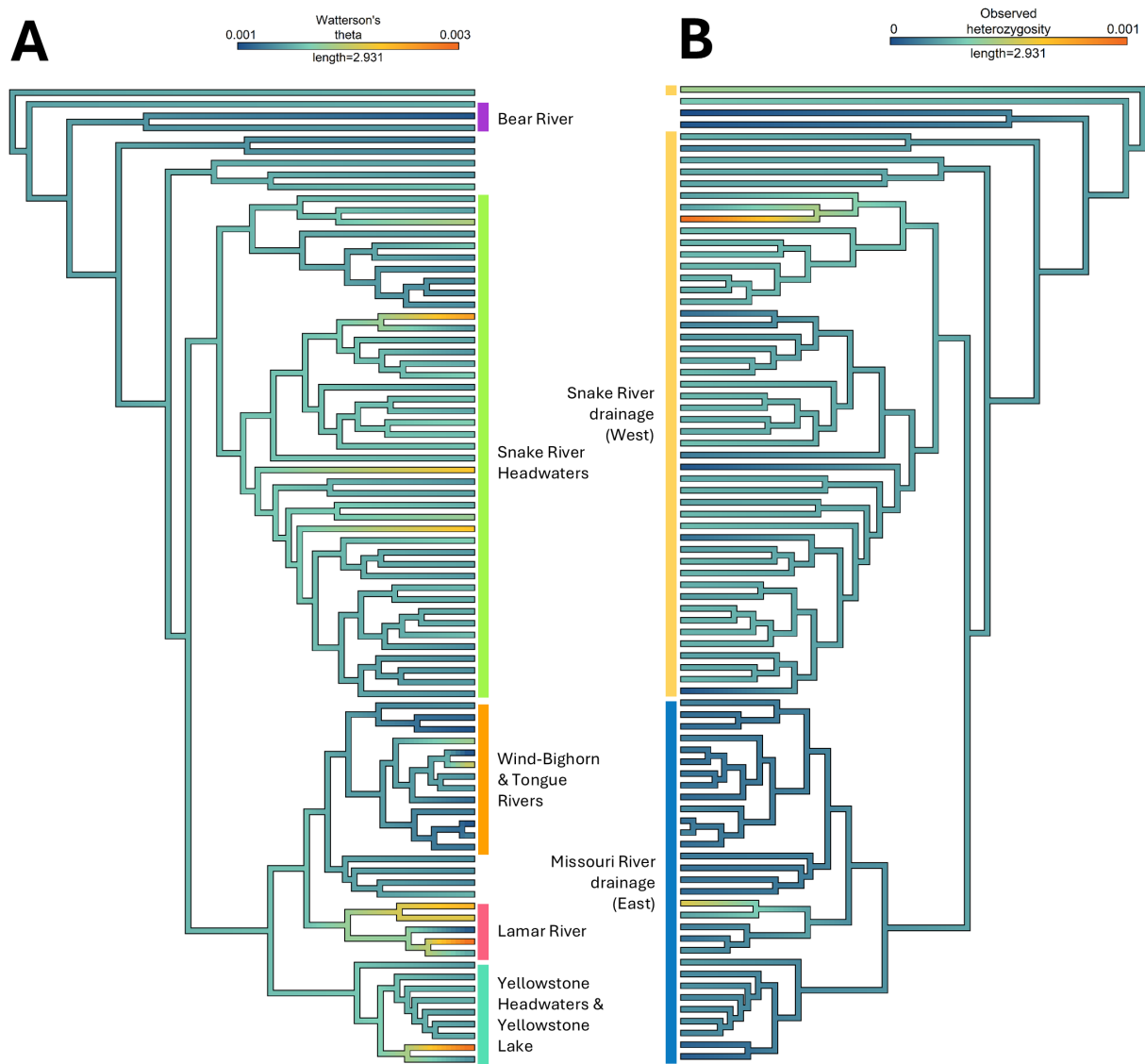


Figure 4.4: Phylogenetic trees showing results from models of continuous trait evolution under an Ornstein-Uhlenbeck process. Branch color corresponds to model-reconstructed diversity values. Panel A shows results for Watterson's theta θ_W and panel B shows results for observed heterozygosity. Phylogeny is from Rosenthal et al. in prep and colored bars on tips correspond to the hydrologic basin of origin for adjacent tips at two different levels of organization. Tree topologies are identical.

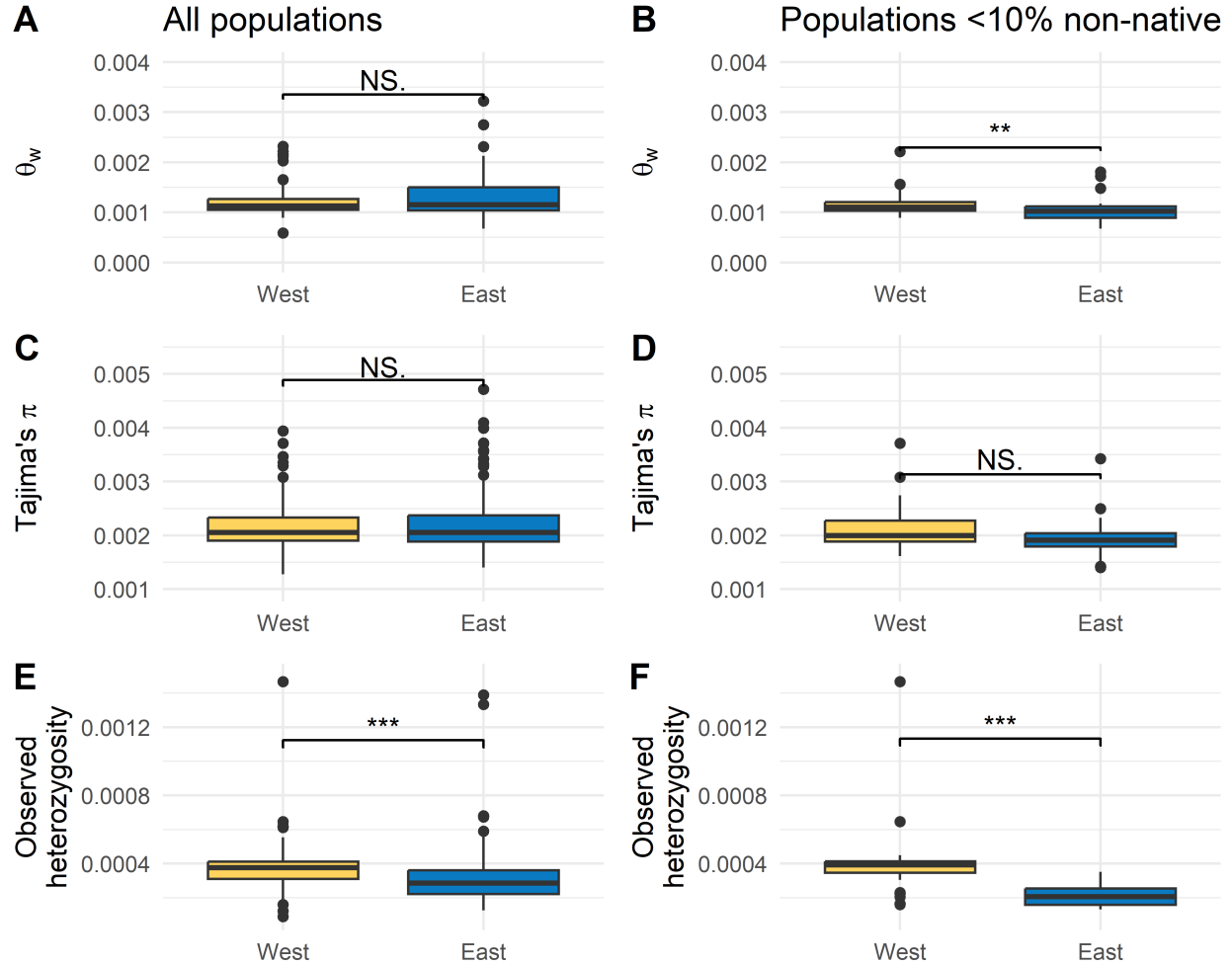


Figure 4.5: Results from t-tests comparing the genetic diversity of populations on either side of the Continental Divide of the Americas across three genetic diversity metrics. Panels A, C, and E show results when all populations are included, and panels B, D, and F show results when only populations with < 10% non-native ancestry are included.

Genetic diversity metric	Geographic scale	Parameter	Estimate	Lower 95% C.I. bound	Upper 95% C.I. bound	Prob. of effect
θ_W	HUC10	Intercept*	-6.67	-7.02	-6.26	
θ_W	HUC10	Upstream area*	0.07	0.02	0.13	98.41%
θ_W	HUC10	Summer temp.	-0.07	-0.13	0.00	97.33%
θ_W	HUC12	Intercept*	-6.64	-7.02	-6.22	
θ_W	HUC12	Upstream area*	0.07	0.02	0.13	99.67%
θ_W	HUC12	Summer temp.*	-0.09	-0.15	-0.02	99.27%
π	HUC10	Intercept*	-6.11	-6.41	-5.77	
π	HUC10	Upstream area*	0.05	0.01	0.10	98.78%
π	HUC10	Summer temp.	-0.03	-0.10	-0.03	87.77%
π	HUC12	Intercept*	-6.09	-6.41	-5.72	
π	HUC12	Upstream area*	0.06	0.01	0.10	99.14%
π	HUC12	Summer temp.	-0.05	-0.11	0.01	94.67%

Table 4.1: Results from Bayesian generalized linear mixed-models of Yellowstone Cutthroat Trout populations without admixture from anthropogenic translocations. Asterisks indicate parameters with mean 95% credible intervals that do not overlap zero. Probability of effect, defined as the proportion of the posterior distribution for that parameter above 0.00, is also reported for coefficients for environmental variables. The summer temperature covariate was centered and scaled, and the upstream area covariate was log-transformed and then centered and scaled.

Genetic diversity metric	Geographic scale	Parameter	Estimate	p-value
θ_W	HUC10	Intercept	-10.83	0.10
θ_W	HUC10	Upstream area	0.06	0.18
θ_W	HUC10	Summer temp.	-0.12	0.5
θ_W	HUC10	Branching prob.	10.58	0.11
θ_W	HUC12	Intercept	-2.16	0.22
θ_W	HUC12	Upstream area	0.04	0.33
θ_W	HUC12	Summer temp.*	-0.36	0.02
θ_W	HUC12	Branching prob.	2.37	0.05
π	HUC10	Intercept	-10.19	0.13
π	HUC10	Upstream area	0.6	0.21
π	HUC10	Summer temp.	-0.09	0.61
π	HUC10	Branching prob.	9.92	0.15
π	HUC12	Intercept	-1.98	0.28
π	HUC12	Upstream area	0.41	0.38
π	HUC12	Summer temp.	-0.29	0.07
π	HUC12	Branching prob.	2.10	0.09

Table 4.2: Results from PGLS models of Yellowstone Cutthroat Trout populations without admixture from anthropogenic translocations. Asterisks indicate parameters with p-values < 0.05 . The summer temperature covariate was centered and scaled, and the upstream area covariate was log-transformed and then centered and scaled. Note that these models were fit with the response variables scaled and centered, and also fit without the Gamma distribution of the Bayesian phylogenetic generalized linear mixed models, and therefore the magnitudes of coefficients are not directly comparable.

Chapter 5

Management recommendations and data availability

Based on the findings of this dissertation, several findings are particularly relevant to natural resource managers, especially those concerned with Yellowstone cutthroat trout populations. I will discuss points relevant to stocking practices and conservation prioritization before mentioning how and where data and samples used in this dissertation can be accessed.

Key Management recommendations

1. Preserve existing basin-level genetic structure.
2. Avoid unnecessary replacement of extant populations merely because they contain hatchery ancestry.
3. Use geographically and genetically appropriate sources for stocking and conservation translocations.
4. Identify small, isolated, and low-diversity populations and potential translocation donors for those populations before they reach a crisis.
5. Preserve tissue and genomic resources for future use to continue assisting decision making.

In chapter 2, there is a clear effect of historic stocking practices spreading ancestry from the Snake and Yellowstone basins outside of their native range. For example, the LeHardy Rapids strain, which contains Yellowstone River ancestry, is maintained despite the majority of stocking for Yellowstone cutthroat trout occurring outside of the Yellowstone River basin. To avoid further anthropogenic homogenization, developing hatchery broods with ancestry from the same basin as the likely stocking location would be prudent. If a new brood is to be developed, I would recommend choosing a source population with moderate to high genetic diversity, or, better yet, founding the brood from a mix of different populations from across the basin of interest (Bell, Kovach, Matthew R. Campbell, et al. [2026](#)). If stocking for conservation purposes is being undertaken, I would recommend targeted translocations from a carefully selected donor population (or a mix of several donors if one suitable donor is unavailable), chosen based on genetic diversity, ancestry, and potential for being locally adapted to the introduction area.

Chapter 2 also highlights an important point regarding conservation prioritization: populations with hatchery admixture are likely nearly as well-adapted to their local environment as their unadmixed counterparts. I would therefore recommend prioritizing the expansion of existing populations/metapopulations or founding new populations rather than replacing existing populations containing hatchery ancestry from within *O. virginalis* bouvieri. Management priorities could change if a particular lineage receives formal conservation recognition: for example, Bighorn basin vs. Yellowstone basin ancestry could have major implications for protection status. Chapter 3 highlights how cryptic diversity within Yellowstone cutthroat trout can be and also demonstrates how most drainage basins (at or below the HUC 8 level) have distinct ancestry that is very detectable through modern genomic methods. Because the relatedness of Yellowstone cutthroat trout populations seems to follow the hierarchical nature of hydrologic basins, ancestry should be protected at the smallest level possible. For example, even if ancestry within the same HUC 8 drainage cannot be used for a conservation initiative, using the next most closely

related appropriate source should be prioritized. Each of these distinct ancestries may reflect adaptation to the local environment in addition to unique genetic variation and evolutionary history. I believe that protecting these genomic resources should be priority, even if some hatchery admixture has affected the population. Lastly, chapter 4 identifies that populations in drainages with higher summer temperatures and in more headwater reaches have lower genetic diversity than their cooler or less isolated counterparts. For conservation populations meeting either of these criteria, I recommend managers attempt to monitor the population closely for signs of inbreeding depression, such as poor population recruitment, sudden declines, or morphological abnormalities: genetic drift will continue to decrease the diversity of these populations over time unless they are of sufficient size. Identifying potential donor populations for genetic rescue of these populations well in advance may also be prudent. When identifying potential genetic rescue donors, it is important to note that rescue from inbred donor populations can be effective, meaning that even small populations can be considered (Bell, Kovach, Robinson, et al. [2026](#)).

Tissue samples for the majority of fish from this project are kept by the Wagner Lab at the University of Wyoming: contact the Wagner Lab for access to these and any other data or samples held there. Tissue samples from the Idaho Department of Fish and Game, Montana Fish, Parks and Wildlife, and Pisces Molecular are not archived by the Wagner Lab (see Table [2.1](#) for population sources). Extracted DNA is also available at the University of Wyoming, though the majority of samples from MTFWP and IDFG do not have extract remaining: contact those departments for access to their samples. The majority of samples from the following populations also do not have extracted DNA remaining: Rawhide Creek, Dick Creek, Piney Creek, Pickett Creek, Greybull River, Wood River, Ishawooa Creek, Upper SF Shoshone River, Marquette Creek, Trapper Creek, S Beaver Creek, N Beaver Creek, Soldier Creek, Muddy Creek, Timber Creek, Stonefly Creek, Bears Ears Lake, E Fk of E Fk Wind River, E Fk Wind River, Bear Creek, Sheridan Creek, Cub Lake, Moose Lake, Dons Lake, Lower Glacier Lake, Phillips Lake,

Horseshoe Lake, Middle Fork Dry Creek, Grassy Lake, Native Lake, Double Lake, Twin Lake, Dinwoody Creek (and all collections within), Honeymoon Lake, Blueberry Lake, Marked Tree Lake, Hidden Lake, Upper Ross Lake, Baptiste Creek, Grave Lake, Stough Creek, Frontier Creek, Baptiste Lake, Baptiste Lake Creek, Little Warm Springs Creek, Papoose Lake, Unnamed Graves Lakes, and Raft Creek. All generated sequence data is archived at the University of Wyoming on its Alcova cluster, which is accessible with permission from Katie Wagner, and will be uploaded to the NCBI Short Read Archive upon acceptance of manuscripts from this dissertation work to a peer-reviewed journal. Photographs provided by sample collectors are also stored by the Wagner Lab.

Bibliography

- Alexander, David H., John Novembre, and Kenneth Lange (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19(9), pp. 1655–1664. (Cit. on pp. 14, 22, 122).
- Alger, Ryan et al. (2018). *Yellowstone cutthroat trout 2018 assessment review data*. URL: <https://app.streamnet.org/files/697/697.zip> (visited on 09/05/2025). (Cit. on pp. 123, 129).
- Alt, David D and Donald W Hyndman (1989). *Roadside geology of Idaho*, Mountain Press. (Cit. on p. 130).
- Altermatt, Florian, Mathew Seymour, and Nicolas Martinez (2013). River network properties shape -diversity and community similarity patterns of aquatic insect communities across major drainage basins. *Journal of Biogeography* 40(12), pp. 2249–2260. (Cit. on p. 143).
- Araki, Hitoshi et al. (2008). Fitness of hatchery-reared salmonids in the wild. *Evolutionary Applications* 1(2), pp. 342–355. (Cit. on p. 29).
- Barreto, Silvia Britto et al. (2022). Drainage rearrangements and in situ diversification of an endemic freshwater fish genus from north-eastern Brazilian rivers. *Freshwater Biology* 67(5), pp. 759–773. (Cit. on p. 116).
- Behnke, Robert (2010). *Trout and salmon of North America*, Simon and Schuster. (Cit. on p. 32).
- Behnke, Robert J (1992). Native trout of western North America. *Am Fish Soc Monogr* 6, pp. i–xx+. (Cit. on pp. 7, 8, 31, 117, 118).
- Bell, Donovan A., Ryan P. Kovach, Matthew R. Campbell, et al. (2026). Long-term genetic outcomes of a mixed-source introduction of Westslope Cutthroat Trout. *Canadian Journal of Fisheries and Aquatic Sciences* 83, pp. 1–10. (Cit. on p. 162).
- Bell, Donovan A., Ryan P. Kovach, Zachary Robinson, et al. (2026). Inbred source populations result in genetic rescue of imperiled trout populations. *Biological Conservation* 314, p. 111677. (Cit. on p. 163).
- Benjamini, Yoav and Yosef Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57(1), pp. 289–300. (Cit. on p. 21).
- Bergeron, Lucie A. et al. (2023). Evolution of the germline mutation rate across vertebrates. *Nature* 615(7951), pp. 285–291. (Cit. on p. 122).
- Bermingham, E. and A. P. Martin (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology* 7(4), pp. 499–517. (Cit. on p. 116).

- Bernatchez, Louis and Julian J. Dodson (1987). Relationship between Bioenergetics and Behavior in Anadromous Fish Migrations. *Canadian Journal of Fisheries and Aquatic Sciences* 44(2), pp. 399–407. (Cit. on p. [117](#)).
- Besnier, Francois et al. (2022). Introgression of domesticated salmon changes life history and phenology of a wild salmon population. *Evolutionary Applications* 15(5), pp. 853–864. (Cit. on p. [28](#)).
- Blanchet, Simon et al. (2020). A river runs through it: The causes, consequences, and management of intraspecific diversity in river networks. *Evolutionary Applications* 13(6), pp. 1195–1213. (Cit. on p. [150](#)).
- Bland, Lucie M. et al. (2015). Predicting the conservation status of data-deficient species. *Conservation Biology* 29(1), pp. 250–259. (Cit. on p. [141](#)).
- Blomberg, Simon P., Theodore Garland JR., and Anthony R. Ives (2003). Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution* 57(4), pp. 717–745. (Cit. on p. [148](#)).
- Brekke, Cathrine and Tim Martin Knutsen (2026). Phenotypic sex determines recombination rate and distribution in sex-reversed rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 612, p. 743143. (Cit. on p. [21](#)).
- Brown, B. L. and C. M. Swan (2010). Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology* 79(3), pp. 571–580. (Cit. on p. [143](#)).
- Bruce, Spencer A. et al. (2020a). Stocked-Fish Introgression into Wild Brook Trout Populations Depends on Habitat. *Transactions of the American Fisheries Society* 149(4), pp. 427–442. (Cit. on p. [28](#)).
- Bruce, Spencer A. et al. (2020b). Stocked-Fish Introgression into Wild Brook Trout Populations Depends on Habitat. *Transactions of the American Fisheries Society* 149(4), pp. 427–442. (Cit. on p. [29](#)).
- Buckley, Sean James et al. (2024). A community-driven captive-breeding and reintroduction program maintains genetic diversity in a threatened freshwater fish. *Conservation Science and Practice* 6(1), e13054. (Cit. on p. [5](#)).
- Buffalo, Vince (2021). Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin’s Paradox. *eLife* 10, e67509. (Cit. on pp. [142](#), [143](#), [150](#)).
- Bürkner, Paul Christian (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). (Cit. on pp. [19](#), [145](#), [147](#)).
- Bürkner, Paul Christian (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal* 10(1), pp. 395–411. (Cit. on p. [19](#)).
- Butt, Nathalie et al. (2021). Importance of species translocations under rapid climate change. *Conservation Biology* 35(3), pp. 775–783. (Cit. on p. [4](#)).
- Campbell, Matthew R. et al. (2011). Historical influences of volcanism and glaciation in shaping mitochondrial DNA variation and distribution in Yellowstone cutthroat trout across its native range. *Transactions of the American Fisheries Society* 140(1), pp. 91–107. (Cit. on pp. [83](#), [123](#), [129](#), [130](#)).
- Campbell, MR et al. (2018). Describing and Preserving the Diversity of Cutthroat Trout in the Yellowstone River, Snake River, and Bonneville Basin. *American Fisheries Society Special Publication* 36, pp. 303–321. (Cit. on p. [15](#)).

- Carrara, Francesco et al. (2012). Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences* 109(15), pp. 5761–5766. (Cit. on p. 143).
- Carvalho, G. R. et al. (1991). Marked genetic divergence revealed by allozymes among populations of the guppy *Poecilia reticulata* (Poeciliidae), in Trinidad. *Biological Journal of the Linnean Society* 42(3), pp. 389–405. (Cit. on p. 5).
- Casanova, Adrián et al. (2022). Genomic Hatchery Introgression in Brown Trout (*Salmo trutta* L.): Development of a Diagnostic SNP Panel for Monitoring the Impacted Mediterranean Rivers. *Genes* 13(2), p. 255. (Cit. on p. 28).
- Cavin, Lionel (2017). 150 million years of freshwater fish biogeography: vicariance or dispersal? *Food Agricultural Sciences and Technology* 3(1), pp. 1–4. (Cit. on p. 116).
- Caye, Kevin et al. (2019). LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies. *Molecular Biology and Evolution* 36(4), pp. 852–860. (Cit. on pp. 20, 26).
- Chang, Christopher C. et al. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4(1), pp. 1–16. (Cit. on p. 14).
- Al-Chokhachy, Robert et al. (2017). Status & Conservation of Yellowstone Cutthroat Trout in the Greater Yellowstone Area. *Yellowstone Science* 25(1), pp. 13–17. (Cit. on pp. 28, 151).
- Christie, Mark R., Michael J. Ford, and Michael S. Blouin (2014). On the reproductive success of early-generation hatchery fish in the wild. *Evolutionary Applications* 7(8), pp. 883–896. (Cit. on p. 29).
- Clarke, Julia G., Adam C. Smith, and Catherine I. Cullingham (2024). Genetic rescue often leads to higher fitness as a result of increased heterozygosity across animal taxa. *Molecular Ecology* 33(19), e17532. (Cit. on pp. 4, 6).
- Claussen, Julie E. and David P. Philipp (2023). Assessing the role of supplementation stocking: A perspective. *Fisheries Management and Ecology* 30(6), pp. 583–591. (Cit. on p. 5).
- Coffman, Alec J. et al. (2016). Computationally Efficient Composite Likelihood Statistics for Demographic Inference. *Molecular Biology and Evolution* 33(2), pp. 591–593. (Cit. on p. 122).
- Comte, Lise and Julian D. Olden (2018a). Evidence for dispersal syndromes in freshwater fishes. *Proceedings of the Royal Society B: Biological Sciences* 285(1871), p. 20172214. (Cit. on p. 117).
- Comte, Lise and Julian D. Olden (2018b). Fish dispersal in flowing waters: A synthesis of movement- and genetic-based studies. *Fish and Fisheries* 19(6), pp. 1063–1077. (Cit. on p. 117).
- Corander, Jukka et al. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9(1), p. 539. (Cit. on p. 15).
- Cortés-Guzmán, Daniela et al. (2024). Dispersal, glacial refugia and temperature shape biogeographical patterns in European freshwater biodiversity. *Global Ecology and Biogeography* 33(9), e13886. (Cit. on p. 116).
- Danecek, Petr, Adam Auton, et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27(15), pp. 2156–2158. (Cit. on pp. 12, 120).

- Danecek, Petr, James K. Bonfield, et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10(2), pp. 1–4. (Cit. on pp. [11](#), [120](#)).
- De Kort, H. et al. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nature Communications* 12(1), p. 516. (Cit. on pp. [142](#), [150](#)).
- DeWoody, J. Andrew et al. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology* 30(17), pp. 4147–4154. (Cit. on p. [141](#)).
- Dias, Murilo S., Jean-François Cornu, et al. (2013). Natural fragmentation in river networks as a driver of speciation for freshwater fishes. *Ecography* 36(6), pp. 683–689. (Cit. on p. [116](#)).
- Dias, Murilo S., Thierry Oberdorff, et al. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters* 17(9), pp. 1130–1140. (Cit. on p. [116](#)).
- Dixo, Marianna et al. (2009). Habitat fragmentation reduces genetic diversity and connectivity among toad populations in the Brazilian Atlantic Coastal Forest. *Biological Conservation* 142(8), pp. 1560–1569. (Cit. on p. [142](#)).
- Drinan, Daniel P. et al. (2011). Genetic variation in westslope cutthroat trout *Oncorhynchus clarkii lewisi*: implications for conservation. *Conservation Genetics* 12(6), pp. 1513–1523. (Cit. on p. [152](#)).
- Eaton, Kendra R., Janet L. Loxterman, and Ernest R. Keeley (2018). Connections and containers: Using genetic data to understand how watershed evolution and human activities influence cutthroat trout biogeography. *PLOS ONE* 13(8), e0202043. (Cit. on p. [7](#)).
- Eldridge, W H, J M Myers, and K A Naish (2009). Long-term changes in the fine-scale population structure of coho salmon populations (*Oncorhynchus kisutch*) subject to extensive supportive breeding. *Heredity* 103(4), pp. 299–309. (Cit. on p. [29](#)).
- Eldridge, William H. and Kerry A. Naish (2007). Long-term effects of translocation and release numbers on fine-scale population structure among coho salmon (*Oncorhynchus kisutch*). *Molecular Ecology* 16(12), pp. 2407–2421. (Cit. on p. [29](#)).
- Ellegren, Hans and Nicolas Galtier (2016). Determinants of genetic diversity. *Nature Reviews Genetics* 17(7), pp. 422–433. (Cit. on p. [142](#)).
- Endicott, Carol et al. (2016). Range-wide Status Assessment for Yellowstone Cutthroat Trout *Oncorhynchus Clarkii* Bouvieri: 2012. *Yellowstone cutthroat trout interagency coordination group, Montana Fish, Wildlife and Parks: Bozeman, MT, USA*. (Cit. on pp. [1](#), [9](#), [22](#)).
- Erdman, Bradley et al. (2022). BROADSCALE Population Structure and Hatchery Introgression of Midwestern Brook Trout. *Transactions of the American Fisheries Society* 151(1), pp. 81–99. (Cit. on pp. [5](#), [28](#), [29](#)).
- Fagan, William F. (2002). Connectivity, Fragmentation, and Extinction Risk in Dendritic Metapopulations. *Ecology* 83(12), pp. 3243–3249. (Cit. on p. [143](#)).
- Ferreira, Dhiego G. et al. (2022). The role of free-flowing tributary rivers in the maintenance of genetic diversity of a migratory fish species living in a river fragmented by dams. *Hydrobiologia* 849(5), pp. 1221–1237. (Cit. on p. [142](#)).
- Finn, Debra S. et al. (2011). Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. *Journal of the North American Benthological Society* 30(4), pp. 963–980. (Cit. on p. [143](#)).

- Fitzpatrick, Sarah W. et al. (2020). Genomic and Fitness Consequences of Genetic Rescue in Wild Populations. *Current Biology* 30(3), 517–522.e5. (Cit. on pp. 4, 21, 31).
- Franke, Mary Ann (1997). A grand experiment. *Yellowstone Science* 5(8), p. 13. (Cit. on pp. 2, 7, 128).
- Frankham, R. (2012). How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. *Heredity* 108(3), pp. 167–178. (Cit. on p. 142).
- Frankham, Richard (2022). Evaluation of proposed genetic goals and targets for the Convention on Biological Diversity. *Conservation Genetics* 23(5), pp. 865–870. (Cit. on p. 6).
- Frichot, Eric and Olivier François (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* 6(8), pp. 925–929. (Cit. on pp. 20, 26).
- Frichot, Eric, Sean D. Schoville, et al. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution* 30(7), pp. 1687–1699. (Cit. on pp. 20, 26).
- Gao, Guangtu et al. (2021). A long reads-based de-novo assembly of the genome of the Arlee homozygous line reveals chromosomal rearrangements in rainbow trout. *G3* 11(4), jkab052. (Cit. on pp. 11, 22, 120).
- Gilk, Sara E et al. (2004). Outbreeding depression in hybrids between spatially separated pink salmon, *Oncorhynchus gorbuscha*, populations: marine survival, homing ability, and variability in family size. *Environmental Biology of Fishes* 69(1), pp. 287–297. (Cit. on p. 6).
- Glémin, Sylvain, Eric Bazin, and Deborah Charlesworth (2006). Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings of the Royal Society B: Biological Sciences* 273(1604), pp. 3011–3019. (Cit. on pp. 142, 150).
- Glover, Kevin A. et al. (2012). Three Decades of Farmed Escapees in the Wild: A Spatio-Temporal Analysis of Atlantic Salmon Population Genetic Structure throughout Norway. *PLoS ONE* 7(8), e43129. (Cit. on p. 28).
- Gompert, Zachariah and C. Alex Buerkle (2013). Analyses of genetic ancestry enable key insights for molecular ecology. *Molecular Ecology* 22(21), pp. 5278–5294. (Cit. on pp. 12, 23).
- González, Alejandra V. et al. (2020). Meta-analysis of the differential effects of habitat fragmentation and degradation on plant genetic diversity. *Conservation Biology* 34(3), pp. 711–720. (Cit. on p. 142).
- Gouskov, Alexandre and Christoph Vorburger (2016). River fragmentation and fish population structure: a comparison of three Swiss midland rivers. *Freshwater Science* 35(2), pp. 689–700. (Cit. on p. 142).
- Gresswell, Robert E. (2011). Biology, status, and management of the yellowstone cutthroat trout. *North American Journal of Fisheries Management* 31(5), pp. 782–812. (Cit. on pp. 6, 28).
- Griffith, Brad et al. (1989). Translocation as a Species Conservation Tool: Status and Strategy. *Science* 245(4917), p. 477. (Cit. on p. 4).

- Griffiths, David (2015). Connectivity and vagility determine spatial richness gradients and diversification of freshwater fish in North America and Europe. *Biological Journal of the Linnean Society* 116(4), pp. 773–786. (Cit. on p. 117).
- Gross, Iwo P., Alan E. Wilson, and Matthew E. Wolak (2024). The fitness consequences of wildlife conservation translocations: a meta-analysis. *Biological Reviews* 99(2), pp. 348–371. (Cit. on p. 29).
- Grundler, Maggie R. et al. (2019). Is genomic diversity a useful proxy for census population size? Evidence from a species-rich community of desert lizards. *Molecular Ecology* 28(7), pp. 1664–1674. (Cit. on p. 142).
- Gutenkunst, Ryan N. et al. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics* 5(10). (Cit. on p. 121).
- Hague, M. T. J. and E. J. Routman (2016). Does population size affect genetic diversity? A test with sympatric lizard species. *Heredity* 116(1), pp. 92–98. (Cit. on p. 142).
- Halbisen, Michael A. and Chris C. Wilson (2009). Variable Introgression from Supplemental Stocking in Southern Ontario Populations of Lake Trout. *Transactions of the American Fisheries Society* 138(4), pp. 699–719. (Cit. on p. 28).
- Hansen, Michael M. et al. (2009). Sixty years of anthropogenic pressure: a spatio-temporal genetic analysis of brown trout populations subject to stocking and population declines. *Molecular Ecology* 18(12), pp. 2549–2562. (Cit. on p. 28).
- Harada, Yasushi, Masashi Yokota, and Masaru Iizuka (1998). Genetic risk of domestication in artificial fish stocking and its possible reduction. *Population Ecology* 40(3), pp. 311–324. (Cit. on p. 5).
- Hayes, Mf and Np Banish (2017). Translocation and reintroduction of native fishes: a review of bull trout *Salvelinus confluentus* with applications for future reintroductions. *Endangered Species Research* 34, pp. 191–209. (Cit. on p. 5).
- Hess, Jon E. et al. (2011). Introgressive hybridization among major Columbia River Chinook salmon (*Oncorhynchus tshawytscha*) lineages within the Klickitat River due to hatchery practices. *Canadian Journal of Fisheries and Aquatic Sciences* 68(11), pp. 1876–1891. (Cit. on p. 29).
- Hohenlohe, Paul A., W. Chris Funk, and Om P. Rajora (2021). Population genomics for wildlife conservation and management. *Molecular Ecology* 30(1), pp. 62–82. (Cit. on p. 1).
- Huelsenbeck, John P., Rasmus Nielsen, and Jonathan P. Bollback (2003). Stochastic Mapping of Morphological Characters. *Systematic Biology* 52(2), pp. 131–158. (Cit. on p. 124).
- Huff, David D. et al. (2011). Mixed-source reintroductions lead to outbreeding depression in second-generation descendents of a native North American fish. *Molecular Ecology* 20(20), pp. 4246–4258. (Cit. on p. 6).
- Hughes, Jane M., Daniel J. Schmidt, and Debra S. Finn (2009). Genes in Streams: Using DNA to Understand the Movement of Freshwater Fauna and Their Riverine Habitat. *BioScience* 59(7), pp. 573–583. (Cit. on p. 143).
- Isaak, Daniel J. and Wayne A. Hubert (2004). Nonlinear Response of Trout Abundance to Summer Stream Temperatures across a Thermally Diverse Montane Landscape. *Transactions of the American Fisheries Society* 133(5), pp. 1254–1259. (Cit. on p. 150).

- Jahner, Joshua P et al. (2026). Resolving the evolutionary history of bighorn sheep to inform future management: an answer to the California bighorn lineage question. *Evolutionary applications* 19(3), e70205. (Cit. on p. 15).
- Jahner, Joshua P. et al. (2019). The genetic legacy of 50 years of desert bighorn sheep translocations. *Evolutionary Applications* 12(2), pp. 198–213. (Cit. on pp. 5, 7).
- Jeon, Jong Yoon et al. (2024). Genomic Diversity as a Key Conservation Criterion: Proof-of-Concept From Mammalian Whole-Genome Resequencing Data. *Evolutionary Applications* 17(9), e70000. (Cit. on p. 151).
- Johnson, James E., Ji X. He, and David G. Fielder (2015). Rehabilitation Stocking of Walleyes and Lake Trout: Restoration of Reproducing Stocks in Michigan Waters of Lake Huron. *North American Journal of Aquaculture* 77(3), pp. 396–408. (Cit. on p. 5).
- Kalyaanamoorthy, Subha et al. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14(6), pp. 587–589. (Cit. on pp. 17, 121).
- Kardos, Marty et al. (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences of the United States of America* 118(48). (Cit. on pp. 6, 31, 141).
- Kim, Daemin et al. (2023). Erosional exhumation of carbonate rock facilitates dispersal-mediated allopatric speciation in freshwater fishes. *Evolution* 77(11), pp. 2442–2455. (Cit. on p. 116).
- Kokkonen, Andrea L. et al. (2024). Using de novo transcriptomes to decipher the relationships in cutthroat trout subspecies (*Oncorhynchus clarkii*). *Evolutionary Applications* 17(7), e13735. (Cit. on pp. 126, 127).
- Korneliussen, Thorfinn Sand, Anders Albrechtsen, and Rasmus Nielsen (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15(1), pp. 1–13. (Cit. on pp. 19, 144).
- Kovach, Ryan P. et al. (2022). Genetic variation in westslope cutthroat trout reveals that widespread genetic rescue is warranted. *Canadian Journal of Fisheries & Aquatic Sciences* 79(6), pp. 936–946. (Cit. on pp. 147, 152).
- Largiadere, C. R. and A. Scholl (1996). Genetic introgression between native and introduced brown trout *Salmo trutta* L. populations in the Rhone River Basin. *Molecular Ecology* 5(3), pp. 417–426. (Cit. on p. 28).
- Lee, Tien Ming and Walter Jetz (2011). Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings of the Royal Society B: Biological Sciences* 278(1710), pp. 1329–1338. (Cit. on p. 141).
- Lehnert, Sarah J. et al. (2020). Multiple decades of stocking has resulted in limited hatchery introgression in wild brook trout (*Salvelinus fontinalis*) populations of Nova Scotia. *Evolutionary Applications* 13(5), pp. 1069–1089. (Cit. on p. 28).
- Létourneau, Justine et al. (2018). Predicting the genetic impact of stocking in Brook Charr (*Salvelinus fontinalis*) by combining RAD sequencing and modeling of explanatory variables. *Evolutionary Applications* 11(5), pp. 577–592. (Cit. on pp. 28, 29).
- Li, Heng (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 00(00), pp. 1–3. (Cit. on pp. 11, 120).
- Licciardi, Joseph M. and Kenneth L. Pierce (2018). History and dynamics of the Greater Yellowstone Glacial System during the last two glaciations. *Quaternary Science Reviews* 200, pp. 1–33. (Cit. on pp. 130, 137, 143, 150).

- Lindsay, J. B. (2016). Whitebox GAT: A case study in geomorphometric analysis. *Computers & Geosciences* 95, pp. 75–84. (Cit. on p. 145).
- Link, Paul K and Ernest R Keeley (2018). Geologic constraints on Snake-Columbia River drainage changes and the associated evolutionary diversification of cutthroat trout over the past 17 million years. *Cutthroat trout: Evolutionary biology and taxonomy*, pp. 53–75. (Cit. on p. 117).
- Lino, Ana et al. (2019). A meta-analysis of the effects of habitat loss and fragmentation on genetic diversity in mammals. *Mammalian Biology* 94, pp. 69–76. (Cit. on p. 142).
- Love Stowell, Sierra M. et al. (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(10), pp. 1565–1574. (Cit. on p. 7).
- Loxterman, Janet L and Ernest R Keeley (2012). Watershed boundaries and geographic isolation: patterns of diversification in cutthroat trout from western North America. *BMC Evolutionary Biology* 12(1), p. 38. (Cit. on pp. 7, 31, 117, 126, 127, 129).
- Luu, Keurcien, Eric Bazin, and Michael G B Blum (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* 17(1), pp. 67–77. (Cit. on pp. 21, 26).
- Mackintosh, Alexander et al. (2019). The determinants of genetic diversity in butterflies. *Nature Communications* 10(1), p. 3466. (Cit. on p. 142).
- Malinsky, Milan (2019). Dsuite - fast D-statistics and related admixture evidence from VCF files. *bioRxiv*, pp. 1–6. (Cit. on p. 123).
- Mandeville, Elizabeth G. et al. (2019). Variable hybridization outcomes in trout are predicted by historical fish stocking and environmental context. *Molecular Ecology* 28(16), pp. 3738–3755. (Cit. on p. 1).
- Marie, A. D., L. Bernatchez, and D. Garant (2010). Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology* 19(10), pp. 2025–2037. (Cit. on pp. 29, 32).
- Markle, DF (2018). An interim classification of the Cutthroat Trout complex, *Oncorhynchus clarkii* sensu lato, with comments on nomenclature. *Cutthroat Trout: Evolutionary biology and taxonomy*. Bethesda, MD: American Fisheries Society, Special Publication 36, pp. 181–197. (Cit. on p. 8).
- McFadden, Ian R. et al. (2023). Linking human impacts to community processes in terrestrial and freshwater ecosystems. *Ecology Letters* 26(2), pp. 203–218. (Cit. on p. 126).
- McLennan, E. A. et al. (2025). Reinforcements in the face of ongoing threats: a case study from a critically small carnivore population. *Animal Conservation* 28(1), pp. 49–65. (Cit. on p. 31).
- McMillan, John R. et al. (2023). A global synthesis of peer-reviewed research on the effects of hatchery salmonids on wild salmonids. *Fisheries Management and Ecology* 30(5), pp. 446–463. (Cit. on p. 5).
- Meek, Mariah H et al. (2023). Understanding Local Adaptation to Prepare Populations for Climate Change. *BioScience* 73(1), pp. 36–47. (Cit. on p. 33).
- Metcalf, J. L. et al. (2012). Historical stocking data and 19th century *span style="font-variant:small-caps;"¿DNA¿/span¿* reveal human-induced changes to native

- diversity and distribution of cutthroat trout. *Molecular Ecology* 21(21), pp. 5194–5207. (Cit. on pp. 6, 7).
- Metcalf, Jessica L., S. Love Stowell, et al. (2012). Historical stocking data and 19th century DNA reveal human-induced changes to native diversity and distribution of cutthroat trout. *Molecular Ecology* 21(21), pp. 5194–5207. (Cit. on p. 31).
- Metcalf, Jessica L., Victoria L. Pritchard, et al. (2007). Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. *Molecular Ecology* 16(21), pp. 4445–4454. (Cit. on p. 29).
- Metcalf, Jessica Lynne (2007). Conservation Genetics of the Greenback Cutthroat Trout (*Oncorhynchus clarkii stomias*). (PhD thesis). University of Colorado. (Cit. on p. 130).
- Minh, Bui Quang et al. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* 37(5), pp. 1530–1534. (Cit. on pp. 17, 121, 144).
- Mitchell, Finis (1985). *Wind River Trails: A hiking and fishing guide to the many trails and lakes of the Wind River range in Wyoming*. 6th ed., Wasatch Publishers, Inc. ISBN: 0-87480-626-7. (Cit. on p. 82).
- Moore, Daniel K et al. (2022). *Origin of Lava-Flow Benches in the Henrys Fork, Warm River, and Fall River Canyons, Idaho and Wyoming, and Origin of Upper and Lower Mesa Falls in Henrys Fork Canyon, Island Park, Idaho*. Technical Report 22-01. University of Idaho, Idaho Geological Survey, Moscow, Idaho. (Cit. on p. 130).
- Moran, Benjamin M. and Eric C. Anderson (2019). Bayesian inference from the conditional genetic stock identification model. *Canadian Journal of Fisheries and Aquatic Sciences* 76(4), pp. 551–560. (Cit. on p. 12).
- Morris, Shane D. et al. (2021). Factors affecting success of conservation translocations of terrestrial vertebrates: A global systematic review. *Global Ecology and Conservation* 28, e01630. (Cit. on p. 4).
- Muneepeerakul, Rachata et al. (2008). Patterns of vegetation biodiversity: The roles of dispersal directionality and river network structure. *Journal of Theoretical Biology* 252(2), pp. 221–229. (Cit. on p. 143).
- Muyle, Aline et al. (2021). Dioecy Is Associated with High Genetic Diversity and Adaptation Rates in the Plant Genus *Silene*. *Molecular Biology and Evolution* 38(3), pp. 805–818. (Cit. on p. 142).
- Nekola, Jeffrey C (1999). Paleoreugia and Neoreugia: The Influence of Colonization History on Community Pattern and Process. *Ecology* 80(8), pp. 2459–2473. (Cit. on p. 116).
- Nonaka, Etsuko et al. (2019). Scaling up the effects of inbreeding depression from individuals to metapopulations. *Journal of Animal Ecology* 88(8), pp. 1202–1214. (Cit. on p. 151).
- Nordberg, Brittany J. et al. (2021). Historical Data Provide Important Context for Understanding Declines in Cutthroat Trout. *North American Journal of Fisheries Management* (Gresswell 2011). (Cit. on p. 7).
- Novak, Ben J., Ryan Phelan, and Michele Weber (2021). U.S. conservation translocations: Over a century of intended consequences. *Conservation Science and Practice* 3(4), e394. (Cit. on p. 4).

- Novak, MARK A, JEFFREY L Kershner, and KAREN E Mock (2005). Molecular genetic investigation of Yellowstone cutthroat trout and finespotted Snake River cutthroat trout. (Cit. on pp. 8, 81, 118, 129, 131).
- Olden, Julian D. et al. (2004). Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution* 19(1), pp. 18–24. (Cit. on p. 4).
- Ommen, Gertjan B van (2005). A short history of genomics. *Part of Everyday Life*, p. 20. (Cit. on p. 1).
- Osborne, Megan J. et al. (2014). Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains fishes. *Molecular Ecology* 23(23), pp. 5663–5679. (Cit. on p. 142).
- Ouborg, N. Joop et al. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics* 26(4), pp. 177–187. (Cit. on p. 1).
- Ozerov, M. Y. et al. (2016). Genomewide introgressive hybridization patterns in wild Atlantic salmon influenced by inadvertent gene flow from hatchery releases. *Molecular Ecology* 25(6), pp. 1275–1293. (Cit. on p. 28).
- Pagel, Mark (1999). Inferring the historical patterns of biological evolution. *Nature* 401(6756), pp. 877–884. (Cit. on p. 148).
- Paradis, Emmanuel and Klaus Schliep (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35(3), pp. 526–528. (Cit. on pp. 18, 121, 146).
- Parchman, Thomas L. et al. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology* 21(12), pp. 2991–3005. (Cit. on pp. 11, 119).
- Paz-Vinas, I. and S. Blanchet (2015). Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *Journal of Evolutionary Biology* 28(4), pp. 986–994. (Cit. on pp. 143, 145).
- Paz-Vinas, I., G. Loot, et al. (2015). Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Molecular Ecology* 24(18), pp. 4586–4604. (Cit. on pp. 143, 150).
- Pederson, J. L. et al. (Jan. 2016). “Chapter 2 - The Bear River’s History and Diversion: Constraints, Unsolved Problems, and Implications for the Lake Bonneville Record”. In: *Developments in Earth Surface Processes*. Ed. by Charles G. Oviatt and John F. Shroder. Vol. 20. Lake Bonneville. Elsevier, pp. 28–59 (cit. on p. 129).
- Pelletier, Tara A. et al. (2018). Predicting plant conservation priorities on a global scale. *Proceedings of the National Academy of Sciences* 115(51), pp. 13027–13032. (Cit. on p. 141).
- Pennell, Matthew W. et al. (2014). geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30(15), pp. 2216–2218. (Cit. on p. 147).
- Pham, Lan, David West, and Gerard P. Closs (2013). Reintroduction of a native galaxiid (*Galaxias fasciatus*) following piscicide treatment in two streams: response and recovery of the fish population. *Ecology of Freshwater Fish* 22(3), pp. 361–373. (Cit. on p. 5).
- Pierce, Kenneth L (1982). History and dynamics of glaciation in the northern Yellowstone National Park area. (Cit. on p. 130).

- Pinheiro, José, Douglas Bates, R. Core Team, et al. (2025). nlme: linear and nonlinear mixed effects models. R package version 3.1-168. *The R Foundation for Statistical Computing (CRAN)*. (Cit. on p. 147).
- PRISM Group, Oregon State University (2023). *Average air temperature, 30 year*. URL: <https://prism.oregonstate.edu> (visited on 07/01/2025). (Cit. on pp. 20, 144).
- Purcell, Shaun et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* 81(3), pp. 559–575. (Cit. on p. 14).
- Rahel, Frank J. (2000). Homogenization of fish faunas across the United States. *Science* 288(5467), pp. 854–856. (Cit. on p. 5).
- Reich, David et al. (2009). Reconstructing Indian Population History. *Physiology & behavior* 461(7263), pp. 489–494. (Cit. on p. 18).
- Revell, Liam J. (2024). phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* 12, e16505. (Cit. on pp. 124, 147).
- Riaz, Maria et al. (2020). Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. *Ecological Applications* 30(2), e02034. (Cit. on p. 5).
- Romiguier, J. et al. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515(7526), pp. 261–263. (Cit. on pp. 142, 150).
- Rosenberg, Noah A. et al. (2002). Genetic Structure of Human Populations. *Science* 298(5602), pp. 2381–2385. (Cit. on p. 15).
- Rougemont, Quentin et al. (2023). Allele surfing causes maladaptation in a Pacific salmon of conservation concern. *PLOS Genetics* 19(9), e1010918. (Cit. on p. 122).
- Santini, Luca et al. (2016). A trait-based approach for predicting species responses to environmental change from sparse data: how well might terrestrial mammals track climate change? *Global Change Biology* 22(7), pp. 2415–2424. (Cit. on p. 116).
- SANTOS, Nuno P et al. (2006). Genetic evidence for limited introgression between wild and stocked individuals in Portuguese brown trout, *Salmo trutta* populations. *Folia Zool* 55(4), pp. 433–443. (Cit. on p. 28).
- Satake, Akiko and Hitoshi Araki (2012). Stocking of captive-bred fish can cause long-term population decline and gene pool replacement: predictions from a population dynamics model incorporating density-dependent mortality. *Theoretical Ecology* 5, pp. 283–296. (Cit. on p. 5).
- Schmetterling, David A. (2001). Seasonal Movements of Fluvial Westslope Cutthroat Trout in the Blackfoot River Drainage, Montana. *North American Journal of Fisheries Management* 21(3), pp. 507–520. (Cit. on p. 117).
- Schwartz, Mark W. and Tara G. Martin (2013). Translocation of imperiled species under changing climates. *Annals of the New York Academy of Sciences* 1286(1), pp. 15–28. (Cit. on p. 4).
- Shastri, Vivaswat et al. (2021). Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources*. (Cit. on pp. 12, 23).
- Shiozawa, Dennis K, R Paul Evans, Derek D Houston, et al. (2018). Geographic variation, isolation, and evolution of cutthroat trout with comments on future directions for management and research. *Cutthroat trout: evolutionary biology and taxonomy. Special*

- Publication of the American Fisheries Society* (36), pp. 129–172. (Cit. on pp. [2](#), [7](#), [31](#), [32](#), [117](#), [127](#), [130](#)).
- Shiozawa, Dennis K, R Paul Evans, Peter Unmack, et al. (2010). CUTTHROAT TROUT PHYLOGENETIC RELATIONSHIPS WITH AN ASSESSMENT OF ASSOCIATIONS AMONG SEVERAL SUBSPECIES. *WEST YELLOWSTONE, MT SEPTEMBER 28-39, 2010*, p. 158. (Cit. on p. [127](#)).
- Skotte, Line, Thorfinn Sand Korneliussen, and Anders Albrechtsen (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics* 195(3), pp. 693–702. (Cit. on p. [15](#)).
- Smith, GR et al. (2002). Biogeography and timing of evolutionary events among Great Basin fishes. *Great Basin aquatic systems history* 33, pp. 175–234. (Cit. on p. [127](#)).
- Smith, Robert B and Lee J Siegel (2000). *Windows into the Earth: The Geologic Story of Yellowstone and Grand Teton National Parks*, Oxford University Press. (Cit. on p. [117](#)).
- Stearley, Ralph F and Gerald R Smith (2016). SALMONID FISHES FROM MIO-PLIOCENE LAKE SEDIMENTS IN THE WESTERN SNAKE RIVER PLAIN AND THE GREAT BASIN. *Miscellaneous Publications of the Museum of Zoology, University of Michigan* (204), pp. 1–43. (Cit. on p. [121](#)).
- Stoddart, D.R. (1981). Biogeography: dispersal and drift. *Progress in Physical Geography: Earth and Environment* 5(4), pp. 575–590. (Cit. on p. [116](#)).
- Stokes, Maya F. et al. (2023). Erosion of heterogeneous rock drives diversification of Appalachian fishes. *Science* 380(6647), pp. 855–859. (Cit. on p. [116](#)).
- Supple, Megan A. and Beth Shapiro (2018). Conservation of biodiversity in the genomics era. *Genome Biology* 19(1), p. 131. (Cit. on p. [1](#)).
- Terui, Akira, Nobuo Ishiyama, et al. (2018). Metapopulation stability in branching river networks. *Proceedings of the National Academy of Sciences* 115(26), E5963–E5969. (Cit. on pp. [30](#), [143](#)).
- Terui, Akira, Seoghyun Kim, et al. (2021). Emergent dual scaling of riverine biodiversity. *Proceedings of the National Academy of Sciences* 118(47), e2105574118. (Cit. on pp. [143](#), [145](#)).
- Terui, Akira, Hirokazu Urabe, et al. (2023). Intentional release of native species undermines ecological stability. *Proceedings of the National Academy of Sciences* 120(7), e2218044120. (Cit. on p. [30](#)).
- Thomas, Chris D (2011). Translocation of species, climate change, and the end of trying to recreate past ecological communities. *Trends in Ecology & Evolution* 26(5), pp. 216–221. (Cit. on p. [4](#)).
- Thomaz, Andréa T. et al. (2015). Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. *Journal of Biogeography* 42(12), pp. 2389–2401. (Cit. on p. [116](#)).
- Travis, Justin M. J. et al. (2013). Dispersal and species' responses to climate change. *Oikos* 122(11), pp. 1532–1540. (Cit. on p. [116](#)).
- Trotter, Pat (2008). *Cutthroat: native trout of the west*, Univ of California Press. (Cit. on p. [7](#)).
- Trotter, Pat et al. (2018). *Cutthroat Trout: evolutionary biology and taxonomy*, American Fisheries Society. (Cit. on p. [6](#)).

- Turner, Thomas F et al. (2006). Life history and environmental variation interact to determine effective population to census size ratio. *Proceedings of the Royal Society B: Biological Sciences* 273(1605), pp. 3065–3073. (Cit. on p. 142).
- U.S. Geological Survey (2018). *High Resolution National Hydrography Dataset Plus*. URL: <https://www.usgs.gov/national-hydrography/access-national-hydrography-products> (visited on 02/29/2024). (Cit. on pp. 14, 144).
- Udall, Randy (1986). Finis Mitchell: Lord of the Winds. *Audubon Magazine*. (Cit. on p. 82).
- Valiquette, Eliane et al. (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(6), pp. 625–644. (Cit. on pp. 28, 29, 32).
- Van Doornik, Donald M. et al. (2013). Genetic Monitoring of Threatened Chinook Salmon Populations: Estimating Introgression of Nonnative Hatchery Stocks and Temporal Genetic Changes. *North American Journal of Fisheries Management* 33(4), pp. 693–706. (Cit. on p. 29).
- Van Orden, Tanner S. et al. (2026). Phylogeny of Cutthroat Trout (*Oncorhynchus clarkii*) Based on Full Mitochondrial Genomes With Comments on Species Delimitation and Taxonomy. *Evolutionary Applications* 19(7), e70304. (Cit. on pp. 126, 127).
- Vellend, Mark (2005). Species Diversity and Genetic Diversity: Parallel Processes and Correlated Patterns. *The American Naturalist* 166(2), pp. 199–215. (Cit. on p. 150).
- Vellend, Mark and Monica A. Geber (2005). Connections between species diversity and genetic diversity. *Ecology Letters* 8(7), pp. 767–781. (Cit. on p. 150).
- Wang, Shaopeng, Bart Haegeman, and Michel Loreau (2015). Dispersal and metapopulation stability. *PeerJ* 3, e1295. (Cit. on p. 30).
- Waters, Jonathan M., Christopher P. Burrige, and Dave Craw (2020). River Capture and Freshwater Biological Evolution: A Review of Galaxiid Fish Vicariance. *Diversity* 12(6), p. 216. (Cit. on p. 129).
- Weeks, Andrew R. et al. (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications* 4(6), pp. 709–725. (Cit. on pp. 4, 32).
- White, Shannon L. et al. (2018). Limited hatchery introgression into wild brook trout (*Salvelinus fontinalis*) populations despite reoccurring stocking. *Evolutionary Applications* 11(9), pp. 1567–1581. (Cit. on p. 28).
- Whiteley, Andrew R. et al. (2015). Genetic rescue to the rescue. *Trends in Ecology and Evolution* 30(1), pp. 42–49. (Cit. on pp. 4, 6).
- Whittaker, R. H. (1972). EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. *TAXON* 21(2-3), pp. 213–251. (Cit. on p. 31).
- Wright, Sewall (1931). Evolution in Mendelian populations. *Genetics* 16(2), p. 97. (Cit. on p. 151).
- Wu, Qiusheng and Andrew Brown (2025). 'whitebox': 'WhiteboxTools' R Frontend. (Cit. on p. 145).
- Wyoming Game & Fish Department (2024). *Fish Stocking Report*. URL: <https://wgfapps.wyo.gov/FishStock/FishStock> (visited on 03/11/2025). (Cit. on p. 7).
- Young, Michael K et al. (2018). The phylogeography of westslope cutthroat trout. In: Trotter, P.; Bisson, P.; Schultz, L.; Roper, B., eds. *Cutthroat Trout: Evolutionary*

- Biology and Taxonomy. Special Publication 36. Bethesda, MD: American Fisheries Society. p. 261-301.*, pp. 261–301. (Cit. on p. [152](#)).
- Zhai, Dongdong et al. (2019). Genetic diversity and population structure of a cyprinid fish (*Ancherythroculter nigrocauda*) in a highly fragmented river. *Journal of Applied Ichthyology* 35(3), pp. 701–708. (Cit. on p. [142](#)).
- Zhang, Chao, Rasmus Nielsen, and Siavash Mirarab (2025). CASTER: Direct species tree inference from whole-genome alignments. *Science* 387(6737), eadk9688. (Cit. on pp. [17](#), [121](#), [144](#)).
- Zimmerman, Shawna J et al. (2019). Evaluation of genetic change from translocation among Gunnison Sage-Grouse (*Centrocercus minimus*) populations. *The Condor* 121(1), duy006. (Cit. on p. [5](#)).