NOAA Processed Report NMFS-NWFSC-PR-2020-06

https://doi.org/10.25923/mv3n-zb79

Reviewing and Synthesizing the State of the Science Regarding Associations between Adult Run Timing and Specific Genotypes in Chinook Salmon and Steelhead:

Report of a workshop held in Seattle, Washington, 27–28 February 2020

June 2020

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northwest Fisheries Science Center

NOAA Processed Report Series NMFS-NWFSC-PR

The Northwest Fisheries Science Center of NOAA's National Marine Fisheries Service uses the NOAA Processed Report NMFS-NWFSC-PR series to disseminate information only. Manuscripts have not been peer-reviewed and may be unedited. Documents within this series represent sound professional work, but do not constitute formal publications. They should only be footnoted as a source of information, and may not be cited as formal scientific literature. The data and any conclusions herein are provisional, and may be formally published elsewhere after appropriate review, augmentation, and editing.

NWFSC Processed Reports are available from the NOAA Institutional Repository, https://repository.library.noaa.gov.

Mention throughout this document of trade names or commercial companies is for identification purposes only and does not imply endorsement by the National Marine Fisheries Service, NOAA.

Recommended citation:

(Ford et al. 2020)[1]

[1] Ford, M., K. Nichols, R. Waples, E. C. Anderson, M. Kardos, I. Koch, G. McKinney, M. R. Miller, J. Myers, K. Naish, S. Narum, K. G. O'Malley, D. Pearse, T. Seamons, A. Spidle, P. Swanson, T. Q. Thompson, K. Warheit, and S. Willis. 2020. Reviewing and Synthesizing the State of the Science Regarding Associations between Adult Run Timing and Specific Genotypes in Chinook Salmon and Steelhead: Report of a workshop held in Seattle, Washington, 27–28 February 2020. U.S. Department of Commerce, NOAA Processed Report NMFS-NWFSC-PR-2020-06.

https://doi.org/10.25923/mv3n-zb79

NOAA FISHERIES

Reviewing and Synthesizing the State of the Science Regarding Associations between Adult Run Timing and Specific Genotypes in Chinook Salmon and Steelhead: Report of a workshop held in Seattle, Washington, 27–28 February 2020

Michael Ford,[1] Krista Nichols,[1] Robin Waples,[1] Eric C. Anderson,[2] Marty Kardos,[3,1] Ilana Koch,[4] Garrett McKinney,[1] Michael R. Miller,[5] Jim Myers,[1] Kerry Naish,[6] Shawn Narum,[4] Kathleen G. O'Malley,[7] Devon Pearse,[2] Todd Seamons,[8] Adrian Spidle,[9] Penny Swanson,[1] Tasha Q. Thompson,[5] Ken Warheit,[8] and Stuart Willis[4]

https://doi.org/10.25923/mv3n-zb79

June 2020

[1]Conservation Biology Division
Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle, Washington 98112

[2]Fisheries Ecology Division
Southwest Fisheries Science Center
110 McAllister Way
Santa Cruz, California 95060

[3]Flathead Lake Biological Station
Division of Biological Sciences
University of Montana
32125 Bio Station Lane
Polson, Montana 59860

[4]Columbia River Inter-Tribal Fish Commission
700 Northeast Multnomah Street, Suite 12
Portland, Oregon 97232

[5]Department of Animal Science
University of California, Davis
2251 Meyer Hall
Davis, California 95616

[6]School of Aquatic and Fisheries Sciences
University of Washington
1122 Northeast Boat Street
Seattle, Washington 98195

[7]State Fisheries Genomics Lab
Coastal Oregon Marine Experiment Station
Department of Fisheries and Wildlife
Oregon State University
2030 Southeast Marine Science Drive
Newport, Oregon 97365

[8]Washington Department of Fish and Wildlife
Natural Resources Building
1111 Washington Street Southeast
Olympia, Washington 98501

[9]Northwest Indian Fisheries Commission
6730 Martin Way East
Olympia, Washington 98516

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northwest Fisheries Science Center

Contents

Figures

Tables

Introduction

Over the past three decades, methods and technology developed as part of concerted efforts to sequence and study the human genome have trickled down for eventual use with non-model species like Pacific salmon (*Oncorhynchus* spp.) and steelhead (the anadromous form of rainbow trout, *O. mykiss*). This trickle-down effect greatly sped up after 2007, following major technological breakthroughs that led to orders-of-magnitude increases in the rate of DNA sequencing and concomitant reductions in cost (http://www.genome.gov/sequencingcosts/). For salmon, it had long been hoped (and expected by many) that genomic tools would provide novel insights into the genetic basis of adaptive traits, which were difficult to study by earlier methods. As discussed in the next section (Background), the conventional view in quantitative genetics has been that most traits are controlled by many genes, each with a small effect (Falconer & Mackay 1996; Mackay et al. 2009). This paradigm was important in guiding identification of conservation units in salmon and steelhead, which in turn helped to determine which units were protected under the U.S. Endangered Species Act (ESA). Like grizzly bears, bald eagles, and alligators, groups of salmon and steelhead are listed and protected under the provision in the ESA that allows Distinct Population Segments (DPSs) of vertebrate species to be listed, even if conspecific populations are not at risk elsewhere.

By the mid-2010s, new genomics data were suggesting that individual genes or genome regions have large effects on some key life history traits in both Atlantic salmon (e.g., Ayllon et al. 2015; Barson et al. 2015) and steelhead (e.g., Miller et al. 2012; Pearse et al. 2014). More recently, two studies documented a strong association between adult migration timing and a small region of the genome (near the GREB1L gene) in both coastal Chinook salmon (*O. tshawytscha;* Prince et al. 2017) and the coastal subspecies of steelhead (*O. mykiss irideus*; Hess et al. 2016; Prince et al. 2017). These latter results have led to a number of inquiries into the status and conservation of diverse run timing in salmonids.

Adult salmon run timing (generally defined as season of entry into fresh water to initiate the last phase of the spawning migration) is a trait of particular importance for many reasons. In both Chinook salmon and steelhead, adults can be found entering fresh water in every month of the year (Figure 1), but the seasonal ranges within most drainages are shorter in duration and often bimodal. These diverse freshwater entry times allow salmon and steelhead to take full advantage of the range of habitats suitable for spawning and rearing. Early-returning

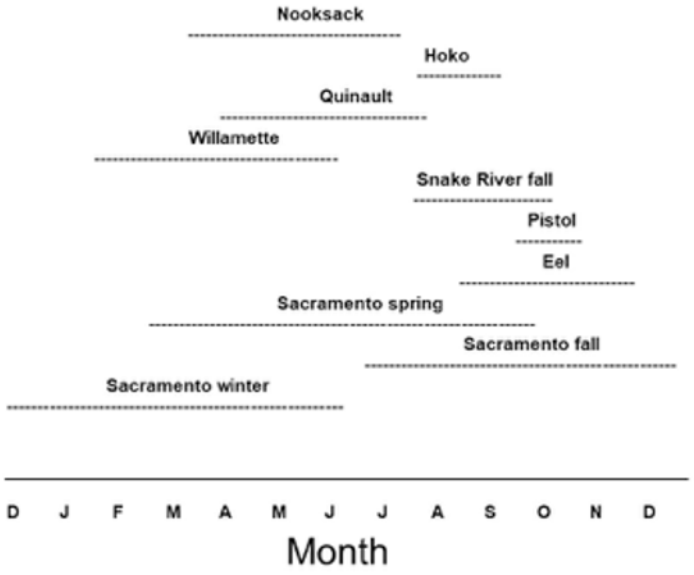

Figure 1. Duration of adult run timing for representative populations of Chinook salmon (reproduced from Waples 2006). The estimated durations for individual populations are approximate and can be affected by methodology, including how far from tidewater run timing is assessed.

spring Chinook salmon and summer steelhead are prized by anglers for their high fat content, and spring Chinook are of special cultural significance to Native American tribes, who value them as the first salmon to return each year. For a century or more, managers have used adult run timing to define salmon and steelhead stocks, regulate harvest, and operate large hatchery programs.

No consensus has been reached in the scientific literature regarding the most appropriate terminology to describe migration timing in Chinook salmon and steelhead (see Healey 1983 and Quinn et al. 2016 for prominent examples)—a situation which has led to confusion for many. We make no attempt to resolve this issue here; instead, we use the terms "early" or "early-migrating" to refer to spring Chinook salmon and summer steelhead and "late" or "late-migrating" to refer to fall Chinook and winter steelhead. These terms are imperfect, as the seasonal run types are themselves variable, but they have the virtues of simplicity and being readily understood. In general, these run-timing designations apply to different populations or stocks, which for the purposes of this document are meant to represent demographically independent units sensu McElhany et al. (2000) (i.e., units for which population dynamics are determined more by local births and deaths than by immigration). As discussed below, however, in some cases early and late run types as defined here might represent phenotypic variation within a single demographic unit, or different extremes along a geographic/temporal cline.

Based on genetic information available at the time most of the ESA status reviews for salmon and steelhead were conducted (1990s), as well as the conventional quantitative-genetic paradigm that assumes traits are controlled by many loci of small effect, populations from the same coastal river basin, but with different adult run timing, were generally considered to be components of diversity within the same salmon and steelhead DPSs. In other words, patterns of genetic variation throughout most of the genome indicate close genetic relationships between fish with different run times returning to the same watershed (e.g., Waples et al. 2004; Moran et al. 2013; Hecht et al. 2015). Results presented by Hess et al. (2016) and Prince et al. (2017) suggested that this paradigm might need to be revisited, because multiple populations of summer steelhead and spring Chinook shared the same alleles of large effect associated with early run timing suggesting that at this specific genomic region (on chromosome 28 in both Chinook salmon and steelhead) patterns of variation reflected run timing rather than geography. Soon after those papers were published, the U.S. National Marine Fisheries Service (NMFS), the agency responsible for implementing the ESA for most marine and anadromous species, received a petition by the Karuk Tribe and Salmon River Restoration Council of northern California to recognize Klamath River spring Chinook salmon as a separate DPS from the more abundant fall Chinook, and to list the new DPS as a threatened or endangered species under the ESA (NMFS 2018). Subsequently, NMFS has received additional ESA petitions for separate listings of Northern California summer-run steelhead (determined to be "Not warranted" (85 FR 6527) and Oregon Coast spring-run Chinook salmon (Native Fish Society 2019).

Although it was clear from the Prince et al. (2017) paper that significant genomic-life history associations occur in both salmon and steelhead, the implications of this result for conservation are not as clear (e.g., Langin 2018). A perspective by Waples and Lindley (2018) identified a number of key questions and the research needs to address them, the

most important of which were to 1) better characterize the distribution of adult run-timing alleles, both geographically and temporally; and 2) more fully explore the run-timing phenotype of individuals that are heterozygotes for "early" and "late" migration alleles. The general results reported in Hess et al. (2016) and Prince et al. (2017) had been known within the salmon genetics community for some time, and by the time the Prince et al. (2017) paper was published, researchers in several laboratories were already collecting new data to further address these key questions (e.g., Narum et al. 2018, Micheletti et al. 2018, Thompson et al. 2019). It soon became apparent that it would be useful for these researchers to meet to share and discuss their published and unpublished results and refine strategies for further research. A workshop was subsequently held on 27–28 February 2020, in Seattle, and what transpired there are the subjects of this report.

The goal of the workshop was to characterize the current state of the science regarding the nature of the associations between genetic variation and run timing in Chinook salmon and steelhead. The study of the genetics of life-history variation as a whole, as well as run timing specifically, continues to advance at a rapid pace, so we did not expect that it would be possible to obtain definitive scientific conclusions. Nevertheless, the published studies, as well as unpublished data, are already starting to influence conservation decisions, so it is important to understand the current results and develop strategies to help guide future research. The workshop was organized as follows:

1. Researchers presented their (un)published data;
2. Participants identified scientific conclusions for which there was general agreement;
3. Participants identified areas of uncertainty or scientific disagreement;
4. Participants discussed types of future research that would most effectively address areas of uncertainty/disagreement;
5. Participants discussed pros and cons of some alternative approaches to conserving run-timing diversity in salmon and steelhead.

Workshop participants (see Table 1) were invited who had relevant genomics data for Chinook salmon or steelhead. In addition, this group included scientists having more general expertise in salmon genetics, more general genomics experience,

Table 1. Workshop participants (*P*) and observers (*O*).

Name	Affiliation
Eric Anderson (P)	NMFS Santa Cruz
Craig Busack (O)	NMFS Portland
Michael Ford (P)	NMFS Seattle
Marty Kardos (P)	University of Montana (now NMFS Seattle)
Ilana Koch (P)	Columbia River Inter-Tribal Fish Commission
Rob Markle (O)	NMFS Portland
Garrett McKinney (P)	NMFS Seattle
Mike Miller (P)	University of California, Davis
Jim Myers (P)	NMFS Seattle
Kerry Naish (P)	University of Washington
Shawn Narum (P)	Columbia River Inter-Tribal Fish Commission
Krista Nichols (P)	NMFS Seattle
Kathleen O'Malley (P)	Oregon State University
Devon Pearse (P)	NMFS Santa Cruz
Gary Rule (O)	NMFS Portland
Todd Seamons (P)	Washington Department of Fish and Wildlife
Adrian Spidle (P)	Northwest Indian Fisheries Commission
Penny Swanson (P)	NMFS Seattle
Tasha Thompson (P)	University of California, Davis
Robin Waples (P)	NMFS Seattle
Ken Warheit (P)	Washington Department of Fish and Wildlife
Stuart Willis (P)	Columbia River Inter-Tribal Fish Commission
Chris Yates (O)	NMFS Long Beach

or experience with conservation and management of salmon and steelhead. Finally, several staff members from the NMFS West Coast Regional Office were observers at the workshop.

It is important to clarify that, although we found it useful to identify areas of scientific agreement where they exist, in general we made no attempt to find consensus among the participants; instead, this report attempts to reflect the full range of views of participants on the issues that were discussed. Furthermore, the group did not attempt to make any specific recommendations to managers regarding what should be done with the information presented; instead, we discussed the likely consequences if different management/conservation strategies were implemented. Nevertheless, we expect that this report will be useful to NMFS and to the broader scientific community in deciding how best to incorporate new genomics information into conservation and management.

Background

Reported declines in Pacific salmon and steelhead populations date back over a century, as do efforts to mitigate the declines (Lichatowich 1999). In the late 20th century, concerns over declining salmon runs in the Columbia River Basin were forestalled temporarily by passage of the 1980 Northwest Power Act; this act called for development of the Columbia River Basin Fish and Wildlife Program, which Lee and Lawrence (Lee & Lawrence 1986, p. 433) argued was "the most ambitious and costly effort at biological restoration on the planet." The Fish and Wildlife Program focused on restoration of wild salmon through flow augmentation and modifications to dams to improve fish survival, and it was intended to double salmon runs within a decade. By 1990 it was clear the prediction that wild salmon runs would be doubled was wildly optimistic, and early drafts of a report (subsequently published as Nehlsen et al. (1991)) documenting over 200 at-risk wild salmon stocks were circulating within the region. In 1990, a Native American tribe and several conservation groups filed formal petitions with NMFS to list five groups of Columbia and Snake River salmon populations as DPSs under the ESA (NMFS 1990). Meanwhile, the remnant winter-run population of Chinook salmon in the Sacramento River was the subject of a series of ESA evaluations after 1985, culminating in the listing of this population as an endangered DPS in 1994 (NMFS 1995).

The 1990 salmon ESA petitions, with the prospect of potentially many more to come, raised an important but complex question: What constitutes a DPS of salmon? The term "distinct population segment" has no clear biological meaning, and the ESA was silent on how to identify DPSs. The U.S. Fish and Wildlife Service, which has ESA jurisdiction for terrestrial species, had made DPS determinations for numerous species individually and on an ad hoc basis, which provided little general guidance for future applications of the DPS provision. With legal deadlines for the 1990 salmon petitions already approaching, NMFS commissioned a science paper (Waples 1991) and subsequently developed a policy (NMFS 1991) that specified two criteria (substantial reproductive isolation, and substantial contribution to the evolutionary legacy of the species as a whole), both of which must be met for a salmon population or group of salmon populations to be considered a DPS. In this context, evolutionary legacy is both the product of past evolutionary events and the reservoir of genetic variation that allows subspecific units to respond to future environmental

challenges. Substantial reproductive isolation was considered necessary for evolutionarily important differences to accrue. This framework equated DPSs with the nascent concept of Evolutionarily Significant Units (ESUs), so both terms are used in describing the same conservation units of salmon and steelhead. During the 1990s, NMFS used the salmon ESU policy to systematically identify salmon and steelhead DPSs in the western contiguous states and determine their status. In 1996, USFWS and NMFS published a DPS policy applicable to all vertebrates that uses a similar two-part test (USFWS & NMFS 1996).

Most salmon and steelhead DPSs cover relatively large geographic freshwater spawning and rearing areas (e.g., Oregon Coast; Puget Sound; Snake River; see Figure 2) and include up to several dozen demographically-independent populations or stocks, which nevertheless share common ecological, phenotypic, and genetic characteristics. Adult run timing was one of the most important life-history traits considered in making DPS/ESU determinations. Populations that migrate from salt water to fresh water many months before spawning (early migration) are found in many anadromous fishes, but the phenomenon is especially common in salmonids (Quinn et al. 2016). The early-migration strategy in returning adults entails several costs to populations that adopt it, including reduced opportunities for growth in the productive marine environment and potentially higher exposure to predation and extreme temperature and flow regimes in the freshwater habitats they must hold in prior to spawning. Because the early-migration strategy is nevertheless widespread, it must generate benefits to offset these costs. Two general types of potential benefits to early migration have been identified (Quinn et al. 2016): 1) If migrating early is the only way to gain access to specific spawning and rearing habitats (e.g., because of temperature/flow conditions that are not available or are less available to later-migrating fish), the strategy could also be successful in spite

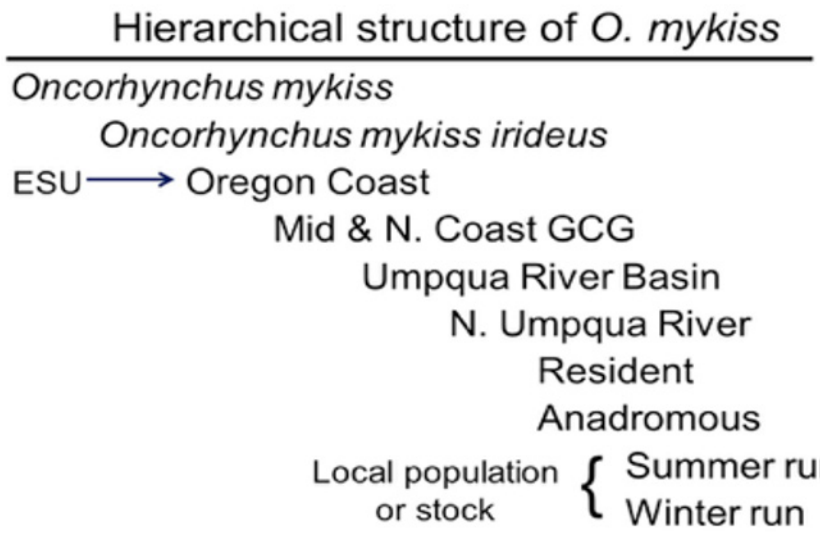

Figure 2. The hierarchical population structure of *O. mykiss* pictured here is typical for most Pacific salmon species. In North America, two subspecies of *O. mykiss* are recognized, with *O. mykiss irideus* being the coastal form. The Oregon Coast, bounded in the south by Cape Blanco and in the north by the Columbia River, is a natural biogeographic unit, and the State of Oregon has identified three "Gene Conservation Groups" (GCGs) within the Oregon Coast. Within the Mid & N Coast GCG the largest river is the Umpqua, which has two major forks. The North Umpqua River has both resident (rainbow trout) and anadromous (steelhead) forms, with the latter divided into summer run and winter run. It is generally agreed that these two adult run-timing forms are different populations/stocks, and they are the focus of day-to-day fishery management by state and tribal comanagers. For purposes of federal protection under the ESA, NMFS has generally recognized larger units that include multiple populations, and in this case all populations within the Oregon Coast as defined above are included in a single DPS/ESU. Modified from Waples (2006).

of the costs; 2) If mortality while holding in fresh water is lower than mortality that would be experienced at sea, and the difference is strong enough to offset foregone opportunities for marine growth, early migration could increase fitness. In both Chinook salmon and steelhead, early-migrating populations have been extirpated at higher rates than late-migrating populations (Gustafson et al. 2007), which suggests that recent anthropogenic changes have skewed the cost-benefit balance against the early migration strategy.

At the time the coastwide status reviews were conducted (1990s), most of the available genetic data were from allozymes (protein polymorphisms genotyped by electrophoresis), and in coastal drainages and in the Lower Columbia River, the following pattern was consistently found: different life-history types of Chinook salmon and steelhead within the same stream were genetically more similar to each other than either was to the same life-history type in another stream (Figure 3). In the largest interior basin (Columbia/Snake), the opposite pattern was found for Chinook salmon, with genetic affinities determined more by run timing than by geography (all interior steelhead are part of the subspecies *O. mykiss gairdneri* and are considered to have a single seasonal (summer) time of freshwater entry).

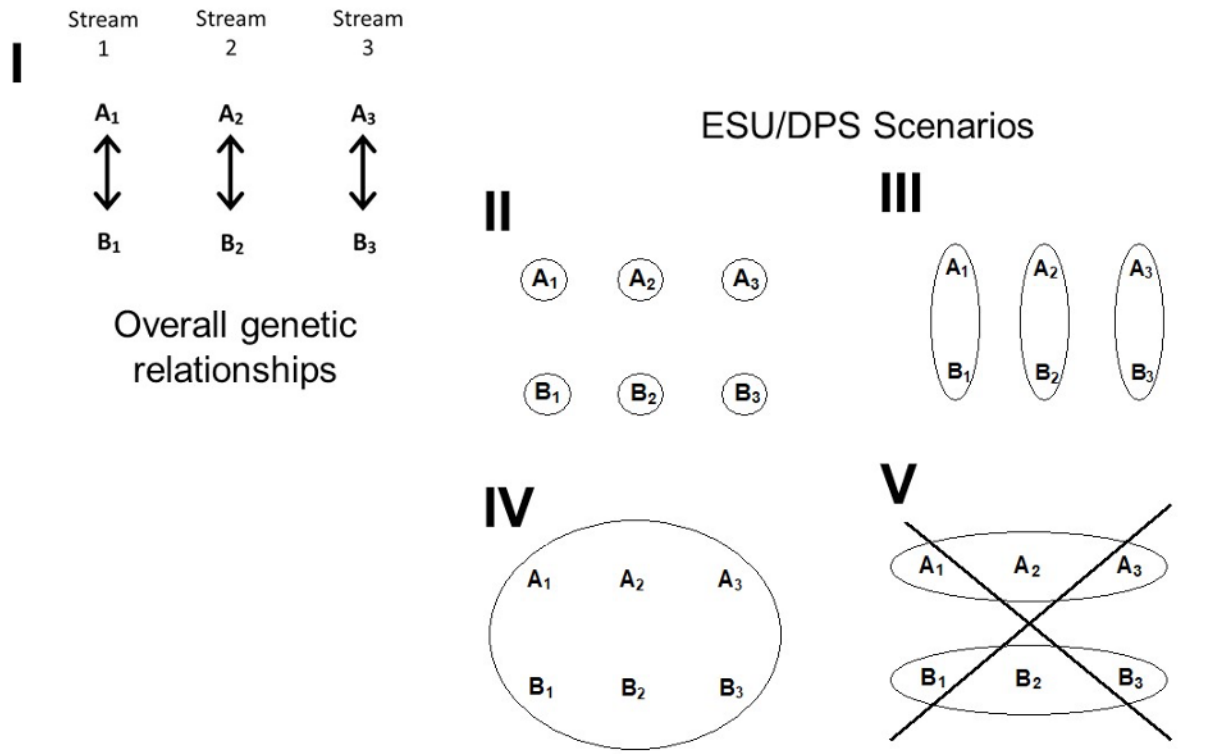

Figure 3. Upper left, I: overall genetic relationships between alternative life-history types (A, B). Data from allozymes, microsatellites, and the vast majority of SNPs show that different life-history types within the same stream are genetically more similar than either life-history type is to the same type in a different stream. This pattern is found for spring vs fall Chinook salmon and summer vs winter steelhead in coastal drainages and the Lower Columbia River, and it also holds for resident and anadromous *O. mykiss*, except when anadromous access to resident spawning areas is blocked. For details, see Busby et al. (1996); Myers et al. (1998); Waples et al. (2004). Bottom right: given these overall genetic relationships, using a lineage-based concept could produce conservation units as depicted in II, III, or IV, depending on whether one is inclined to be a lumper or a splitter, but scenario V would create artificial units that don't correspond to overall genetic lineages. Modified from Waples (2006).

These same general geographic/ life-history patterns were also found in subsequent analyses of genetic datasets based on microsatellites (e.g., Moran et al. 2013) and single-nucleotide polymorphisms (SNPs; e.g., Narum et al. 2008; Hecht et al. 2015; Arciniega et al. 2016). Historical population genetic structure in the extensive Sacramento-San Joaquin system is largely unknown, as major anthropogenic changes started in the 19th century. Long before any genetic samples were taken, most spring Chinook salmon populations had been extirpated, and fall-run populations had been homogenized (Williamson and May 2005).

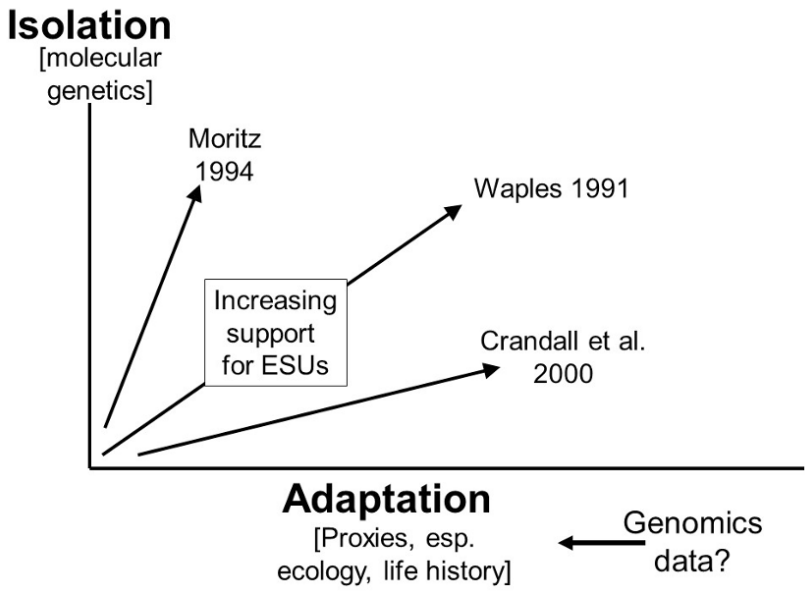

Figure 4. Two major axes of diversity that characterize the most widely-used ESU concepts. Moritz's (1994) method based on reciprocal monophyly of mtDNA places almost all emphasis on the isolation axis. Crandall et al.'s (2000) method based on exchangeability places more emphasis on adaptations. The NMFS salmon ESU framework places relatively equal weight on the two axes. Isolation is most easily documented using molecular genetic methods, whereas inferences about adaptations have traditionally relied on proxies like ecology and life history. The advent of new genomics tools raises the possibility that it will be possible to study adaptations directly at the DNA level. Modified from Waples and Lindley (2018).

Based on these results, it was concluded that adult run-timing differences had evolved independently many times within both Chinook salmon and steelhead, most likely with the more specialized early migration type evolving from local populations of the more generalist late migration type (Thorgaard 1983; Busby et al. 1996; Myers et al. 1998; Waples et al. 2004). Furthermore, it was thought that the early-migration phenotype could evolve from mature migrators on relatively short timescales (perhaps around 100 years; Waples et al. 2004). Therefore, in defining ESUs of coastal and lower Columbia River populations in both species, it was concluded that adult run-timing differences reflected diversity within ESUs (as illustrated in Figure 2). This was consistent with the standard quantitative genetic paradigm, which holds that most traits are controlled by many genes of small effect (Falconer & Mackay 1996; Mackay et al. 2009). Empirical studies from many species support this paradigm; for example, recent DNA studies have confirmed that height in humans is associated with many thousand SNPs spread throughout the genome (Wood et al. 2014).

The two criteria used to identify salmon ESUs/DPSs define two axes of diversity that can be used to map other widely-used methods to define conservation units (Figure 4; Waples 2006). Whereas Moritz's (1994) ESU framework depends almost entirely on the isolation axis and that of Crandall et al. (2000) gives greater attention to adaptations, the NMFS ESU framework places relatively equal weight on the two axes. In defining ESUs of salmon, molecular genetic data were particularly important for identifying reproductively isolated groups.

For the most part, however, inferences about adaptations had to rely on proxies, especially life history and other traits (which are influenced by both genes and the environment) and ecological features of the habitat (which can reflect different selective regimes that promote local adaptation). Now that genomics data are relatively cheaply and easily obtainable, it is becoming possible to evaluate the extent to which genomics data can take the place of these proxies and allow us to study adaptations at the level of DNA (Waples et al. 2020).

Presentation Summaries

A key component of the workshop was presentation of both previously published and currently unpublished results related to the genetic basis of adult return time, from laboratories across the Pacific Northwest. This section summarizes these presentations and their principal conclusions. The first two presentations summarized published information; the remaining presentations discussed unpublished data for Chinook salmon and steelhead. Note that any conclusions summarized in this section reflect those of the individual presentations. A later section of this report summarizes the workshop discussions and the conclusions of the workshop as a whole.

Background information and considerations for marker discovery and validation

Tasha Thompson, UC Davis

This presentation summarized the results of two recent papers that explored the genetic basis of adult migration phenotypes in steelhead and Chinook salmon. The first paper (Prince et al. 2017) used a genome wide association study (GWAS) based on RADseq data to identify regions of the genome statistically associated with premature vs. mature migration (i.e., summer- vs. winter-run in steelhead, and spring- vs. fall-run in Chinook). The primary goal of the study was to explore the genetic basis of the bimodal adult migration phenotypes typical of coastal rivers. The study used confidently-phenotyped samples (i.e., samples whose migration phenotype could be determined with high confidence based on information such as migration date, carcass recovery date/location, etc.) from coastal rivers spanning California, Oregon, and Washington. The principal results and conclusions of Prince et al. (2017) were:

- For both species, signals of population structure using RADseq markers across the whole genome were explained by geography rather than run-type, similar to what had been observed in previous studies based on fewer genetic markers.
- The GWAS in both species found that the same single region on chromosome 28 near or within the GREB1L gene was highly associated with adult migration phenotype in each species and in all locations within each species.
- A phylogenetic analysis of the data suggested a monophyletic origin of premature (early return) alleles within each species but different origins between species.
- The paper also reanalyzed data from an earlier steelhead study (Hess et al. 2016) and concluded that the same premature allele had been identified in that population (Klickitat River steelhead) and that heterozygous steelhead had an intermediate return-time phenotype.

The second paper presented (Thompson et al. 2019) extended the study of Prince et al. (2017) by conducting more extensive marker development and validation in the GREB1L region of Chinook salmon, and then using these new genetic markers to further explore the genetics of Chinook adult run timing in two systems where the spring-run has substantially declined—the Rogue and Klamath Rivers. In the Rogue River, the migration timing of

Chinook salmon in the upper basin has changed dramatically since the construction of Lost Creek Dam in 1977. Prior to the dam, Chinook counts at the entrance to the upper basin peaked in late May or early June; after the dam, counts peaked in late August to early October. In the Klamath River, the spring-run was historically equally or more abundant than the fall-run but has since declined to only a few hundred wild individuals in recent years, with complete extirpation of the spring-run in several major tributaries and the mainstem. The principal results and conclusions of Thompson et al. (2019) were:

- The more thorough development and validation of markers in the GREB1L region in coastal Chinook was critically important. The newly-identified markers were more tightly associated with run timing than the original markers found in the Prince et al. (2017) study. The Prince et al. (2017) markers were on the outskirts of the associated region, were not completely diagnostic, and had very high false-positive rates for the spring-run allele in some populations. This illustrates that thorough marker development and validation is essential, as using markers that are associated but not in complete linkage disequilibrium with the causative variant(s) can create noise which may lead to erroneous conclusions.
- Chinook sampled at the entrance to the upper basin of the Rogue during three time periods in 2004 had three different genotypic distributions at newly-developed GREB1L markers: late May samples were almost exclusively homozygous for the 'early' allele, early August samples were mostly heterozygous, and early October samples were mostly homozygous for the 'late' allele.
- Mid-September samples collected near the river mouth were entirely homozygous for the 'late' allele, suggesting that the early-October fish sampled further upstream but with 'early' alleles had likely entered fresh water earlier and then held below the collection site. This observation highlights the importance of carefully accounting for error and noise in phenotyping.
- The high frequency of heterozygotes in the early August samples indicated heterozygotes have in intermediate migration phenotype relative to homozygous early and late individuals.
- Selection modeling based on the time trend of reduction in the spring-run migration phenotype indicated that the 'early' allele would be predicted to persist for a long time if recessive (at a frequency of ~5.0%) but could be quickly lost if it is dominant or co-dominant and there is selection against early return timing.
- The predictions from modeling were empirically tested in two tributaries of the Klamath River, the Shasta and Scott, from which the spring-run phenotype was extirpated in the 1930's and 1970's respectively. The Salmon River (Klamath Tributary), which still hosts wild spring Chinook, served as a positive control. The 'early' allele frequency in the Salmon River was ~20%, consistent with the continued presence of the spring run in this watershed. In contrast, the 'early' allele was nearly absent (~0.2%) in the Shasta and Scott Rivers, where the spring run has been extirpated. The conclusion from this observation is that heterozygotes for the 'early' allele cannot be expected to act as a sustainable reservoir for future restoration of the spring-run Chinook in areas where the early run-timing phenotype has been extirpated.

Overview of adult migration timing in steelhead and Chinook of the Columbia River

Shawn Narum, CRITFC

This presentation summarized results of three recent papers on the genetic basis of run-timing variation in Columbia River steelhead and Chinook salmon. The primary results and conclusions of these papers are briefly summarized here.

Hess et al. (2016) conducted a GWAS analysis of adult run-timing variation in Klickitat River steelhead using RADseq data. The study found 3 strongly associated genetic markers, 1 of which mapped to the GREB1L region on chromosome 28 and the other 2 were unmapped at the time of the study. Subsequently, genome alignment indicates that all 3 markers are in the GREB1L region, and the same SNPs were identified in other steelhead populations examined by Prince et al. (2017).

Micheletti et al. (2018) used a whole-genome, pool-sequencing approach to greatly expand the number of genetic markers used for GWAS of adult return timing in Klickitat steelhead (68% of the genome surveyed compared to <1% surveyed in RADseq based methods). This study confirmed that the GREB1L region and the intergenic region between GREB1L and ROCK1 was highly associated with adult return time in steelhead. Based on analysis of PIT-tagged fish, this region highly associated with adult migration timing. Genotype frequencies for a candidate SNP marker from the intergenic region were surveyed throughout steelhead populations in the Columbia River and unexpected genetic variation was found for inland summer steelhead that had a high frequency of mature genotypes typically found in winter run steelhead in the coastal lineage. This was unexpected since inland steelhead populations primarily exhibit summer/fall entry to fresh water (Busby et al. 1996) and winter run fish are not observed in populations upstream of the Klickitat River. Based on analyses of individually PIT-tagged fish that were genotyped with this same candidate SNP, there was a consistent significant association with genotypes at this SNP marker and arrival timing to spawning grounds in steelhead returning to populations throughout the Columbia River Basin from both coastal and inland lineages of this species. This suggested a stronger relationship between the candidate marker and arrival-timing phenotypes as opposed to freshwater entry timing in steelhead of the Columbia River. A total of 13 markers have been developed spanning the GREB1L, intergenic, and ROCK1 region of Chr28 for future analyses in steelhead populations.

Narum et al. (2018) conducted a similar pool-seq whole genome sequencing study in Columbia River Chinook salmon, analyzing ~8M single nucleotide polymorphisms (SNPs) in samples from three major Chinook lineages: coastal, interior stream-type and interior ocean-type. The study included early (premature) vs. late (mature) migrating groups within each of the three lineages. The same GREB1L/ROCK1 region on chromosome 28 was highly associated with adult migration timing in all three lineages, including the final bimodal migration to the spawning grounds in the interior stream-type population that had an initially unimodal (early) return time to fresh water. Results indicated that ROCK1 may be particularly important for arrival migration timing in Chinook salmon, and demonstrated unexpected genetic variation for interior stream-type. The paper also reanalyzed previously

collected RADseq data to summarize patterns of a GREB1L region SNP in Chinook salmon samples from California to Alaska, finding the SNP to be polymorphic in many areas throughout the range. This SNP was one of the original Chinook SNPs identified in Prince et al. (2017) and is not as highly associated with return time as some subsequently identified variants but provided an initial geographic survey of genotype frequencies in this candidate region. Approximately 30 markers have been developed spanning the GREB1L, intergenic, and ROCK1 region of Chr28 for future analyses of Chinook salmon.

Genetic analyses of Chinook salmon run timing and associated traits

Eric Anderson, SWFSC, NMFS

This presentation provided an overview of work that has been ongoing at the Southwest Fisheries Science Center over the last two years. One paper describing that work is currently in review.

The first half of the presentation dealt with a whole-genome resequencing data set that includes 16 individuals from each of 10 collections: two spring-fall pairs from the Klamath Basin (Trinity River spring and fall, and Salmon River spring and fall), and 6 collections from the Sacramento-San Joaquin Basin, including two spring run collections, two fall-run collections, and one collection from each of a late-fall run and a winter run. Inspection of allele frequency differences revealed only a single region, roughly 140 kilobases (Kb) long, within the genome where all spring-run were fixed for specific variants which were different from the variants shared by all fall-run (and late-fall run) fish from the collections. Within a 30 Kb stretch, the winter run (another early-migrating stock) shared the same fixed variants with the spring-run fish. This region was termed the region of strongest association (RoSA).

Statistical phasing was used to infer haplotypes within and flanking the RoSA of all the fish in the sample, except for 14 that had very low sequencing read depths. This phased data set enabled a number of analyses. First, an heuristic approach, taking account of recombination, to infer local coalescent trees revealed two clear haplotypic lineages in the RoSA: an early (E) and a late (L) lineage. Within each lineage there were two main, distinct groups of haplotypes, with the Sacramento Basin possessing representatives from each group, while haplotypes of only one of the subgroups was observed in the Klamath, indicating greater diversity within haplotypic lineages in the Sacramento than in the Klamath Basin. Inclusion of sequence data from the Johnson Creek fish used for a recent Chinook salmon reference genome indicated that this early migrating fish of the interior Columbia River possesses a haplotype of the E lineage within the RoSA. Trees based on genetic distances between the sequences show the E and L lineages clustering with themselves, yet both clusters are greatly separated from coho salmon, the closest relative of Chinook salmon. The genetic divergence between E- and L-lineage haplotypes is about 6% of the divergence between either lineage and coho salmon (*Oncorhynchus kisutch*).

Through whole genome resequencing we identified two single nucleotide polymorphisms (SNPs) in the RoSA that are non-synonymous mutations within an exon of the GREB1L gene and which are fixed for alternate variants between the E and the L haplotype lineages. At both of these sites, the ancestral allele (i.e., the one carried also by coho salmon) occurs in the E lineage. While these sites are interesting candidates for causative polymorphisms, we also found a 10 Kb duplication flanking the RoSA that is strongly associated with migration timing. The region where this duplication occurs could be involved in regulation of either GREB1L, ROCK1, or both.

The second half of the presentation documented the SWFSC's extensive surveys of variation within the RoSA using targeted genotyping. We developed microhaplotype assays for 8 SNPs that reliably identify the haplotypes of a fish as belonging to the E or the L lineage. Genotypes from over a thousand fish from Oregon and California indicated that L-lineage haplotypes are found in fall-run fish, while E-lineage haplotypes are found in spring-run fish, and heterozygotes are found in all basins where both spring- and fall-run fish are known to occur. In the Klamath and Siletz Basins, heterozygotes were quite common amongst fish that were phenotypically recorded as spring run, and much less common amongst fish recorded as phenotypically fall run. This pattern was reversed in the Central Valley of California, where fish recorded as phenotypically spring run were rarely heterozygous at the RoSA, though heterozygotes were found amongst fish phenotypically recorded as fall run. This may indicate that the dominance pattern of variation at the RoSA is different in the different basins; however, it might also indicate a simple difference in how spring- and fall- run phenotypes are categorized in the two different basins.

The RoSA markers were applied to some 500 fish caught over a five month period (across two years, 2009 and 2010) in the Yurok tribal fishery in the Klamath River estuary. Fish are taken in this fishery directly upon (or very soon after) their arrival in fresh water. In both years there was a striking association between RoSA genotype and the time of freshwater entry, with RoSA genotypes accounting for 85% of the variance in freshwater entry timing. Entry times for the two different homozygous genotypes (EE and LL) were almost perfectly disjunct, with only a single day (July 16) upon which a single representative of both genotypes was caught. The heterozygotes returned to freshwater at an intermediate time which overlapped completely with either the EE or LL homozygotes. Although the timing of return to fresh water was clearly associated with RoSA genotype, we also looked for indications that RoSA genotype might predict differences in reproductive status (measured by gonadosomatic index, GSI) or fat content (measured by non-water fraction of the liver, NWF). Since both of these measures vary over time, we applied a linear mixed model to test for effects of RoSA genotype on these measures, while accounting for the date of sampling. Despite an abundance of power for detecting any direct effects of RoSA genotype on GSI or NWF, we found none, indicating that RoSA genotype might directly influence timing of freshwater return, but might not directly affect reproductive maturation or fat levels in the fish. Rather, these two characteristics (differences in reproductive phenology and fat content) of spring-run Chinook salmon may be observed simply because fish with different RoSA genotypes are typically sampled in fresh water at different times of the year.

We compared measurements of female gonadosomatic index (GSI) at the time of freshwater entry for fish en route to the Trinity River Hatchery (TRH) to measurements of GSI for fish that were deemed ripe and ready to spawn at TRH some time later (Figure 5). While, on average, spring-run fish have lower GSI values than fall-run fish at the time of freshwater entry, *both* the EE and LL fish see significant increases in GSI between the time of freshwater entry and the time of spawning. The authors who introduced the term "premature migration" for the migration of spring Chinook (and early migrating forms of other species) took pains to qualify that the terminology does not necessarily mean migration before maturity (Quinn et al. 2016); however the nomenclature has (not surprisingly) led some to describe fall-run Chinook salmon as being sexually mature before migration to freshwater (Prince et al. 2017, Thompson et al. 2018). Such a characterization is in clear disagreement with the data presented in Figure 5, and exemplifies the sort of confusion that the terminology of "premature-migrating" can engender.

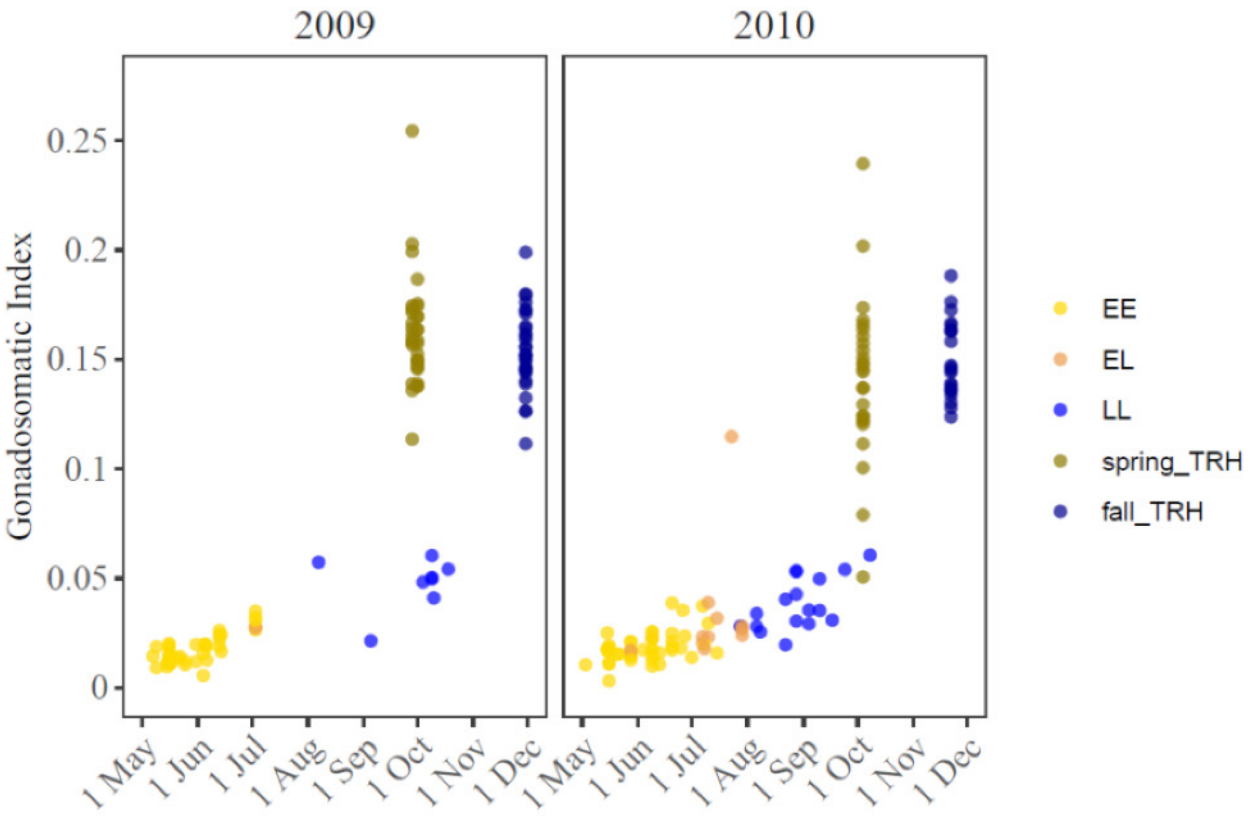

Figure 5. Female gonadosomatic index of Chinook salmon spawned at Trinity River hatchery (spring_TRH and fall_TRH) and female samples from the Klamath River estuary that assigned to the Trinity River by genetic stock identification analysis. Almost all estuary-sampled fish have considerably lower gonadosomatic index than active spawners at the Trinity River hatchery, regardless of their RoSA genotype. While LL homozygotes have, on average, slightly higher gonadosomatic index values than EE homozygotes (because LL fish enter freshwater later), both genotype classes undergo substantial, additional gonadal development prior to spawning. Figure courtesy E. Anderson.

RoSA genotypes were obtained from several hundred carcasses in the Salmon River (Klamath) over a number of years. The genotypes indicate that there is likely considerable spatial and temporal overlap in the spawning of EE and LL homozygotes, and that a large fraction of the spawners are heterozygous (EL) at the RoSA. In three years with sufficiently large samples to make a comparison, the frequencies of the three different genotypes (EE, LL, and EL) are in proportions close to what would be expected from random mating between fish of the three genotypes. The sample of carcasses results from separate surveys that are designed to target spring-run and fall-run fish separately, so it is difficult to infer exact patterns of spawning from these data, but, a survey design that focuses on late and early spawning periods is unlikely to artificially and greatly inflate the proportion of heterozygotes, so it is clear that numerous heterozygotes occur on the spawning grounds.

These patterns indicate the current distribution of the genotypes, but do not necessarily reflect their distribution during pre-European-contact times of the past. We assessed this issue by analyzing the frequency of fish with evident recombinations close to the RoSA using a coalescent-with-recombination simulation. Recombinations in that region today indicate

the occurrence of the E and L haplotypes, together in a heterozygous (EL) fish, at some time in the past. The observed frequency (around 0.22) of recombinants in the Klamath River is very unlikely to have arisen due to the occurrence of heterozygous (EL) fish solely in the last 200 years. Rather, concordant with the incredibly close genetic relationship between the ecotypes in the remainder of the genome outside of the RoSA, interbreeding between fall-run and spring-run fish has occurred historically in the Klamath and is a characteristic feature of the evolution of the two ecotypes in the Klamath. It was noted that our model was not designed to estimate changes in introgression rates between the pre-European-contact era and the present, but rather to reject the null hypothesis of a scenario in which there was no introgression between the spring and fall run prior to the last 200 years. Therefore, our results do not preclude the possibility that rates of interbreeding between spring- and fall-run fish have increased in the Klamath over the last 200 years. On the other hand, the data do allow a comparison between the Sacramento Basin and the Klamath Basin of historical frequencies of heterozygotes. It is quite clear that, historically, there were fewer heterozygotes (and consequently, likely less exchange between the spring and fall run fish) in the Sacramento Basin than there was historically in the Klamath Basin. This is concordant with data from the two basins throughout the genome outside of the RoSA.

Spatio-temporal distribution of Chinook carcasses from the Klamath and Rogue Rivers

Michael Miller, UC Davis

General Introduction

The presentation noted that premature migration is a paradox because there are so many disadvantages to this strategy, as summarized above and by Quinn et al. (2016): "This 'premature migration' reduces growth opportunities at sea, compels them to occupy much less productive freshwater habitats, and exposes them to extremes of flow and temperature, disease, and predation." The presentation also noted that the primary hypothesis for the advantage of premature migration, also from Quinn et al. (2016), is access to specific spawning and rearing habitats that are difficult for mature migrating fish to access because of physical factors such as temperature or flow.

Rogue River Introduction and Results

The upper Rogue was historically dominated by spring-run individuals. In other words, the upper Rogue was predominately spring-run habitat. Lost Creek Dam (LCD) was built in the late 1970s and caused decreased water temperatures and increased flows in the summer. Rogue Chinook have experienced a dramatic shift in migration timing since LCD was built, with fall Chinook and heterozygotes increasing at the expense of spring Chinook. Post-LCD fish counts (i.e., data from Gold Ray Dam) and 2014 carcass genotyping reveal that the extent of spatiotemporal segregation between spring and fall-run spawning has declined substantially since LCD was built. In addition to the artificial flow/temperature regime from LCD, modifications to Rainie Falls may also contribute to heterozygote and fall-run access to the upper Rogue. Lastly, the impact of Cole Rivers Hatchery (which has a spring-run hatchery program) is unknown but important to consider.

Salmon River Introduction and Results

The Salmon River has the last viable population of wild spring run in the Klamath Basin. The Salmon River has a natural(ish) flow regime but several historical low flow barriers have been modified. Carcass genotyping reveals that historical low-flow barriers that were modified in the 1980s no longer hinder fall-run migration, as many fall-run carcasses are found above these modified barriers. However, the upper South Fork Salmon remains predominately spring-run habitat, likely due to the long high-gradient section above Matthews Creek which appears to exclude most heterozygous and fall-run individuals. Lastly, the spawning time and location of heterozygotes is more similar to fall-run than spring-run, even though that is not the case for freshwater entry (where heterozygotes have an intermediate phenotype).

General Discussion

See Spatial distribution of run type genotypes in juvenile *O. mykiss* from the Eel and North Umpqua Rivers.

GREB1L variation in Chinook salmon from the Rogue and Sandy Rivers

Kathleen O'Malley, OSU/ODFW

This presentation described the spatial and temporal distribution of genetic variation in Chinook salmon in two Oregon Rivers, the Rogue and Sandy. The genetic data consisted of 298 genomically distributed and presumably selectively neutral SNPs (Hess et al. 2015), a single marker associated with sex (Hess et al. 2015), and the two SNPs located ~30 kb apart and just upstream of GREB1L (Thompson et al. 2019). In the Rogue River, tissue sample collections included (1) fin clips from 162 returning unmarked, naturally produced Chinook salmon caught in the lower river in 2019, (2) 445, 485 and 485 unmarked carcass samples from 10 survey reaches in the upper river in 2016, 2017 and 2018, respectively, and (3) fin clips from 1,575 fish used as broodstock at Cole Rivers Hatchery in 2018. In the Sandy River, samples were from collected from unmarked carcasses in 2015, 2016, and 2017 (n = 174, 187 and 159, respectively).

Results of the Rogue River study were summarized as follows:

- All but one naturally produced Chinook captured in the lower Rogue before May 23rd were homozygous for the early GREB1L allele (snp640165 and snp670329).
  - These results suggest that current ODFW fishing regulations, which prohibit harvest of naturally produced Chinook salmon (River Mile 0-125) prior to June 1, are protective of the majority of homozygous spring fish. Samples will be collected again in 2020 to determine if these findings are consistent across years.
- Results based on carcass sampling (2016-2018) showed that homozygous spring fish (snp640165) were frequently found in upstream locations from Cole Rivers Hatchery to Takelma Park. These homozygous spring fish were primarily sampled earlier in the season, in weeks 37 to 41 (Sept 10th – Oct 8th) (Figure 6).
- Most homozygous fall fish (snp640165) were sampled in week 40 (Oct 1st) or later.
- These results provide the first comprehensive, multi-year analysis of the spatial and temporal distribution of GREB1L alleles across the entire spawning period for Chinook salmon in the upper Rogue River (RM 125.5-157). ODFW sampled carcasses from every 4th fish in each year (2016-2018). In total, tissue samples were collected from 1,415 fish spanning the 10 survey reaches with sampling beginning on Sept 10th each year and concluding on November 4th. In an earlier analysis, Thompson et al. (2019) noted spatial and temporal overlap in homozygous spring, heterozygous, and homozygous fall fish. However, because of the small sample size (N=86) in 2014 and the numerous differences in sampling protocol between the 2014 and 2016-18 studies, significant caution should be used when making comparisons between these years. Specifically, the samples in 2014 were collected during a period (Sept 22-Oct 29) that did not include carcasses from the first three weeks of spawning, when spring-run fish typically predominate. Additionally, in 2014 surveyors were intentionally selecting fish based on appearance of the carcass (versus simply sampling every 4th fish encountered, as in 2016-18). Last, the small sample size in 2014 likely does not adequately capture diversity present within time periods or reaches, and thus cannot be considered representative of the run.

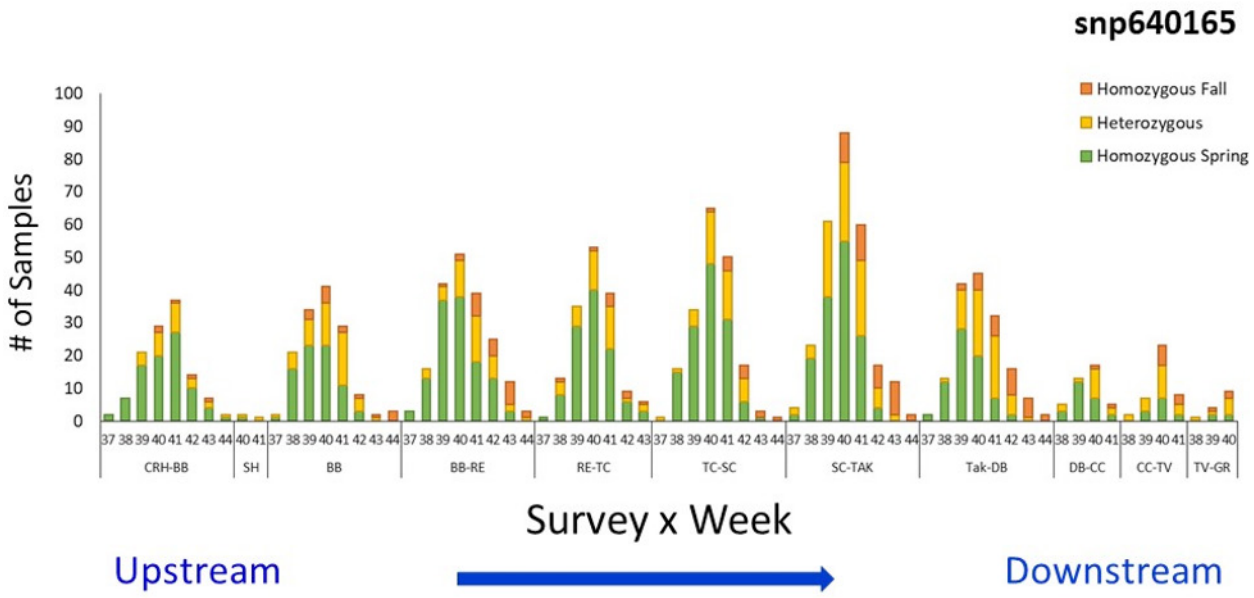

Figure 6. Distribution of GREB1L SNP1 genotypes across survey reaches and time with all three years combined (2016–2018). Greb1L SNP1 (snp640165) is more diagnostic of adult migration phenotype in Rogue River Chinook salmon than SNP2 (snp670329) (T. Thompson, pers. comm.). The Julian week when carcass samples were collected is on the x-axis and ranges from 37 (Sep 10th–16th) to 44 (Oct 28th–Nov 4th), grouped by survey reach. The most upstream survey location is Cole Rivers Hatchery (CRH) and the furthest downstream location is the old Gold Ray Dam site (GR). Number of carcass samples collected is on the y-axis. Figure courtesy K. O'Malley.

- Most of the Cole Rivers Hatchery broodstock were homozygous for the early GREB1L allele. However, a fraction (11%) of the hatchery broodstock samples were heterozygous, and very small fraction (0.3%) were homozygous fall. The homozygous fall fish were collected after August 15th. These results are based on snp640165 which is reportedly more diagnostic of adult migration phenotype than snp670329 (per comm T. Thompson).
- The results from the Sandy River study were summarized as follows:
- There were two genetic clusters, early and late, based on 260 presumed neutral SNPs (Figure 7). These two clusters separate based on collection date and location but not based on variation at GREB1L.
- The early spawning collection contained a mix of GREB1L genotypes (homozygous spring, heterozygous, and homozygous fall) whereas the late spawning collection consisted primarily of homozygous fall GREB1L genotypes.
- Genetic stock identification indicated that the early and late carcass sample collections correspond to spring (Willamette River spring-run) and fall (West Cascade fall-run) genetic reporting groups, respectively.

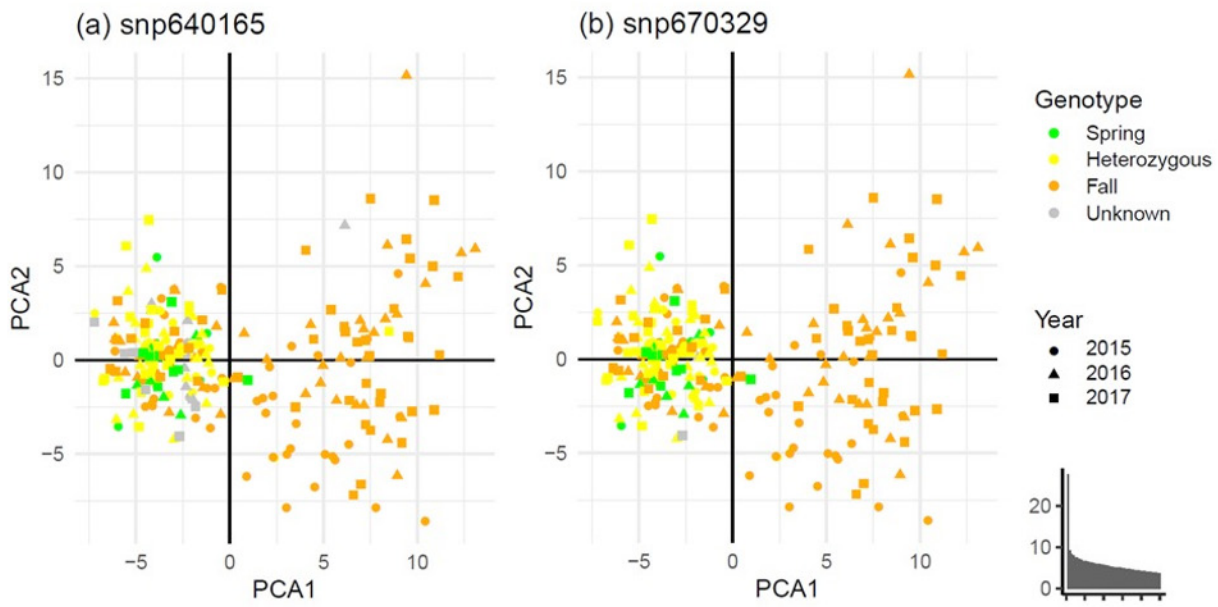

Figure 7. Principal components analysis of Sandy River Chinook salmon carcass samples (all years, n =237) genotyped at 260 presumed neutral SNPs performed in R package `adegenet`. The first PCA axis explained 5.4% of the variation and the second axis explained 1.8% of the variation. Eigenvalues for the first 50 axes are shown in the inset bar plots. Shapes indicate the year samples were collected. (a) Points are shaded to indicate the GREB1L genotype (snp640165) of each carcass sample. (b) Points are shaded to indicate the GREB1L (snp670329) of each carcass sample. Figure courtesy K. O'Malley.

Analysis of run-type field calls in the Chehalis River Basin, Washington

Tasha Thompson, UC Davis

This presentation described patterns of genetic variation at run-timing-associated loci and neutral genomic loci from Chinook salmon in the Chehalis River presented in a report for Washington Department of Fish and Wildlife (Thompson et al., 2019b). The Chehalis River is the largest watershed found entirely in Washington State. It has a long history of environmental degradation and habitat modifications that have likely influenced past and present fish distributions, including extensive splash-dam logging operations in the late 19th and early 20th centuries, contemporary dams, and modifications to natural barriers.

The Chehalis River has both spring and fall-run Chinook salmon, but accurately estimating the abundance of each run type is difficult due to the lack of a common access point (e.g., a weir or fish ladder). Instead, the runs are estimated from spawning surveys, but the rate of classification errors (e.g. counting fall as spring or vice versa), and thus the accuracy of these methods, is not known. Genetic markers for run type may therefore be a useful tool for improving monitoring of run-specific abundance. The principal results and conclusions of this presentation were:

- **The GREB1L run-type markers developed in Thompson et al. (2019; see above) work well to distinguish spring- and fall-run Chinook in the Chehalis.** The markers were validated with a set of confidently-phenotyped samples collected from fisheries in the mainstem Chehalis River in May/June (spring-run control) or Grays Harbor in October (fall-run control). The markers were found to be highly associated with run type in this population, with spring-returning fish nearly entirely homozygous for the early allele and fall-returning fish nearly entirely homozygous for the late allele.
- **The accuracy of fall-run field calls was relatively high, but the accuracy of spring-run field calls was low.** A sample set consisting of 300 "fall" field-called fish and 146 "spring" field-called fish collected between 2001 and 2016 was genotyped at the validated run type markers. Of the samples field called as fall-run, 89% were homozygous for the fall-run allele (10% were heterozygous and 1% were homozygous spring-run). Of the samples field called as spring-run, only 48% were homozygous for the spring-run allele (18% were heterozygous and 34% were homozygous fall-run). The high error rate among spring-run field calls suggests there are likely fewer spring-run Chinook in the Chehalis Basin than are currently estimated. However, the samples in this data set were generally collected ad hoc during spawning surveys (as opposed to being collected systematically across space and time), so further work is needed to determine the extent of this bias.
- **The accuracy of spring-run field calls exhibited strong spatial patterns.** Spring-run field calls were most accurate in the Newaukum River (an upper basin tributary), and least accurate in the Satsop and other lower basin tributaries. Heterozygotes were primarily observed in tributaries where homozygous spring-run individuals were also observed, and samples collected in the lower watershed (below the Skookumchuck River) were almost exclusively homozygous fall-run regardless of field call. These results demonstrate that run-type field call accuracy can vary substantially across space.

- **Most accurate spring-run field calls were collected ~15-20 years ago.** Therefore, it is unknown whether the reported patterns represent current conditions because a change in the relative abundance of the spring-run over time would likely influence the accuracy of field calls. Lack of consistency in sample collection effort over the years precludes direct comparisons between genotype frequencies in the oldest and newest samples in our data set, so a new systematic survey of run-type genotype distribution in the basin would be extremely valuable.
- **Annual analysis of fry or smolt samples along with greater analysis of carcass samples could potentially be used to more accurately monitor the abundance of the spring- and fall-run types in this and other systems.**
- **A RADseq analysis of population structure at neutral loci found that the greatest differentiation between spring and fall Chinook salmon existed in samples from the Newaukum River.** The presentation concluded that as spring-run fish decline, introgression from fall-run fish is likely to increase, resulting in less differentiation between the groups than may have existed historically (when spring-run were more abundant).

Validation and association of candidate markers for adult migration timing and fitness in Chinook salmon

Ilana Koch and Shawn Narum, CRITFC

To provide a more thorough test of individual phenotypic association within and among lineages of Chinook Salmon, 33 candidate markers were developed across a 220 Kb region on chromosome 28 previously associated with migration timing. Markers spanned GREB1L, intergenic, and ROCK1 genes on chromosome 28. Along with these candidate markers, neutral markers were genotyped in individuals from representative collections that exhibited phenotypic variation in migration time from each of three lineages of Chinook Salmon. Association tests confirmed the majority of markers on chromosome 28 were significantly associated with migration timing and the strongest association was consistently observed for markers within the ROCK1 gene and the intergenic region between GREB1L and ROCK1. Candidate markers alone explained a wide range of phenotypic variation for migration timing in Lower Columbia and Interior ocean-type lineages (29% and 78%, respectively), but less for the Interior stream-type lineage (5%). Individuals that were heterozygous at markers within or upstream of ROCK1 had migration phenotypes that suggested a pattern of dominant inheritance for early migration across populations. Finally, previously published fitness estimates from the Interior stream-type lineage enabled tests of association with migration timing and two candidate markers, which revealed that fish with homozygous mature genotypes had slightly higher fitness than fish with premature genotypes, while heterozygous fish were intermediate.

Additional results and conclusions:

- Association tests confirmed that the majority of chromosome 28 markers were significantly associated with migration timing.
  - The strongest association was found for markers within the ROCK1 gene and the intergenic region between GREB1L and ROCK1.
- SNPs explained a large percent of phenotypic variation for Lower Columbia and Interior ocean-type lineages, but a lower percentage in the Interior stream-type.
- Annual migration timing patterns were highly variable in the Interior stream-type lineage.
  - The environment could contribute substantially to phenotypic variation in this population.
- For Lower Columbia and Interior ocean-type lineages, the percent of premature alleles in the late-migrating populations was close to zero at or near the top significant SNPs.
  - However, there were premature alleles found in the Interior stream-type within late migrating fish.
- Heterozygous individuals were skewed towards premature migration timing, potentially representing dominant inheritance for premature alleles in ROCK1.
- Individuals that demonstrate earlier migration timing in the Interior stream-type potentially have lower fitness relative to late-migrating individuals.
  - There were significant differences between alternate homozygotes (not heterozygotes).

Summary of run timing genome association study in Wenatchee spring Chinook salmon

Michael Ford, NWFSC, NMFS

The NWFSC, in collaboration with WDFW, has conducted a long-term pedigree study of spring Chinook salmon in the Wenatchee River. All spring Chinook salmon migrating upstream of Tumwater Dam are sampled at the dam, including measurement of length, weight, age, phenotypic sex, date of sampling at the dam, and natural or hatchery origin. A fin clip is also taken for genetic analysis. After sampling, the fish were either released above the dam to spawn naturally, or in some cases were collected for hatchery broodstock. The date of sampling at the dam provides a measure of upstream migration timing to the spawning areas, and that is the primary trait of focus in this analysis.

Starting in 2004 and continuing to the present (2019), approximately 55,000 upstream migrating spring Chinook salmon have been sampled at Tumwater Dam. Most of these (through return year 2016 and ongoing) have been genotyped either for 16 microsatellite loci or 96 SNP loci, or both, primarily for the purpose of parentage analysis. In addition, a sub-sample of 575 fish have been genotyped using RADseq, and these fish form the basis of the analysis described here.

The RADseq genotyped samples were initially selected to be approximately representative of the run-timing distribution, but due to DNA quality issues the final set of genotyped samples varied among years with respect to how closely the sample distribution matched the entire return time distribution.

A GWAS identified numerous SNPs potentially associated with variation in return time, including 6 SNPs with -log(p-values) >10 and MAF > 0.05, out of 32342 total mapped SNPs in the analysis. One of the most significant SNPs, position 12242006 on chromosome 28, has been previously identified as a variant associated with Chinook salmon return time in a number of prior studies (SNP "2", scaffold79929e:595121 in Prince et al. (2017) and Ots 28 position 11033626 in Narum et al. 2018). Only the chr28:12243006 SNP is within a characterized gene – an intron of GREB1 gene (Prince et al. 2017). Note that other SNPs in this same genomic region have subsequently been identified that are more tightly associated with return time in some populations (Thompson et al. 2019, Narum et al. 2018).

Combined across all years, there is a clear relationship between the alternative genotypes at chr28:12243006 and return time, and a similar pattern is seen for both sexes and for both hatchery and natural fish. When plotted separately for each return year, considerable year-to-year variation becomes apparent. The other highly significant SNPs show varying degrees of association with return time. If effects across loci are assumed to additive and the total number of "late" alleles are plotted against return time in each year, a pattern of roughly additive variation is seen only in the 2008 return year.

Summary:

- A GREB1L variant previously associated with run timing segregates in Wenatchee spring-run Chinook and is also associated with run timing.
- The "late" allele is at relatively high frequency (>20%), and higher than in later returning summer/fall fish returning to the same river.
- Other loci also contribute to run-timing variation.

GREB1L variation in California steelhead

Devon Pearse, SWFSC, NMFS

This presentation described data and results on the relative positions of SNPs mapped to the rainbow trout genome (Pearse et al. 2019a) located in the GREB1L region of chromosome Omy28, providing a direct comparative basis among studies that have used different SNPs in this region (Hess et al. 2016; Prince et al. 2017; Pearse et al. 2019b). The presentation then described results on SNP variation in more than 30 geographically distributed samples of wild and hatchery *O. mykiss*, as well as temporally-distributed samples from three locations in the Klamath and Eel Rivers. The presentation discussed the dynamic variation in natural river features, highlighting the importance of interannual and longer-term variation in flow, river geomorphology, and other factors influencing relative fitness and demographic variation among life-history ecotypes and the importance of the portfolio effect in maintaining biodiversity and genomic variation.

Summary and primary conclusions:

- Variation in the GREB1L region associated with run timing is *genomically* distributed across a complex ~200KB region on chromosome Omy28. The causative variation is not known within this region. Different studies have used different SNPs within this genomic region; it is important to standardize across studies to make valid comparisons.
- Variation in the GREB1L region associated with run timing is *geographically* distributed throughout California and other areas, with multiple haplotypes and genotypes observed in many populations, including those above and below barriers to anadromous migration and in coastal and inland rivers.
- Variation in the GREB1L region associated with run timing is *temporally* distributed, with overlap in sampling of both alleles and all three genotypes observed within and among samples in the Klamath and Eel Rivers.
- Relevant link, with downloadable PDFs of the Panel Report (Pearse et al. 2019b) and 12-month finding on the NC steelhead petition: https://www.fisheries.noaa.gov/action/12-month-finding-petition-list-summer-run-steelhead-northern-california-endangered-under

Spatial distribution of run type genotypes in juvenile *O. mykiss* from the Eel and North Umpqua Rivers

Michael Miller, UC Davis

General Introduction

See Spatio-temporal distribution of Chinook carcasses from the Klamath and Rogue Rivers.

North Umpqua Introduction and Results

The North Umpqua River supports a world-famous steelhead fishery, with summers being the most unique/famous run and providing fishing opportunities early-summer through late-fall. Although Steamboat Creek is often regarded as the most important tributary for summers (because many adults over-summer in the creek), the spawning and rearing distribution of summer and winter steelhead in the North Umpqua is not well understood. Steamboat Creek has potential high-flow barriers (e.g., Little Falls, Steamboat Falls) that could be hindering winter-run access to its upper reaches. However, a fish ladder built at Steamboat Falls in the late 1950s and modified in 2012 could be facilitating winter-run access. There is a striking difference in the spatial distribution of winter and summer juveniles throughout the North Umpqua basin: middle and upper Steamboat Creek appear to be exclusively used by summer-run, whereas other locations throughout the North Umpqua Basin are primarily winter-run. Thus, the fish ladder on Steamboat Creek does not appear to have facilitated winter-run access, consistent with the idea that falls/rapids in lower Steamboat are acting as high-flow barriers which exclude winters (even during relatively dry winters). Lastly, the impact of the summer-run program at Rock Creek Hatchery (e.g., in contributing to the number of heterozygotes and summers outside of Steamboat Creek) is not known.

Eel River Introduction and Results

The Eel River supports the southernmost extant population of summer-run steelhead. There are strong spatial differences in the spawning and rearing distribution of summer and winter steelhead, with summer-run juveniles being primarily found above flow dependent barriers in both the Van Duzen and Middle Fork Eel rivers. The Mainstem Eel above Scott Dam appears to have historically supported summer-run steelhead (also in an area above a flow dependent barrier), and their alleles have persisted since dam construction. The summer-run allele does not apprear to have been maintained in the South Fork since the phenotype was extirpated, suggesting the summer-run allele does not persist in the absence of the summer-run phenotype in anadromous waters, even when a relatively healthy resident population is present.

General Discussion

Our results support Quinn et al. (2016)'s primary hypothesis that access to exclusive or nearly-exclusive spatiotemporal habitat (i.e., specific habitat that is difficult for mature migrators to access due to physical factors such as temperature or flow) can be a major advantage of premature migration that offsets its disadvantages. Premature migrators still

have exclusive or nearly-exclusive spatiotemporal habitat in some places (e.g., Steamboat Creek on the North Umpqua [steelhead] and the upper South Fork Salmon [Chinook]), but the amount and/or degree of such habitat has likely been substantially reduced in many locations (e.g., most of the Salmon and Rogue rivers). In locations with historically exclusive or nearly-exclusive premature migrator habitat that is now accessible to mature migrators, management and restoration actions that decrease mature migrator access (e.g., promoting natural flow and temperature regimes, restoring flow dependent barriers, etc.) could greatly benefit premature migrators. In places where premature migrators retain exclusive or nearly-exclusive habitat (e.g., Steamboat Creek), maintaining this exclusiveness (e.g., don't modify barriers) and improving the quality of this habitat (e.g., through habitat restoration projects) will promote the persistence of the premature migrators into the future.

Mature migrator access to habitat that was historically exclusive or nearly-exclusive to premature migrators could be especially problematic when the premature migrating population is much smaller than the mature migrating population. In other words, the fewer premature relative to mature migrating individuals there are in a location, the more potential there is for mature individuals to "swamp out" the premature individuals (e.g., through the creation of heterozygotes). For example, if there are 1,000 spring-run individuals, 2,000 fall-run individuals, and 5% of the fall-run individuals make it into the spring-run habitat, any particular spring-run individual would have only an approximately 10% chance of mating with a fall-run individual. However, if there are 100 spring-run individuals, 2,000 fall-run individuals, and 5% of the fall-run individuals make it into the spring-run habitat, any particular spring-run individual would have an approximately 50% chance of mating with a fall-run individual. Lastly, if there are 100 spring-run individuals, 2,000 fall-run individuals, and 50% of the fall-run individuals make it into the spring-run habitat (because the historical factors that had previously hindered fall-run access had been modified), any particular spring-run individual would have an approximately 90% chance of mating with a fall-run individual. These types of situation are expected to be especially problematic to premature migrating populations.

Migration timing marker associations in Columbia River steelhead

Stuart Willis, CRITFC

As life-history diversity plays a critical role in supporting the resilience of exploited populations, understanding the genetic basis of those life-history variations is important for conservation management. However, effective implementation of these discoveries requires a robust understanding of the strength and universality of genetic associations. Here, we examine genetic variation of single nucleotide polymorphism markers in candidate regions associated with migration phenology in steelhead (*Oncorhynchus mykiss*) from the Columbia River. We found chromosome 28 markers explained a significant amount of variance in migration timing in both coastal and inland steelhead. However, the strength of association was much greater in coastal steelhead, suggesting that genomic background and the challenging migration pathways that inland steelhead must traverse may moderate the effects of this region. While these results lend support to the use of these candidate regions in predicting life-history characteristics, we suggest that further data on stock specific associations and haplotype frequencies will be useful in guiding implementation of genetic assays to inform management.

Recolonization following dam removal: Observations on genetic and life-history variation in *Oncorhynchus mykiss* in the Elwha River, Washington, USA

Krista Nichols, NWFSC, NMFS

This presentation described a study using RADseq and GT-seq data to monitor recolonization of steelhead into the upper Elwha River following dam removal (which was completed in 2014). The data consisted of samples from below, above and between the two old dam sites, from before (2004-2012), during (2013-2014), and after dam removal (2015-2019), for a total of 2710 samples genotyped for RADseq loci and 591 for 288 GT-seq loci. The RADseq dataset encompassed samples collected both before and after dam removal, while the GTseq dataset included 2018 and 2019 steelhead and upper watershed collections from within the Olympic National Park.

Summary and conclusions:

- A preliminary GWAS analysis using RADseq markers in 2015 returning adult steelhead identified the GREB1L marker on Omy28 as significant (include samples size and numbers of summers/winters here).
- The 'early' GREB1L alleles overlapping between the RADseq and GTseq datasets include only marker Omy28_11667915. Prior to dam removal, *O. mykiss* sampled from the portion of the Elwha River and tributaries above the dams had a high proportion of 'early' alleles, while returning steelhead below the dams (pre-dam), had few to none of these alleles. Once the dams were removed, more of the early alleles appear in the returning steelhead, suggesting that the upper watershed is producing many of the hundreds of steelhead (including many summers) that have returned in recent years.
- Recolonization of steelhead in the Elwha River to the upper watershed has been rapid following dam removal.
- Genetic results support 're-awakening' of summer steelhead, likely owing to the harboring of alleles for run timing in freshwater resident populations above the dams.
- Polymorphisms on Omy28 and Omy05 are the primary drivers of genetic differentiation within the system, but the inversion polymorphism on Omy05 does not conform to association with residency/anadromy.

Genetic variation in Puget Sound steelhead

Ken Warheit, WDFW

The presentation described RADseq data from Skagit and Nooksack River steelhead, including both wild fish and highly domesticated (Chambers Creek) hatchery fish. The wild samples in both watersheds (n = 24 and 14, respectively) were entirely homozygous for a "late"-associated allele at the GREB1L region. The Chambers Creek hatchery fish from both watersheds had all three genotypes (EE, EL, LL; total n = 54 and 23 for the Skagit and Nooksack, respectively). Separate samples of summer run hatchery steelhead originating from the Lower Columbia River (Skamania, Kalama) were entirely homozygous for an early-associated GREB1L allele.

Workshop Discussion

Following the presentations, the workshop participants discussed several topics, including identifying areas of scientific agreement and disagreement. We also discussed areas of scientific uncertainty and research that could be conducted to address those uncertainties. The workshop ended with a discussion of the conservation implications of the new work on the genetic basis of adult run-timing variation. Most of the workshop discussion focused on Chinook salmon, although many of the issues raised apply to steelhead as well. Issues that were discussed that apply specifically to steelhead are summarized separately at the end of this section.

Areas of Scientific Agreement, Disagreement, and Uncertainty

Prior to the workshop, the organizers identified and distributed a number of questions for the participants to consider, which form the subheadings below. Note that the considerations below are for both steelhead and Chinook salmon, except where noted.

• Is the GREB1L/ROCK1 region responsible for adult migration timing, and if so by what mechanism?

Areas of agreement

A single region in the genome has a strong statistical association with adult run timing. Based on both published and unpublished data, multiple studies have identified one ~200 Kb region of chromosome 28 (near two genes called GREB1L and ROCK1) with alleles strongly associated with various measures of adult run timing in multiple populations of Chinook salmon and steelhead. The populations examined range from coastal populations in California and Oregon, to interior Columbia River, to the Straits of Juan de Fuca and Puget Sound. The initial studies (Hess et al. 2016; Prince et al. 2017) were based on relatively sparse (<1%) genome coverage using RADseq markers, and identified only a few associated markers in the region. Subsequent studies with more complete genome coverage either for the whole genome or targeted at the GREB1L region (Micheletti et al. 2018; Narum et al. 2018; Thompson et al. 2019) have found even stronger statistical associations between GREB1L/ROCK1 region markers and run-timing phenotypes. The causal genetic variants remain unknown, but the peaks of statistical association with run timing are generally highest in the intergenic region between the GREB1L and ROCK1 genes (i.e., in the regulatory regions of these genes).

The migration phenotype measured across prior studies is not standardized, and efforts should be made to do so. The workshop participants emphasized that migration is a complicated trait and that it is important to be clear what phenotype is being described. For example, the strength of genetic by life-history associations may differ for the same population when run timing is measured at different locations, such as the river mouth or further upstream. Some participants also noted that the 'decision' on when to migrate

takes place months before freshwater entry, making even the point of freshwater entry a somewhat arbitrary marker of a complex phenotype. In general, participants agreed that, at least for coastal populations, genetic associations with adult run timing are likely to be strongest when run timing is measured as the time of freshwater entry, while measurements taken at more upstream locations are likely to be statistically noisier and require further investigation. For example, some spring Chinook salmon may over-summer below a data collection site and therefore not be observed until making an upriver push just before spawning (e.g., Thompson et al. 2019).

Marker development, validation, and standardization is extremely important. Marker choice is very important—for example, some run-type markers have weaker associations for spring-run alleles in some populations, especially if they flank the GREB1L/ROCK1 region. Furthermore, marker development and validation for steelhead is not as well developed as it is for Chinook salmon, but markers for steelhead that span GREB1L/ROCK1 have been developed by multiple labs and are being tested broadly. Further testing and standardization of markers is a crucial area of ongoing/future work, as is the use of standard nomenclature for and alignment of markers (based on genome alignment) across studies (see Appendix).

Areas of uncertainty

The causal variant(s) for adult run timing remain to be identified. The primary uncertainty related to this question is that the causal variants within the region remain unknown. Inferences regarding the effects of the region on run timing are solely based on statistical associations, which can be affected by a variety of population genetic processes, as well as experimental design. Based on the studies to date, it is nearly certain that the GREB1L/ROCK1 region contains causal variants, but there are numerous SNPs within this region, most of which are likely to be neutral with respect to the physiology of run timing and only associated through linkage. Population history (e.g., genetic drift, and the locations of recombination events flanking the causal locus/loci) can affect these patterns of linkage disequilibrium. Direct tests for functional significance are an important future area for research. However, many experimental approaches to directly test function in model organisms will be extremely challenging or impossible to apply to the study of migration characteristics of salmon. Functionally, the GREB1L gene appears to be a good candidate to influence traits related to sexual maturity, but there is at least one other characterized gene in the same region (ROCK1) and the details of how genetic variation in this region contributes functionally to the variation in physiology necessary for run-timing variation remains unknown. It is also possible that each of these genes may play distinct roles in different phases of adult migration such as returning from the ocean, freshwater entry, and arrival for spawning. Additionally, there may be many loci outside of the GREB1L/ROCK1 genomic region with small effects on run timing. Quantifying the relative contributions of the GREB1L/ROCK1 region, loci outside this region and environmental effects on the adult migration phenotypic variation (in multiple populations across the range), could advance understanding of the scope for adaptation to ongoing and future selection pressures. Ideally, such work should be conducted using accurately and thoroughly phenotyped samples to provide a means of studying different aspects of the “early” and “late” phenotypes.

- **What is the distribution of genetic variation for adult migration timing in space and time?**
- **Do the genes associated with migration timing have the same effect in populations inhabiting different environments and with different genetic backgrounds?**

Areas of agreement

The GREB1L/ROCK1 association with run timing is best characterized in US West coastal populations for both Chinook salmon and steelhead, and to some degree in the Columbia River basin. For Chinook salmon, the GREB1L/ROCK1 region and its association with adult run timing has been best characterized in US West Coast and Columbia River watersheds, including the Sacramento/San Joaquin, Klamath, Rogue, Nooksack, Puyallup, and Chehalis Rivers on the coast and the Cowlitz, Lewis, McKenzie, Clearwater, Deschutes, Yakima, Methow Rivers, Johnson Creek and Priest Rapids Hatchery in the Columbia; smaller samples have also been analyzed from some other rivers. Characterization of GREB1L/ROCK1 variation in steelhead has also largely focused on coastal and Columbia River watersheds. While there are many more locations that will be interesting and important to explore further (e.g., Canada and Alaska), GREB1L/ROCK1 variants are strongly associated with early vs. late migration across diverse geographic and ecological environments. Most workshop participants agreed that the frequency of 'early' alleles among wild coastal Chinook populations is likely to vary over various time scales, due to a combination of random (genetic drift) and directional (selection, migration) factors.

Areas of uncertainty

Our current understanding of both the contemporary and historical distribution of genetic variation in GREB1L/ROCK1, in association with run timing, is confounded by issues with phenotyping, influence of hatchery populations, and anthropogenic activities influencing access to habitat across space and time. Even in watersheds that have been relatively well studied, issues related to uncertain phenotyping, the uncertain influence of hatchery populations, and in some cases sparse temporal sampling leads to uncertainty about the true frequency of alternative GREB1L/ROCK1 alleles in many Chinook salmon and steelhead populations. Salish Sea and more northern populations in British Columbia and Alaska also remain relatively poorly characterized for variation in the GREB1L/ROCK1 region, although one GREB1L-linked marker has been shown to be polymorphic in Chinook salmon throughout the North American range (Narum et al. 2018). This marker is not within the region of highest association in well studied populations, however, so variation in this particular SNP does not necessarily imply widespread variation in the genetic variants responsible for variation in run timing.

Studies on Columbia River populations also suggest that the genetic basis of run timing across the species range may be more complex than is understood to date. For example, some interior spring/summer Chinook populations, such as Johnson Creek, segregate for variants at GREB1L/ROCK1 that are evolutionarily related to 'early' and 'late' alleles in coastal populations (Narum et al. 2018). The Johnson Creek population has a unimodal

"early" freshwater time of entry (compared to any coastal or Columbia River fall-run populations), but exhibits bimodal return timing to spawning grounds that is associated with early and late alleles at GREB1L/ROCK (Narum et al. 2018; Koch and Narum, in press). Results from these studies indicate that freshwater entry timing and arrival at spawning grounds may be two different phenotypes which are highly correlated in coastal and interior ocean-type lineages, but are uncoupled for interior stream-type populations. There are also distinct patterns of linkage disequilibrium in this region of chromosome 28 that suggest two distinct haplotype blocks for the interior spring/summer lineage instead of one block as seen in other lineages of Chinook salmon (Koch and Narum, in press). Thus, these two candidate genes and their regulatory regions may have a distinct effect on each of these two phenotypic traits with ROCK1 more directly associated with timing of arrival to spawning grounds than freshwater entry. Finally, it is clear that variation at the GREB1L/ROCK1 region contributes to both variation within and among the major Columbia River Chinook salmon lineages, but current studies do not rule out the possibility that other genomic regions also play a role for some of the major run-timing difference between the major lineages. Interior Columbia steelhead (which all have relatively early freshwater entry times) also appear to have a complex and not-fully-understood relationship between adult migration phenotypes, related traits, and variation in the GREB1L/ROCK1 region (Micheletti et al. 2018 and Willis presentation).

- **What is the pattern of dominance among haplotypes in the GREB1L/ROCK1 genomic region?**
- **What phenotype do heterozygotes express, and what is their fitness compared to homozygotes?**

Areas of agreement

Heterozygotes are likely an important mechanism for the spread and maintenance of the early migration alleles over long time scales. The workshop participants agreed that understanding the dominance patterns for run-timing variation associated with the GREB1L/ROCK1 region is important for evaluating the likely persistence of the 'early' allele(s) if environmental conditions favor late run timing. The participants also agreed that heterozygotes were likely an important mechanism for the spread and maintenance of spring-run Chinook salmon among coastal populations and are likely to be important in the future as habitat favoring early return times is restored. The participants also agreed that the dominance patterns at the GREB1L/ROCK1 region may be complicated and depend on both the evolutionary lineage within a species and how the phenotype is characterized (e.g., freshwater entry vs. spawning time and location).

For Chinook salmon, the empirical data so far appear to be consistent with either an additive model or dominance of the early allele. No existing data sets have found strong evidence that the early phenotype is recessive. In addition, in some coastal locations where spring-run Chinook have been largely extirpated (Shasta River, Scott River, Iron Gate, and Wynoochee Rivers) or were never known to exist (Eel River?), the early alleles are absent or extremely rare (Thompson et al. 2019), which would not be expected if the premature

allele were recessive. For both species, robust conclusions regarding dominance can be difficult in light of uncertainties (discussed above) associated with accurately assigning run-timing phenotypes to individual fish.

Areas of uncertainty

It may be too simplistic to focus on dominance of migration timing alone since genetic variation at the GREB1L/ROCK1 region also could influence other traits that are more difficult to study. In addition, it seems likely that successful expression of the spring/early run time phenotype requires a host of additional adaptations, such as appropriate egg and juvenile growth regimes for the upstream habitats and that the genetic basis of these adaptations has not been characterized.

- **In what circumstances is it reasonable to conclude that the current distribution of GREB1L genes accurately reflects historical (pre-European contact) patterns?**
- **When/where is that not a good assumption?**

Areas of agreement

Interaction between individuals with variable run timing has occurred historically, is expected, and likely varies depending on historical environmental conditions. However, anthropogenic impacts have also likely changed these interactions in many locations. Workshop participants agreed that some interbreeding between spring and fall coastal Chinook salmon certainly occurred naturally, but that the degree of interbreeding in many coastal areas has likely increased over the past 100 years as spring run have declined and habitat alterations and other human actions have increased the potential for spawning overlap between spring and fall runs.

Participants agreed that interbreeding between runs likely occurred historically (i.e., pre-European immigration) in many or most locations, but estimating precise natural/historical levels of interbreeding is challenging. For example, an analysis of recombination patterns in the Salmon River (Klamath) rejected the hypothesis that zero interbreeding occurred between spring and fall runs prior to 200 years ago, but did not distinguish between levels of historical interbreeding (e.g., 1% vs 25%; Anderson presentation). In addition, salmon habitat is dynamic over a variety of temporal scales even in pristine watersheds, and thus natural levels of interbreeding have likely varied over time. However, in many locations, there are strong indications that human-driven habitat modifications have increased opportunities for interbreeding. Substantial numbers of heterozygotes have been observed in contemporary samples from the Salmon (Klamath, CA), Rogue (OR), and Chehalis (WA) River basins, indicating high levels of current and/or recent interbreeding among fall and spring-run fish. For example, in the Salmon River, the mature, heterozygous, and premature genotypes were found in nearly Hardy-Weinberg equilibrium proportions in one data set, suggesting spring and fall Chinook salmon are currently interbreeding at a high rate. However, documented habitat alteration in the Salmon River (e.g., modification of Bloomer

Falls and other low flow barriers that previously hindered fall-run migration) has likely increased the opportunity for interbreeding compared to historical times (Olson and Dix 1991). In the Rogue River, data from an upper-basin fish counting station collected from 1942 to 2009 suggest a major increase in the frequency of fall-run fish accessing historical spring-run habitat after a dam was constructed and a concomitant increase in intermediate migrators (i.e., putative heterozygotes) (ODFW 2000; Thompson et al. 2019). Importantly, in these Rogue River data, the frequency of fall run and intermediate migrators in the Upper Rogue was consistently low across almost 40 years of data before a substantial increase corresponding to the construction of Lost Creek Dam in 1977. In the Chehalis basin, U.S. Fish and Wildlife surveys also noted a loss in the spatiotemporal segregation between spring and fall-run spawning after a dam was built (Hiss et al. 1985), and a substantial proportion of heterozygotes observed in the Chehalis (Thompson presentation) were sampled near this dam. Therefore, it seems reasonable to conclude that, although the degree of demographic interaction between spring and fall fish naturally varies over time and that some degree of interbreeding is normal and expected, human activities have notably increased interbreeding in many locations.

Areas of uncertainty

It is unclear how much demographic isolation from fall run is required for spring Chinook salmon to persist. See the more extensive discussion in Conservation Implications.

• How common are large-effect genes?
• Is it likely that strong associations will be found between specific alleles and many other phenotypic/life-history traits in salmon?

Areas of agreement

Loci of large effect have been identified for other salmonid life-history traits. Single loci of large effect have been found for a number of other traits, but none as consistently across populations and species as GREB1L/ROCK1. For example, Barson et al. (2015) described a locus in Atlantic salmon that explained 39% of the phenotypic variance in age at maturity in Atlantic salmon, but a subsequent large-scale GWAS study (Sinclair-Waters et al. 2020) found a mixed genetic architecture that involves a polygenic component as well as large-effect loci. In *O. mykiss*, a region on chromosome 5, recently identified to have two chromosomal inversions in close proximity, is associated with anadromy and residency in southern portions of the species' range (Pearse et al. 2019a and references therein), but this pattern has not been found in some other geographic areas (K. Nichols presentation, and Weinstein et al. 2019). This same genomic region had been previously identified with strong associations to development rate (explaining up to 30% of the phenotypic variance) (Miller et al. 2012 and references therein), and age at maturity in the species (Haidle et al. 2008). Variation in Y chromosome haplotypes have also been significantly associated with male age at maturation in Chinook salmon (McKinney et al. 2019), but the effect sizes have not been characterized. There are also many studies, from quantitative trait loci (QTL) analyses to genome-wide association studies that identify additional loci of smaller effect, or traits with smaller signals throughout the genome (e.g. Brieuc et al. 2015).

Areas of uncertainty

More data are needed from whole genome sequencing to know the extent to which complex traits are controlled by single genes of large effect, or many loci of smaller effect and how this various among populations. As whole genome sequencing becomes more economically feasible, we may better understand the extent to which genes of large effect have been missed with reduced representation methods, and how traits vary in their genetic architecture.

- **Prince et al. (2017) concluded that the haplotypes associated with early migration timing evolved only once within each species. Is that the case, or are the genetic variants more evolutionarily labile?**

Areas of agreement

The evolutionary history of the GREB1L/ROCK1 region is complex and has not been well characterized throughout each species' entire range. But it is clear that the early and late haplotypes that have been well characterized evolved long ago in each species' evolutionary history. It is also clear, based on available data, that the allelic variants for early migration have not arisen independently via new mutations from the genomic background of late migration individuals in each watershed.

Summary of Future Research Needs

As a result of discussion at the workshop, the participants outlined the following areas for future research:

- Better standardization and characterization of adult migration phenotypes in multiple populations and lineages, including when the 'decision' to migrate is made, how it relates to the timing of sexual maturity and the relationship(s) between the date of freshwater entry and subsequent upstream movements.
- More thorough marker development and validation (see next section). Ideally, identification of the functional variant(s) in the GREB1L/ROCK1 region that cause alternative migration phenotypes.
- Greater understanding of the physiological mechanisms leading to alternative migration phenotypes.
- Tests for association of GREB1L/ROCK1 variation on phenotypes other than adult run timing, such as timing of sexual maturity or other life-history traits.
- More thorough evaluations of the genetics of run-timing variation, throughout the geographic range of Chinook salmon and steelhead, as well as studies in other salmon species in order to develop broad baseline data on the historical and current distribution of alleles at this locus. Current studies have been primarily focused on a limited number of West Coast and Columbia River populations. These investigations should include characterization of the full suite of genetic variants (and their effect sizes) contributing to run timing,

- More thorough characterization of GRE1L/ROCK1 haplotype diversity and the phenotype and dominance pattern of each identified haplotype in multiple populations of both species, across their range.
- Perform comparative analyses on systems with early-run and late-run populations that have been differentially impacted by human activities resulting in differing levels of interbreeding between life-history types, to determine how interbreeding might affect persistence of run type alleles.

Conservation Implications

Areas of agreement

The workshop participants agreed that spring Chinook salmon and summer steelhead occupy a specialized ecological niche—upstream areas accessible primarily during spring flow events—that has made them particularly vulnerable to extirpation or decline due to habitat degradation. The participants also agreed that diversity, including diversity in adult run timing, is important for the long-term viability of many if not all salmon ESUs. This emphasis on diversity is reflected in the recovery plans for listed coastal Chinook ESUs. For example, the Lower Columbia River and Puget Sound Chinook ESUs, as well as the Northern California steelhead DPS, each contain both early and late-run populations, and the recovery plans for these ESU/DPSs requires both run types to be recovered in order for the ESU/DPSs to be considered recovered (Shared Strategy Development Committee 2007; Dornbush 2013; Pearse et al. 2019b). This requirement was reaffirmed by the recent rejection of a petition to separate summer- and winter-run steelhead within the Northern California steelhead DPS[1].

After discussion on whether conservation strategies might need to change based on the GREB1L/ROCK1 findings, the participants generally agreed that using patterns of genetic variation throughout the genome remains important for identifying conservation units, rather than identifying units based solely on small genomic regions associated with specific traits.

The participants generally agreed that the evaluation of risk to early returning population groups (spring Chinook, summer steelhead) needs to consider what we now know about the genetic basis of adult return time. In particular, under the paradigm in which run timing was assumed to be influenced by a great many loci each of small effect, it seemed reasonable to conclude that late run populations (fall Chinook, winter steelhead) would be the optimal source for recolonization of early runs, should these be extirpated. Under the new paradigm in which run timing in some populations appears to be largely influenced by a single genomic region of large effect, the workshop participants concluded that it is now far less clear that local late-run fish will contain the necessary genotypes to restore early run populations, particularly given that surveys from locations where the spring run has either been extirpated (i.e., Scott, Shasta, Iron Gate, Wynoochee) or likely didn't historically exist (Eel River) have found that spring-run alleles are absent or extremely rare.

[1] https://www.fisheries.noaa.gov/action/12-month-finding-petition-list-summer-run-steelhead-northern-california-endangered-under

The participants generally agreed that the finding that the early run trait has a simple genetic basis implies that it is at greater risk of loss than if it were highly polygenic because loss of the 'early' allele(s) equates to the loss of the phenotype. The exception to this conclusion would be if fish containing the early allele could readily migrate from a reservoir elsewhere, either from another population within an ESU or from a different ESU, and successfully reproduce in the new location; in the case of steelhead, reservoirs could possibly exist in previously dammed habitats that restricted migration to and from the ocean. The status of one ESU may therefore depend on the status of other ESUs with which it may exchange occasional migrants. For example, if an 'early' allele at the GREB1L/ROCK1 region is required for expression of spring run timing in coastal Chinook salmon and these alleles are lost from an ESU, the only opportunities to regain those alleles would be by mutation (expected to be exceedingly rare) or immigration of the allele in fish from other ESUs. The reduction in frequency of the spring-run trait in multiple coastal Chinook salmon ESUs may therefore be an indicator of greater risk to spring run in each ESU than would be the case if the allele were common in multiple ESUs. If spring-run populations are at risk in multiple ESUs, this is therefore something that should be considered in evaluating risk.

Areas of uncertainty

One area of uncertainty and potential disagreement at the workshop was the degree to which run-timing diversity in spring Chinook salmon is partitioned among populations versus among individuals within a population. Most status reviews and recovery plans for ESA-listed coastal Chinook salmon (Myers et al. 1998; Shared Strategy Development Committee 2007; Dornbush 2013) consider spring versus fall run timing to be a characteristic of a population, although there is clearly substantial variation among individuals. In contemporary samples from many coastal drainages, however (Klamath, Rogue, Chehalis) heterozygous genotypes at the GREB1L/ROCK1 region are relatively common, and overall genetic differences between spring and fall Chinook salmon are small. These findings represent clear evidence for substantial ongoing or recent interbreeding among spring and fall-run fish in multiple populations. Although not precisely quantified at the workshop, the overall proportions of heterozygotes in at least some portions of the Klamath, Rogue, Sandy and Chehalis Rivers appear to indicate levels of interbreeding more consistent with spring and fall-run fish currently being ecotypic variants that are part of the same demographic population rather than as two different populations.

The extent to which observed contemporary levels of interbreeding between individuals with early and late run timing would be typical under historical environmental conditions is unknown. The dynamic nature of the Pacific Northwest environment and geology makes it reasonable to conclude that the direction and amount of interbreeding between early and late runs has been variable over many timescales. However, there is clear documentation that anthropogenic activities have increased opportunities for interbreeding between ecotypes, at least in some locations. For example, high rates of interbreeding between spring and fall-run fish in the Upper Rogue River appears to be due to changes in water temperature and flow associated with an upstream dam that has allowed fall-run fish to access what was historically spring-run habitat (Thompson et al. 2019). Workshop participants also cited numerous habitat changes in the Klamath and Chehalis

Rivers that likely have increased interbreeding between spring and fall-run Chinook salmon, including modifications to natural low-flow barriers to allow fall-run fish greater access to upstream habitats and/or blockage of upstream habitat (Wendler and Deschamps 1955; Hiss et al. 1985; Olsen and Dix 1991), both of which would be expected to increase relative degree of overlap and thus opportunities for interbreeding between runs. Analysis of recombination events in the Klamath River whole-genome sequencing data indicated that some level of interbreeding between the run types was occurring prior to 200 years ago, but the level of historical interbreeding or the degree to which it has increased has not been quantified (Anderson presentation). However, the type of habitat that creates flow-dependent partial migration barriers is naturally dynamic, so it is reasonable to conclude that the nature and extent of interbreeding has also been variable over space and time.

Understanding the conservation implications of dominance patterns at the GREB1L/ROCK region is also important and is complicated because of tradeoffs between the probability of persistence of the early-run allele and the feasibility of starting new early-run populations. If the early allele(s) is dominant, the early phenotype will be expressed in both the heterozygote and the early-early homozygote, and will be positively selected for in locations that favor expression of the early-run phenotype. If the early allele(s) is recessive, it could still persist in the heterozygote form in environments that are not favorable to the early run phenotype. On the other hand, the relationships might be additive, or partially dominant, in which case the heterozygote phenotype would be intermediate to the homozygote phenotypes.

The dominance-recessive relationships might influence the success of colonization events. In cases where the late allele is dominant, heterozygotes would have the same run timing as late-late homozygotes. A single generation of mating would produce early-early homozygote offspring, based on the frequency of the allele p in the population. If p is small, which is likely in the case of recolonization, and the late allele is dominant, it would be difficult for the early-run phenotype to become successfully established because early-early homozygotes would be extremely rare. The colonization scenario would be more likely to succeed under an additive mode of inheritance (heterozygotes have intermediate run timing) or with dominance of the early allele, because only one early allele would be needed to create an early phenotype. Current data (discussed in the previous section) tends to support the idea that the early allele(s) are dominant or co-dominant, suggesting that conserving existing and restoring lost habitats and environmental conditions in which the early run phenotype is favored will be very important for conservation of the early-run alleles and phenotypes.

Regardless to what extent current levels of interbreeding are a consequence of human mediated habitat alterations, such interbreeding and the common occurrence of heterozygotes at the GREB1L/ROCK1 region presents challenges for status monitoring, recovery planning, and other management actions. For status monitoring, if spring-run fish are currently not demographically independent from fall-run fish, then the standard population modeling approaches for assessing risk are not appropriate. Furthermore, it might not always be clear whether some interbreeding with fall-run fish helps to maintain the viability of the spring-run phenotype, or whether fall-run fish are a competitive threat that is displacing the spring run from areas they historically

occupied – or whether both factors are operating. From a broader perspective, the declining frequency of spring run (and increasing frequency of fall run) individuals could be seen as either genetic change and biodiversity loss that could be detrimental to the species' future evolution (e.g., it could result in a lack of standing genetic variation for responding to increasing summer temperature) or the natural consequence of a species adapting to its current, anthropogenically-modified environmental conditions. Despite the complexity of these issues, the workshop participants concluded that maintaining run-timing diversity within ESUs is important for the viability and long-term conservation of the species.

Improved strategies are needed for monitoring run timing and associated genetic variation. The workshop participants discussed several strategies for improved monitoring as well as potential types of conservation measures to increase spring-run abundance. Based on presentations of genotype surveys in several watersheds (Klamath, Rogue, Chehalis) containing both spring and fall-run Chinook salmon, monitoring trends in GREB1L/ROCK1 genotypes might provide a more accurate measure of spring run relative abundance than phenotypic monitoring. This is particularly the case if the phenotypic monitoring is conducted well after the presumed time of freshwater entry, such as is typical with spawning ground surveys. One potential idea that was discussed was monitoring trends of 'early' allele and genotype frequencies in a population-genetic context, such that fish with homozygous 'early' genotypes would receive double the weight of fish with heterozygous genotypes in abundance surveys. This approach could be particularly useful in areas where spring and fall-run fish are commonly interbreeding.

What conservation measures can be put into place now with existing knowledge? Conservation measures for spring run that were discussed included potentially shaping fisheries to focus disproportionately on fish with fall run timing, restoring access to spring-run habitat that has been blocked, considering restoring natural barriers that have been modified to increase fall-run access to historically spring-run habitats, and restoring more natural flow regimes (e.g., low summer flows that prevent mature migrating individuals from encroaching on premature habitat). Workshop participants agreed that the presence of heterozygotes does not in itself indicate a threat to the viability of spring run as these heterozygotes contain alleles that may be important to spring-run restoration. Some workshop participants also noted, however, that in some cases the presence of high proportions of heterozygotes might represent a departure from the historical conditions and a warning sign that the spring-run phenotype is at risk.

Issues specifically associated with steelhead

Most of the issues discussed above with reference to spring and fall-run Chinook salmon also apply to summer and winter run steelhead, though in steelhead there is even more need to better characterize the relationship between variation at GREB1L/ROCK1 and run timing throughout the range of the species. However, there are some important differences and additional uncertainties associated with the genetics and conservation of run-timing variation in steelhead compared to Chinook.

One major factor to consider regarding the conservation implications of the genetics of run-timing diversity in steelhead is the existence of conspecific resident rainbow trout populations that may effectively act as reservoirs for the 'early' GREB1L/ROCK1 alleles. Such variation appears likely to have been responsible for the rapid reappearance of summer steelhead in the Upper Elwha River following dam removal, for example (Nichols presentation). 'Early' GREB1L alleles have also been observed at high frequencies in resident populations in other coastal streams (Pearse et al. 2019b), including those that are completely fixed for the 'rearranged' (in some studies associated with the 'resident' life history) haplotype of chromosome Omy05 (Pearse et al. 2014). The ability of this allele to persist in resident life-history forms is therefore likely to reduce the overall risk of loss of GREB1L diversity in areas where summer steelhead have been extirpated or are in decline but resident forms persist above a barrier (i.e., are protected from ongoing interbreeding with winter steelhead). The participants noted, however, that it is unclear what functional significance variation at GREB1L has for resident *O. mykiss*, which don't migrate from the ocean but might perform seasonal migrations within freshwater watersheds.

Another factor to consider for steelhead compared to Chinook is the generally greater amount of life-history diversity found in *O. mykiss*. In addition to the resident life history, the species also exhibits a broader, but poorly characterized, run-timing distribution in some locations. For example, fall-run steelhead on the Rogue and Klamath make up a substantial portion of the return, but are generally grouped with either the summer or winter runs (Busby et al. 1994). In addition, novel life-history patterns exist, such as the 'half pounder' pattern found in the Klamath and Rogue Rivers and other streams in Northern California and Southern Oregon. Along with greater life-history diversity patterns of variation within the GREB1L/ROCK1 region also appear to be more complicated in steelhead than in Chinook, although the functional significance, if any, of most of this variation remains unknown. Conducting more extensive surveys of both phenotypic and genotypic variation in *O. mykiss* was identified by the participants as a high priority for future research.

References

Arciniega M, Clemento A, Miller MR, Peterson M, Garza JC, Pearse DE. (2016) Parallel evolution of the summer steelhead ecotype in multiple populations from Oregon and northern California. Conservation Genetics 17, 165–175.

Ashcraft S, Holt C, Scharpf M, Zimmerman M, VanBuskirk N. (2017) Spawner Abundance and Distribution of Salmon and Steelhead in the Upper Chehalis River, 2013–2017 (No. FPT 17-12). Retrieved from Washington Department of Fish and Wildlife website: https://wdfw.wa.gov/publications/01970/

Ayllon F, Kjaerner-Semb E, Furmanek T, et al. (2015) The vgll3 Locus Controls Age at Maturity in Wild and Domesticated Atlantic Salmon (*Salmo salar* L.) Males. PLoS Genetics 11, DOI: 10.1371/journal.pgen.1005628

Barson NJ, Aykanat T, Hindar K, et al. (2015) Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528, 405–408.

Brieuc MSO, Ono K, Drinan DO, Naish KA. (2015) Integration of random forest with population-based outlier analyses provides insight on the genomic basis and evolution of run timing in Chinook salmon (*Oncorhynchus tshawytscha*). Molecular Ecology 24, 2729–2746.

Busby PJ, Wainwright TC, Waples RS. (1994) Status review for Klamath Mountains Province steelhead. U.S. Dep. Commerce, NOAA Technical Memorandum NMFS-NWFSC-19, 130p.

Busby PJ, Wainwright TC, Bryant GJ, et al. (1996) Status Review of West Coast Steelhead from Washington, Idaho, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-27.

Campbell, L., Claiborne, A., Ashcraft, S., Zimmerman, M., & Holt, C. (2017). Investigating Juvenile Life History and Maternal Run Timing of Chehalis River Spring and Fall Chinook Salmon Using Otolith Chemistry. WDFW.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. (2000) Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15, 290–295.

Dornbush P. (2013) ESA Recovery Plan for: Lower Columbia River coho salmon, Lower Columbia River Chinook salmon, Lower Columbia River Chum Salmon, and Lower Columbia River steelhead. Prepared by the National Marine Fisheries Service. 503 p.

Falconer DS, Mackay TFC. (1996) Introduction to Quantitative Genetics, Fourth Edition Longman Limited, Essex.

Gustafson RG, Waples RS, Myers JM, et al. (2007) Pacific salmon extinctions: Quantifying lost and remaining diversity. Conservation Biology 21, 1009–1020.

Haidle L, Janssen JE, Gharbi K, et al. (2008) Determination of quantitative trait loci (QTL) for early maturation in rainbow trout (*Oncorhynchus mykiss*). Marine biotechnology 10, 579–592.

Healey MC. (1983) Coastwide distribution and ocean migration patterns of stream- and ocean-type Chinook salmon, *Oncorhynchus tshawytscha*. Can. Field-Nat. 97, 427–433.

Hecht BC, Matala AP, Hess JE Narum SR. (2015), Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. Molecular Ecology, 24, 5573–5595

Hess JE, Zendt JS, Matala AR, Narum SR. (2016) Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing. Proceedings of the Royal Society B—Biological Sciences 283, 20153064.

Hess JE, Campbell NR, Matala AP, Hasselman DJ, Narum SR. (2015) Genetic assessment of Columbia River stocks, 4/1/2014–3/31/2015 Annual Report, 2008-907-00, https://www.critfc.org/wp-content/uploads/2016/04/16-03.pdf.

Hiss JM, Meyer JH, Boomer RS. (1985) Chehalis Spring Chinook Progress Report 1984. US Fish and Wildlife Service, Olympia, WA

Langin K. (2018) A different animal. Science 360, 590–592.

Lee KN, Lawrence J. (1986) Adaptive management: learning from the Columbia River Basin fish and wildlife program. Environmental Law (Portland) 16, 431–460.

Lichatowich J. (1999) Salmon without rivers. A history of the Pacific salmon crisis Island Press, Washington, D.C.

Mackay TFC, Stone EA, Ayroles JF. (2009) The genetics of quantitative traits: challenges and prospects. Nature Reviews Genetics 10, 565–577.

McClelland EK, Naish KA. (2010) Quantitative trait locus analysis of hatch timing, weight, length and growth rate in coho salmon, *Oncorhynchus kisutch*. Heredity 105, 562–573.

McElhany P, Ruckelshaus MH, Ford MJ, Wainwright TC, Bjorkstedt EP. (2000) Viable salmonid populations and the recovery of evolutionarily significant units. US Dept. Commer. NOAA Tech. Memo. NMFS-NWFSC, 42, p. 156.

McKinney GJ, Seeb JE, Pascal CE, et al. (2019) Y-chromosome haplotypes drive variation in size and age at maturity in male Chinook salmon. bioRxiv, 691063.

Micheletti SJ, Hess JE, Zendt JS, Narum SR. (2018) Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous steelhead. BMC Evolutionary Biology 18, DOI: 10.1186/s12862-018-1255-5.

Miller MR, Brunelli JP, Wheeler PA, et al. (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. Molecular Ecology 21, 237–249.

Moran, P., D.J. Teel, M.A. Banks, T.D. Beacham, M.R. Bellinger, S.M. Blankenship, J.R. Candy, J.C. Garza, J.E. Hess, S.R. Narum, L.W. Seeb, W.D. Templin, C.G. Wallace, and C.T. Smith. (2013) Divergent life-history races do not represent Chinook salmon coast wide: the importance of scale in Quaternary biogeography. Canadian Journal of Fisheries and Aquatic Sciences 70, 415–435.

Moritz C. (1994) Defining 'evolutionarily significant units' for conservation. Trends in Ecology and Evolution 9, 373–375.

Myers JM, and 10 others. (1998) Status review of chinook salmon from Washington, Idaho, Oregon and California. NOAA Technical Memorandum NMFS-NWFSC-35.

Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, DeKoning J, Elz A, Guthrie CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF, Garza JC. (2008) Differentiating salmon populations at broad and fine geographic scales with microsatellites and SNPs. Molecular Ecology 17, 3464–3477.

Narum SR, Di Genova A, Micheletti SJ, Maass A. (2018) Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon. Proceedings of the Royal Society B-Biological Sciences 285, DOI: 10.1098/rspb.2018.0935.

Nehlsen W, Williams JE, Lichatowich JA. (1991) Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries 16, 4–21.

NMFS. (1990) Listing endangered and threatened species: Petitions to list Snake River spring, summer, and fall chinook salmon, and Lower Columbia River coho salmon. Federal Register 55, 3742.

NMFS. (1991) Notice of policy: Policy on applying the definition of species under the Endangered Species Act to Pacific salmon. Federal Register 56, 58612–58618.

NMFS. (1995) Endangered and threatened species: Status of Sacramento River winter-run chinook salmon. Federal Register 59, 440–450.

NMFS. (2018) 90-day petition finding, request for information and initiation of status review. Federal Register 83, 8410–8414.

ODFW. (2000) Effects of Lost Creek Dam on spring Chinook salmon in the Rogue River. Phase II completion report (Oregon Dept of Fish and Wildlife, Salem, OR).

Olson AD, Dix OJ. (1991) Lower Salmon River Sub-basin Fish Habitat Condition and Utilization Assessment 1990/1991 Final Report. USDA – Forest Service, Klamath National Forest.

Pearse DE, Miller MR, Abadia-Cardoso A, Garza JC. (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. Proc Roy Soc B 281, 20140012.

Pearse DE, Barson NJ, Nome T, et al. (2019a) Sex-dependent dominance maintains migration supergene in rainbow trout. Nature Ecology & Evolution 3, DOI: 10.1038/s41559-019-1044-6.

Pearse DE, Garza JC, Myers J, Spence B. (2019b) Northern California steelhead DPS-Configuration Review-Panel Report. Available: https://www.fisheries.noaa.gov/action/12-month-finding-petition-list-summer-run-steelhead-northern-california-endangered-under.

Prince DJ, O'Rourke SM, Thompson TQ, et al. (2017) The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. Science Advances 3, e1603198.

Quinn TP, McGinnity P, Reed TE. (2016) The paradox of "premature migration" by adult anadromous salmonid fishes: patterns and hypotheses. Canadian Journal of Fisheries and Aquatic Sciences 73, 1015–1030.

Shared Strategy Development Committee. (2007) Puget Sound salmon recovery plan. Volume 1. Plan adopted by the National Marine Fisheries Service (NMFS) January 19, 2007. Available at http://www.westcoast.fisheries.noaa.gov/protected_species/salmon_steelhead/recovery_planning_and_implementation/puget_sound/puget_sound_chinook_recovery_plan.html.

Sinclair-Waters M, Ødegård J, Korsvoll SA, Moen T, Lien S, Primmer CR, Barson NJ. (2020) Beyond large-effect loci: large-scale GWAS reveals a mixed large-effect and polygenic architecture for age at maturity of Atlantic salmon. Genetics Selection Evolution, 52, 1–11.

Thompson TQ, Bellinger MR, O'Rourke SM, et al. (2019) Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. Proceedings of the National Academy of Sciences of the United States of America 116, 177–186.

Thompson TQ, O'Rourke SM, Brown SK, Seamons T, Zimmerman M, Miller MR. (2019) Run-type genetic markers and genomic data provide insight for monitoring spring-run Chinook Salmon in the Chehalis Basin, WDFW contract #18-11697, Final report submitted to Washington Department of Fish and Wildlife, 26 pp.

USFWS, NMFS. (1996) Policy regarding the recognition of distinct vertebrate population segments under the Endangered Species Act. Federal Register 61, 4722.

Veale AJ, Russello MA. (2017) Genomic Changes Associated with Reproductive and Migratory Ecotypes in Sockeye Salmon (*Oncorhynchus nerka*). Genome Biology and Evolution 9, 2921–2939.

Waples RS. (1991) Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review 53, 11-22.

Waples RS. (2006) Distinct Population Segments. In: The Endangered Species Act at Thirty: Conserving Biodiversity in Human-Dominated Landscapes. pp. 127–149. Island Press, Washington, D.C.

Waples RS, Lindley ST. (2018) Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids. Evolutionary Applications 11, 1518–1526.

Waples RS, Naish KA, Primmer CR. (2020) Conservation and Management of Salmon in the Age of Genomics. Annual Review of Animal Biosciences 8, 117–143.

Waples RS, Teel DJ, Myers JM, Marshall AR. (2004) Life-history divergence in Chinook salmon: historic contingency and parallel evolution. Evolution 58, 386–403.

Weinstein SY, Thrower FP, Nichols KM, Hale MC. (2019) A large-scale chromosomal inversion is not associated with life history development in rainbow trout from Southeast Alaska. Plos One https://doi.org/10.1371/journal.pone.0223018.

Wendler, HO, Deschamps G. (1955) Logging dams on coastal Washington streams. Fish. Res. Pap. 1: 27-38, Washington Department of Fisheries, Olympia, WA.

Wood AR, Esko T, Yang J, et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nature Genetics 46, 1173–1186.

Glossary

Sequencing Methods

Amplicon Sequencing (AmpliconSeq)
General term for sequencing of PCR products.

Genotyping-in-thousands by sequencing (GTseq)
Variant of amplicon sequencing that uses dual-index barcodes to allow sequencing of hundreds to thousands of individuals and for hundreds of loci at one time.

Next-generation sequencing (NGS)
General term describing rapid and high-volume DNA sequencing technologies that superseded traditional DNA sequencing.

PoolSeq
A cost-saving approach whereby pools of DNA from multiple individual are sequenced to estimate allele frequencies of the pool rather than obtaining individual genotypes. Typically used as a variant of *whole-genome sequencing*.

RAD-capture (RAPTURE)
Method of sequencing subsets of RAD-seq loci using capture baits to target only the desired loci. This method allows sequencing of hundreds of samples and thousands to tens of thousands of loci at one time but is limited to RAD-seq derived loci.

Read Depth Coverage
The number of overlapping sequence reads at a particular site in the genome. For example 10x coverage for a SNP means ten sequence reads overlap. Higher coverage yields greater confidence in genotype calls but there is a tradeoff that higher depth per individual means fewer individuals can be sequenced per lane.

Reduced-representation sequencing (RRS)
General term describing methods that sequence subsets of the genome. Commonly used to economically genotype thousands to tens of thousands of SNPs.

Restriction-site associated DNA sequencing (RAD-seq)
Reduced-representation sequencing method that sequences DNA adjacent to restriction sites. This method has many variants that differ in details of sample preparation and in number of loci generated. The number of loci generated typically range from thousands to tens of thousands. Variants include the original RAD-seq method (RAD or traditional RAD), 2bRAD, GBS, SBG, CRoPS, RRL, MSG, ezRAD, and ddRAD. See Andrews et al. 2016 for overview of different methods.

SNP array
Analysis platform used to assess SNP loci genotypes in a high-throughput manner (even millions of loci).

Whole-genome sequencing (WGS)
General term for methods that sequencing the entire genome of an organism, as opposed to methods that sequence subsets of the genome (see *Reduced-representation sequencing*).

Genetic Variants

Allozymes (allo enzymes)
Variant sites at protein-coding loci that are detected with protein electrophoresis based on differences in amino acid sequence.

Genetic architecture
The underlying genetic basis of a phenotypic trait; the number and effect sizes of genes, their interactions within and between each other, and their inheritance patterns.

Genome-wide association study (GWAS)
Study of genetic variation spanning the genome to detect variants associated with specific phenotypic traits.

Haplotype
A set of DNA variants that are inherited together. A haplotype can refer to a combination of alleles or to a set of *SNPs* found on the same chromosome.

Homeolog
A special case of a *paralog* arising through genome duplication. In salmon genomics this commonly refers to chromosome pairs that still show signals of retained tetraploidy following the salmonid whole-genome duplication.

Inversion
A chromosome rearrangement in which a segment of a chromosome is reversed end to end. Inversions inhibit recombination, leading to divergence between inverted and ancestral chromosome types.

Microhaplotype
DNA sequence variation comprised of two or more close *SNPs*. For *NGS* data this often means within a single sequence read, commonly 100 or 150 base pairs for Illumina sequencers.

Microsatellites
Noncoding regions of DNA that contain variable numbers of short (usually 2–4 base pairs), repeated DNA sequences.

Paralog
DNA sequence copies created by a duplication event within the same genome. This is a general term and may refer to copies that arose through tandem duplication of sequence or whole-genome duplication.

Single-nucleotide polymorphisms (SNPs)
Single DNA base pairs that are variable within the target population; most SNPs only have two variant alleles.

GREB1L Terms

Region of strongest association (RoSA)
The GREB1L/ROCK1 region showing the statistically strongest association with run timing (a term coined by Anderson et al. (pers. comm.).

Appendix

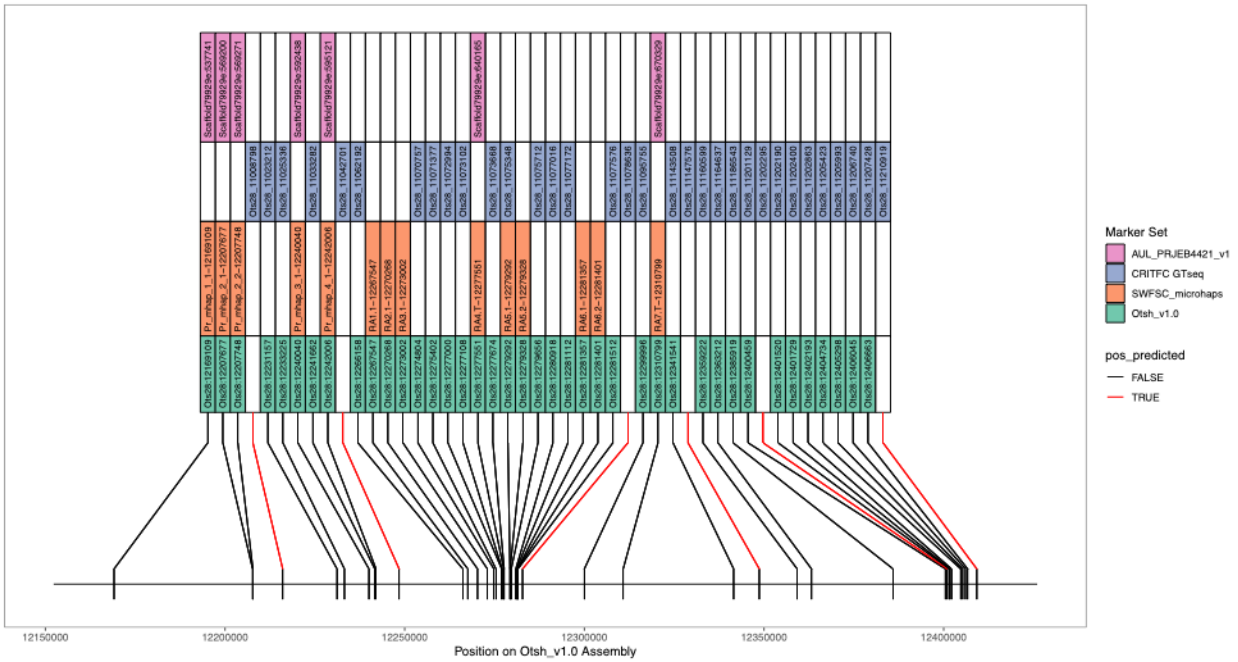

Figure A-1. Markers developed across studies and laboratories for the GREB1L/ROCK1 region on chromosome 28 in Chinook salmon. Marker sets are either the Chinook genome coordinates (Otsh_v1.0) or panels developed by individual laboratories (CRITFC = Columbia River Inter-Tribal Fish Commission, SWFSC = Southwest Fisheries Science Center, AUL = *O. mykiss* scaffold-aligned RADseq markers). Figure courtesy E. Anderson.

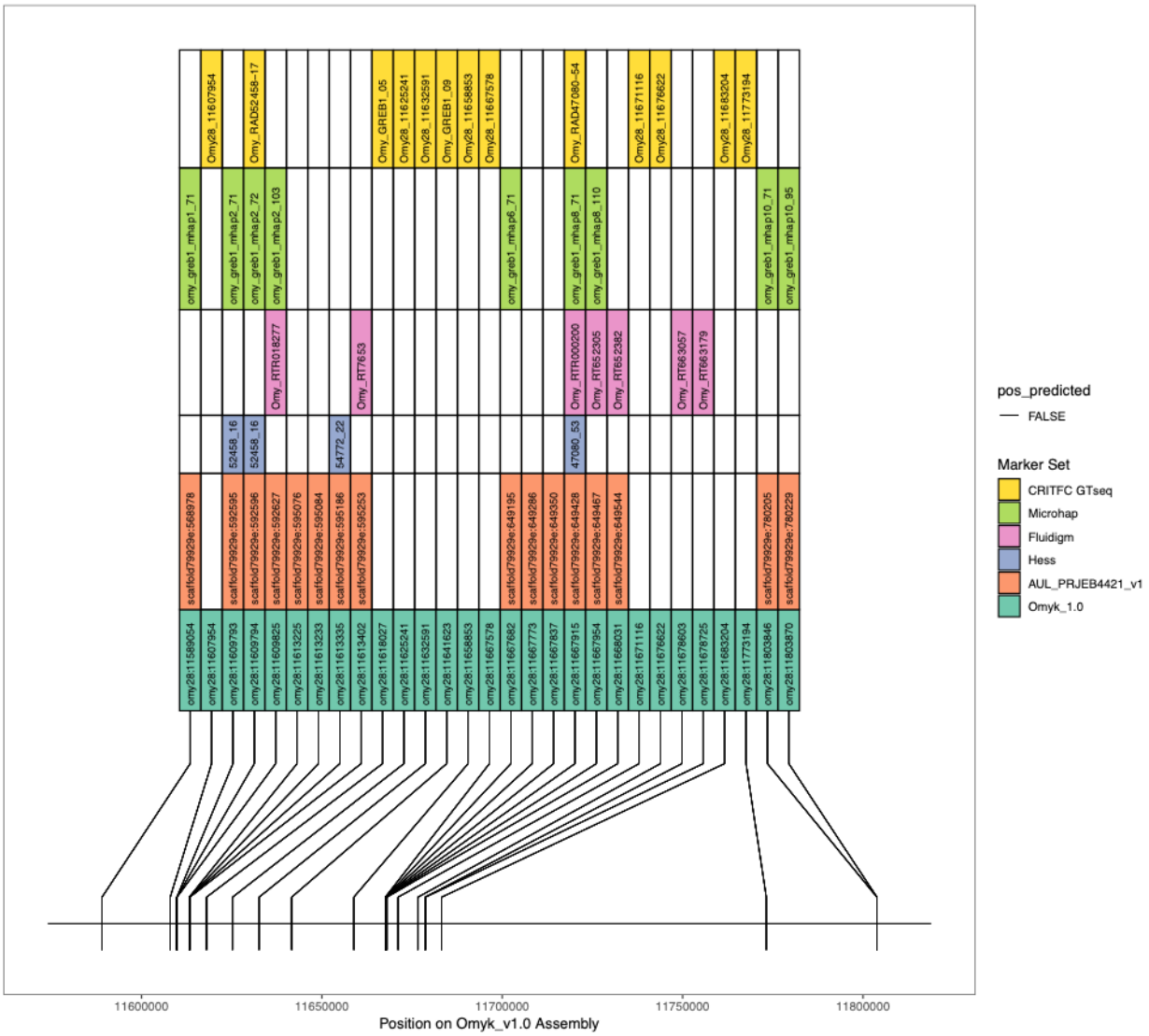

Figure A-2. Markers developed across studies and laboratories for the GREB1L/ROCK1 region on chromosome 28 in *Oncorhynchus mykiss*. Marker sets are either the *O. mykiss* genome coordinates (Omyk_1.0) or panels developed by individual laboratories or publications (CRITFC and Hess = Columbia River Inter-Tribal Fish Commission, Microhap and Fluidigm = Southwest Fisheries Science Center, AUL = *O. mykiss* scaffold-aligned RADseq markers). Figure courtesy E. Anderson.

U.S. Secretary of Commerce
Wilbur L. Ross, Jr.

Acting Under Secretary of Commerce
for Oceans and Atmosphere
Dr. Neil Jacobs

Assistant Administrator for Fisheries
Chris Oliver

June 2020

fisheries.noaa.gov

OFFICIAL BUSINESS

National Marine
Fisheries Service
Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle, Washington 98112