Independent Scientific Review Panel
for the Northwest Power & Conservation Council
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204
www.nwcouncil.org/fw/isrp

Memorandum (ISRP 2010-21) **June 18, 2010**

To: Bruce Measure, Chair, Northwest Power and Conservation Council

From: Eric Loudenslager, ISRP Chair

Subject: Follow-up Review of CRITFC Fish Accord Proposal, Influence of Environment and Landscape on Salmonid Genetics, 2009-005-00

Background

This is a follow-up review to an earlier ISRP and Council review of the Columbia River Inter-Tribal Fish Commission's Accord Proposal: Influence of Environment and Landscape on Salmonid Genetics, 2009-005-00. The proposal has two basic objectives: 1) Environment and Landscape Genetics – Evaluate genetic structure of natural populations of salmonids relative to their environment and identify candidate markers associated with traits that are related to adaptation of steelhead and Chinook salmon populations (i.e., smoltification and thermal tolerance); and 2) Controlled Experiments – laboratory/hatchery experiments with controlled environmental variables to validate phenotypic response of fish with given genotypes. The information gained from this proposal is intended to facilitate understanding of adaptation of natural salmonid populations to their environment. CRITFC believes this information should benefit future management of natural, supplemented, and reintroduced populations.

The proposal was originally submitted to the ISRP for review in November 2008, and on December, 12, 2008, we requested additional information before we could determine if the proposal met scientific criteria. On January 28, 2009 the ISRP and the Council received a response from the CRITFC, and on February 19, 2009, we provided a review of the proposal and response (ISRP 2009-3). We found that the proposal "Does Not Meet Scientific Review Criteria" because it lacked adequate detail to meet certain review standards. We recommended that if this project proceeds, a detailed study design should be prepared and reviewed. We laid out five points for each of the major project objectives that needed to be addressed in a revised proposal/study plan.

On March 11, 2009, the Council recommended that CRITFC continue to design the project for implementation, conditioned on the understanding that the implementation of this project be dependent on a review by a final ISRP and Council review. In addition, the Council anticipated that the project proponent will participate in development of a regional approach to monitoring, evaluation, research and reporting strategies and that some changes to the scope and intent of this project may be adjusted when the regional strategy is in place.

On May 11, 2010, the Council forwarded CRITFC's revised proposal that was revamped to provide the details requested by the ISRP and the Council. Our review below is organized by the two major objectives and our previous concerns.

ISRP Recommendation

Does Not Meet Scientific Review Criteria

This revised proposal is an improvement on the original but still does not have the level of detail essential for technical review (see comments below). Additionally, the revision did not provide a point-by-point response to the individual issues raised by the ISRP in the February 19, 2009 review. If a subsequent proposal is developed, in addition to a revised narrative, the ISRP requests that the proponents provide a document that succinctly responds to the individual points raised by the ISRP.

The ISRP believes that a well-crafted investigation could lead to an increased understanding of the genetic and environmental causation of the anadromy dichotomy in O. mykiss. The project may very well separate resident from migrant genotypes, even within a single interbreeding population, and distinguish between non-interbreeding resident and migrant ecotypes in the same watershed. The project could improve our understanding of whether geographic variation is (at least) triggered by differences in temporal-thermal profiles. At the very least, the project could have an impact on our understanding of life-history variation and evolution of these traits.

The ISRP encourages the proponent to consider further developing the investigations outlined in this proposal.

Additional recent relevant technical literature, not cited in the proposal, is provided.

ISRP Review Summary

The landscape genetic objective

In February 2009, the ISRP recommended that the study design needs to include five items:

1. The specific hypotheses for the focal populations the analysis is intended to address

The revised proposal still does not include clear hypotheses. There is a general objective to test the concordance of steelhead and Chinook salmon phylogenetic structure with landscape features, and then use recently developed analytical methodologies to evaluate whether marker genes may be associated with directional or balancing selection. But there are no hypotheses or linkage to management plans or decisions for specific independent populations, MPGs, or ESUs.

A lack of hypotheses associated with specific populations is especially problematic, with the attempt to identify SNPs reflecting adaptation to environmental factors. The basic approach to accomplish this objective appears to be a simple comparison of SNPs among populations of

steelhead and Chinook exposed to varying thermal conditions; differences in SNPs are assumed to represent a response to thermal conditions. However, these differences in SNPs may not be caused by selective pressures caused by temperature. SNPs could well be a response to some other environmental variable, for which data are not available. The laboratory work with steelhead should help to identify genetic markers that are related to thermal tolerance. Nevertheless, it is unclear how the analysis of the archived tissue samples will contribute to advancing our understanding on this topic.

2. The field locations where genotypic data will be taken

Some information on the locations from which the archived steelhead tissue samples were collected is provided. However, this information is not provided for the Chinook samples. The proposal simply states "Chinook salmon from thermally adapted populations will also be evaluated in reference to others in the Columbia River." Where are these populations located?

Moreover, how do you know that a given population is "thermally adapted"? There is ample evidence that fish in warm reaches behaviorally thermoregulate by seeking out areas where cool water collect. The availability of these refuges varies. Therefore, it is possible that salmon or steelhead residing is an area where average water temperatures are warm may not be exposed to a significant selective force from temperature. This site variability could complicate the interpretation of the SNPs correlation analysis.

3. The sources and type of genotypic data for each site

A critical deficiency is an absence of a discussion of whether the proponents are using SNPs as markers to locate QTLs for traits of interest (in this case smolting in steelhead and thermal tolerance in steelhead and Chinook) or whether they are actually searching for SNPs that are responsible for the phenotypic traits. SNPs are found in non-coding and coding portions of DNA. When they are located in coding segments, they sometimes alter a gene or its expression; other times they do not. SNPs are used in marker-assisted plant and animal breeding, and are associated with several human diseases.

The proponents plan to screen 100 SNP loci in steelhead and Chinook salmon. For either purpose – to evaluate QTLs or to identify specific loci putatively responsible for phenotypic variation in thermal tolerance or anadromy – the proponents have not established that 100 genes are sufficient. The identification of QTLs for smolt traits in steelhead used 260 or so markers, selected to provide coverage of the 29 linkage groups in O. mykiss (Nichols et al. 2008). Furthermore, there is no explanation of how those hundred were chosen. They could be "targets of availability" or tightly targeted to a set of interesting genes. One either needs a very large panoply that covers the genome or a targeted set that identifies credible candidate loci. The proponent needs to explain to the ISRP, which is it here? And why those and how were they chosen?

There is no explanation in the proposal of how the investigators will determine whether a SNP associated with a trait of interest is a marker for a QTL or actually part of the gene responsible for the trait. If the SNP is a marker for a QTL, then the follow-up evaluation of estimating the

SNP allele frequency across the basin does not make sense. The linkage disequilibrium of a SNP and a QTL for a trait, which provides the empirical statistical association, are unlikely to remain in across different locations. It will likely depend on how close they are. If really close, the disequilibrium could be fairly persistent. Further, QTLs may vary in their expression from one location to another. Considerable foundation investigations will be needed to justify using SNP allele frequencies across the basin to model an adaptation framework.

Most SNPs are synonymous and do not change the protein, even when found in coding sequence, though they may change the regulation (which could be an important distinction for the traits of interest). Determining whether the SNP is resident within the gene of interest, whether it is in a coding sequence, whether it changes the amino acid composition and/or regulation of the protein, or whether it is simply a hitchhiker in the near vicinity of the locus of real interest, is a non-trivial post-discovery challenge. The real payoff comes, if and when they can translate a SNP marker into physiological/developmental *understanding*. There is inadequate description of that follow-up process, and the ISRP needs at least some sense of how that is to be pursued.

In the abstract the investigators state that the first part (landscape genetics) will be conducted in stages: first, mining extant genetic data; second, genotyping new genetic markers (SNPs); third, screening archived tissue samples with these new markers; then analyzing all these markers for concordance or discordance with environmental and geological characters across the landscape. The experimental design on page 8 does not provide details on mining extant data or screening archived samples. Also, the abstract states that putative candidate markers identified in the landscape analysis will be further explored in the expression of traits objective, as a validation under controlled circumstances. Yet, all the controlled experiments are limited to *O. mykiss*. None are described using Chinook salmon.

It would probably be legitimate to confine attention to *O. mykiss* for this proposal.

4. The type and location of environmental data

No information is provided on the location or source of environmental data that will be used in the analysis. It is indicated that precipitation and air temperature estimates can be generated using the Oregon Weather Service PRISM model. But converting air temperature to water temperature is a complicated and uncertain undertaking. Actual water temperature data may be available for many of the sites where the archived tissue samples were collected. The availability of good environmental data should be a key factor in selecting the tissue samples to analyze.

5. The specific correlative analyses that will be performed on the data with an explanation of how the analysis of genetic data with environmental data from those sites resolves the questions posed in the hypothesis

A broad description of tests between genetic data and environmental data is provided, along with a discussion of tests that will evaluate whether any loci are putatively under stabilizing or disruptive selection. There is no explanation, however, of how these analyses will be used to test hypotheses regarding specific populations or how inference from the analysis will inform management decisions, even though the nature of the results is unpredictable at this point.

The expression of traits objective

In February 2009, the ISRP recommended that study design needs to explain the specific methods that will be employed, specifically by addressing the five items below.

1. The breeding design that is going to be used to identify QTLs

The general strategy of using presumptively genetically homogenous populations of resident and anadromous and warm and cool adapted *O. mykiss* is appropriate. However, the explanation of the experimental protocols are not detailed enough to conclude they are sufficient. For example, paragraph 4 on page 12 states: "Specific candidate markers that have shown statistical evidence for selection in Objective 1 (landscape genetics objective) will be chosen for further evaluation in gene expression studies." What is missing is an explanation of how the test fish used in the controlled experiments will be produced to actually carry the alleles of interest discovered in objective 1. Further, in the same paragraph the proposal states: "For example, if glutamate dehydrogenase is indicated to be under differential directional selection among desert and montane populations of *O. mykiss*, the expression of this gene will be tested under controlled temperatures in a wet lab setting." Nowhere in the experimental design for objective 1 was there mention of any specific genes associated with the SNPs. There was no discussion of screening for alleles (SNPs) in any functional genes. This information should be included in the proposal. It is not clear how RNA expression for SNP regions will be accomplished.

2. How alleles at QTLs will be identified

There is no discussion of whether the investigation will evaluate QTLs for thermal tolerance and smolt traits or SNPs that directly influence these traits.

3. How the frequencies of these alleles will be estimated in natural populations

A description of how SNP alleles are to be estimated is included, but that does not suffice to establish that the SNP alleles are representing phenotypic expression and presumptive adaptation at locations throughout the basin. SNPs that are used as markers for QTLs, established from pedigree studies in a laboratory (hatchery), can be used to estimate SNP allele frequencies in natural populations, but the geographic structure of the QTLs has to be separately specified. Whether the two patterns are coordinate is, of course, the issue of greater concern, and that can only be assessed empirically. An association between a QTL and a SNP in a family pedigree does not automatically translate into a population association across the Columbia River Basin, though it certainly does provide us with a legitimate target for sampling and evaluation.

4. How populations that are going to be screened for QTL allele frequencies and expression of HSP and Na/K ATPase will be selected

A portion of the investigation will employ hatchery stocks maintained at Washington State University with known QTLs for smolt traits. Thermal tolerance evaluation will require collecting gametes and rearing fish from locations with specific thermal regimes. While basic

information is provided, the proposal still lacks detail that permits adequate review. The discussion of screening field locations for SNP and QTL frequencies is inadequate.

5. How many populations will be screened?

The screening will be focused on experimental stocks with presumed genotypic and phenotypic traits, rather than populations. The general breeding design is described. It is not evident that the fish in the test populations from Washington State University carry the SNPs that will be identified in the Landscape portion of this proposal. SNPs are normally identified, in the first instance, as highly polymorphic segregating alleles in more than one starting stock or population, and while persistence of that polymorphism over geographic space is certainly not guaranteed, experience seems to show that enough of them will persist geographically that one could be confident that starting with 100 polymorphic SNPS would provide useful geographic signatures for a substantial number of them. Whether the same SNPs will "light up" any particular QTL in other places remains an open question.

Additional Literature

Giger, T., and 7 co-authors. 2008. Population transcriptomics of life-history variation in the genus Salmo. Molecular Ecology 17:3095-3108.

Santure, A. W., and 5 co-authors. 2010. On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. Molecular Ecology 19:xxx-xxx (in press).