Accord Project Sponsors ISRP Response Report

Date: 7/15/2010

Project Number	2009-005-00
Proposer	Columbia River Inter-Tribal Fish Commission
Project Title & Brief Description	Influence of Environment and Landscape on Salmonid Genetics
Contact Name	Shawn Narum
Contact email	nars@critfc.org

ISRP Review History:

Original Narrative submission date: December, 12, 2008

The proposal was originally submitted to the ISRP for review in November 2008, and on December, 12, 2008, ISRP requested additional information before they 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, ISRP provided a review of the proposal and response (ISRP 2009-3). ISRP found that the proposal "Does Not Meet Scientific Review Criteria" because it lacked adequate detail to meet certain review standards. ISRP recommended that if this project proceeds, a detailed study design should be prepared and reviewed. ISRP 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. ISRP review comments below are organized by the two major objectives and ISRP previous concerns.

Date ISRP Review comments were received: Most recent comments June 18, 2010.

ISRP Review results: *[Check appropriate box]*

☐ Meets scientific criteria.

☐ Meets scientific criteria (qualified).

☐ Response requested - meets scientific criteria (qualified).

X Response requested – does not meet scientific criteria.

Response to ISRP Summary: [*Please check appropriate box and respond below in: Response to ISRP Comments]*

☐ The narrative will be revised and resubmitted by (insert target date).

X A response to ISRP comments is provided in this document.

[*Your response should include 1) areas of agreement with ISRP comments, i.e., additional information and/or any changes in the project scope of work and, 2) areas of disagreement, i.e., state why you believe there is sufficient data or sound science to proceed, and/or provide additional information which supports your perspective].*

Response to ISRP Comments:

(ISRP comments in Black, CRITFC response in Green)

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.

The primary hypothesis of this study is that functional/adaptive genetic variation is at least partially responsible for a variety of phenotypic and life history traits in salmonids in the Columbia River Basin. Many of these traits are important for recovery and conservation of steelhead and Chinook salmon ESUs that are currently listed under the ESA. This includes all five threatened ESUs for *O. mykiss*: Lower Columbia River, Middle Columbia River, Upper Columbia River, Upper Willamette River, and Snake River) and five of eight ESUs for Chinook salmon (endangered - Upper Columbia R. spring-run; threatened – Snake R. spring/summer-run; threatened – Snake R. fall-run; threatened – Lower Columbia R., threatened – Upper Willamette R., threatened) (Busby et al 1996; Federal Register 71(3) 2006).

This research affects all stocks in the Columbia River Basin and therefore is considered a basinwide application. The project addresses needs for management of natural and supplemented populations, along with reintroduction programs. These needs have been identified in multiple "Reasonable and Prudent Alternatives" (RPA) in the BiOp:

> -page 57, RPA No. 41, Preserve genetic resources – This study will provide a basis for monitoring biologically relevant genetic diversity in salmonid populations in contrast to selectively neutral levels of genetic diversity that are typically evaluated.
> -page 69, RPA No. 50, Fish population status monitoring – Once candidate markers are identified that correspond to specific traits, we plan to monitor populations throughout the Columbia River Basin to evaluate potential for thermal adaptation and smoltification.
> -page 89, RPA No. 63, Monitor hatchery effectiveness – The study will provide resources for evaluating the maintenance of biologically relevant genetic diversity in hatcheries.
> -page 89, RPA no. 64, Investigate hatchery critical uncertainties – Candidate markers can be highly useful for hatchery programs that are uncertain of which broodstock fish may offer the best chance at recovery for anadromous steelhead or those that are adapted to local environmental conditions and pathogens.

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.

The approach taken in this study is to identify functional genetic variation associated with traits of interest and includes: 1) landscape and environmental genetics through genome scans of natural populations (archived and new tissue samples) with SNP markers from functional genes; and 2) validation of genetic markers in controlled laboratory experiments (e.g., temperature stress). In objective 1, we recognize that differences in SNP allele frequencies may not be caused by selection pressure caused by temperature, so rigorous statistical analyses (i.e., Spatial Analysis Method SAM, Joost et al. 2008) have been performed to identify candidate markers associated with temperature *and* other environmental variables in natural populations (Narum et al. 2010 –submitted to Molecular Ecology; also see Section 1 of BPA Annual Report by Narum 2010). A similar approach has been taken to identify SNP markers associated with anadromy in *O. mykiss* (Narum et al. 2010 – submitted to Trans. Am. Fish. Soc.; also see Section 2 of BPA Annual Report by Narum 2010). Experiments in controlled environments (objective 2) will allow us to further screen SNP markers for validation and test for gene expression patterns associated with a phenotypic response (e.g., thermal tolerance or smoltification). Our intent is to then monitor allele frequencies of validated candidate gene markers in salmonid populations throughout the Columbia River Basin. However, we expect that candidate markers may not be useful in all populations due to recombination and other disassociation of linkage and further research will be needed to fill-in candidate markers for these exceptions. We also plan to develop models that incorporate allele frequencies of *multiple* candidate markers, environmental features, and phenotypic data to better understand local adaptation of natural populations where appropriate. This information will help us to better understand functional and biologically important variation among salmonid populations in the Columbia River that will extend our knowledge beyond connectivity (gene flow and interbreeding) that is currently estimated with neutral (non-adaptive) genetic variation. We expect that this work will have profound implications on strategies to conserve biologically relevant genetic variation of salmonids, including designation of ESU boundaries, maintenance of adaptive variation in hatchery programs, and tools to facilitate broodstock selection in hatcheries attempting to rear the most biologically appropriate stocks for a particular area (e.g., marker-assisted breeding).

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?

We will use populations of Chinook salmon throughout the Columbia River basin that occupy variable climates. Initial efforts will place focus on populations from rivers that are known to deviate from typical temperature regimes, such as fall Chinook from Crab Creek (potentially warm adapted) and Clearwater River (potentially cool adapted) relative to other natural populations (i.e., Hanford Reach).

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.

Until we perform the research, we don't know if a population is thermally adapted. We will focus on potentially adapted populations from rivers that are known to deviate from typical temperature regimes as mentioned above. Thermal refuges are certainly a possibility that we will consider in interpretation of results.

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.

We are interested in identifying candidate markers that may be associated with traits of interest, and we are well aware that SNPs may be located in coding regions as synonymous or non-synonymous polymorphisms or in cis-regulatory regions (see Narum and Campbell 2010, J. of Heredity). In many cases we expect that candidate SNP markers in our studies will be physically linked to the actual polymorphism responsible for a trait, although it is possible that we identify the underlying genetic variation.

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?

This study differs from classical QTL studies that use a large number of unknown gene markers to find QTLs for a trait of interest. In our study, we have incorporated SNPs that have been discovered from known genes/ESTs from GenBank (numerous papers on SNP discovery from our laboratory: Campbell and Narum 2008, Trans. Am. Fish. Soc.; Campbell et al. 2009, Mol. Ecol. Res., Campbell and Narum 2009, Mol. Ecol. Res.; Narum and Campbell 2010, J. of Heredity; Campbell and Narum 2010a submitted to Mol. Ecol. Res.; Campbell and Narum 2010b submitted to Trans. Am. Fish. Soc.). Thus, our panel of ~100 markers is targeted to a set of

interesting genes that may have important roles in a variety of biologically important traits. See Sections 1 & 2 of BPA Annual Report Narum 2010 for a list of gene markers.

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.

Candidate markers will be empirically evaluated for their potential utility across the Columbia River Basin. While not all candidate markers may be useful throughout the range of a species, we expect that multiple candidate markers for polygenic traits will offer enough power to be useful for broad evaluations. Linkage disequilibrium will break down from recombination (dependant upon the physical distance of candidate markers and the actual polymorphism) and level of gene flow with other populations. This suggests that not all markers will be of use for broader applications, but we will need to test this application to find these exceptions. Populations of exception may require additional study to find candidates applicable to those regions.

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.

We are well aware of differing roles of SNPs in certain positions of the genome. Our numerous SNP discovery publications provide adequate discussion of this issue (Campbell and Narum 2008, Trans. Am. Fish. Soc.; Campbell et al. 2009, Mol. Ecol. Res., Campbell and Narum 2009, Mol. Ecol. Res.; Narum and Campbell 2010, J. of Heredity; Campbell and Narum 2010a submitted to Mol. Ecol. Res.; Campbell and Narum 2010b submitted to Trans. Am. Fish. Soc.). As the ISRP suggests, we are indeed interested in validating candidate markers into "physiological/developmental understanding" and our experiments in objective 2 are intended to further this pursuit. Future work may also include intensive Next Generation sequencing of candidate regions to identify polymorphisms throughout a large genomic target region.

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 is not clear to me how the experimental design on page 8 does not provide specific details on screeing archived populations since there are multiple paragraphs regarding this issue. However, see Sections 1 & 2 of BPA Annual Report Narum 2010 for further details.

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

The initial 1-2 years of this project are targeted towards *O. mykiss*, but future years of this 10 year project will allow for evaluation of Chinook salmon as well. Further details regarding upcoming screening of Chinook salmon populations is provided in response to #2 above.

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.

It is not accurate to say "***no*** information is provided on the location or source of environmental data" as a paragraph is provided on page 6 under the heading "Landscape Data". In some circumstances, long term environmental data will be more critical to characterize climate regimes rather than debating the use of air vs. water temperatures that are typically highly correlated (see Section 1 of BPA Annual Report Narum 2010 for further details). However, we agree that better environmental data will lead to more clear interpretation of results.

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.

Extensive information is presented in the narrative regarding statistical analyses, however it may be more clear to see the results of these tests in Section 1 & 2 of BPA Annual Report Narum 2010.

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.

In question are the fish to be used in the controlled experiments for thermal adaptation. As stated in the narrative, gametes will be collected from one representative population from desert and montane environments based on screening completed in Objective 1. This initial screening should provide adequate confidence that alleles of interest will be carried in the test populations. The SNP discovery reference of Campbell et al. 2009 was provided for further information regarding SNP markers and corresponding ESTs/genes, but our laboratory has several SNP discovery publications that should resolve any questions on this matter (Campbell and Narum 2008, Trans. Am. Fish. Soc.; Campbell et al. 2009, Mol. Ecol. Res., Campbell and Narum 2009, Mol. Ecol. Res.; Narum and Campbell 2010, J. of Heredity; Campbell and Narum 2010a submitted to Mol. Ecol. Res.; Campbell and Narum 2010b submitted to Trans. Am. Fish. Soc.). A table of 96 SNP markers and their corresponding gene name is available in Section 1 & 2 of BPA Annual Report Narum 2010. RNA expression will be quantified in various tissues through real-time quantitative PCR completed on an existing ABI 7900HT instrument located at the Hagerman Fish Culture Experiment Station.

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.

We are indeed interested in validating candidate markers into "physiological/developmental understanding" as mentioned above and our experiments in objective 2 are intended to further this pursuit. Future work may also include intensive Next Generation sequencing of candidate regions to identify the actual polymorphisms throughout a large genomic target region.

Candidate markers will be empirically evaluated for their potential utility across the Columbia River Basin. While not all candidate markers may be useful throughout the range of a species, we expect that multiple candidate markers for polygenic traits will offer enough power to be useful for broad evaluations. Linkage disequilibrium will break down from recombination (dependant upon the physical distance of candidate markers and the actual polymorphism) and level of gene flow with other populations. This suggests that not all markers will be of use for broader applications, but we will need to test this application to find these exceptions.

Populations of exception may require additional study to find candidates applicable to those regions.

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.

We fully agree that candidate markers will need to be empirically evaluated for their potential utility across the Columbia River Basin. The genetic data for this will come from our SNP baselines that are under development in BPA project 2008-907-00. As indicated by ISRP, candidate markers provide us with legitimate targets for further 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.

In question are the fish to be used in the controlled experiments for thermal adaptation and smoltification studies. As stated in the narrative, gametes will be collected from one representative population from each desert and montane environments based on screening completed in Objective 1. This initial screening should provide adequate confidence that alleles of interest will be carried in the test populations. The narrative also provides ample information regarding the clonal lines that will be used to evaluate smoltification.

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.

Through extensive SNP discovery efforts in our lab, we have identified numerous polymorphic segregating alleles from broad ascertainment panels (Campbell and Narum 2008, Trans. Am. Fish. Soc.; Campbell et al. 2009, Mol. Ecol. Res., Campbell and Narum 2009, Mol. Ecol. Res.; Narum and Campbell 2010, J. of Heredity; Campbell and Narum 2010a submitted to Mol. Ecol. Res.; Campbell and Narum 2010b submitted to Trans. Am. Fish. Soc.). A total of 96 SNPs were chosen for screening based on levels of polymorphism observed in fish specifically from the Columbia River Basin. It is our intent to screen these SNPs in archived tissues (Objective 1) and also in fish under controlled environments (Objective 2). We plan to test both natural fish and aquaculture strains as available for each trait of interest as discussed above. Initial screening of natural populations should provide adequate confidence that alleles of interest will be carried, and inclusion of additional stocks will give an indication whether these candidate markers will transfer to other stocks.

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).

These references provided by ISRP are noted.

REFERENCES:

Campbell, N. R., and S. R. Narum. 2008. Identification of novel SNPs in Chinook salmon and variation among life history types. Transactions of the American Fisheries Society 137:96-106.

Campbell, N.R., K. Overturf, and S. Narum. 2009. Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9:318-322.

Campbell, N.R., and S. Narum. 2009. Identification and characterization of heat shock response related single nucleotide polymorphisms in O. mykiss and O. tschawytscha. Molecular Ecology Resources 9:1460-1559.

Narum, S.R., and N.R. Campbell. 2010. Sequence divergence of heat shock genes within and among three Oncorhynchids. Journal of Heredity doi:10.1093/jhered/esp081

Campbell, N.R., and S. Narum. Submitted. Relative Genetic diversity estimates among four Pacific salmonids. Transactions of the American Fisheries Society.

Campbell, N.R., and S. Narum. Submitted. Development of 54 novel SNP assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. Molecular Ecology Resources.