

**Connectivity and Gene Flow among Oregon Chub Populations
In the Middle Fork Willamette River**

Draft Report Submitted

11-4-15

Final Report Submitted

12-11-15

Patrick DeHaan and Brice Adams
U.S. Fish and Wildlife Service
Abernathy Fish Technology Center
1440 Abernathy Creek Rd
Longview, WA 98632
patrick_dehaan@fws.gov, brice_adams@fws.gov
360-425-6072

Brian Bangs and Paul Scheerer
Oregon Department of Fish and Wildlife
Native Fish Investigations Project
28655 Hwy 34
Corvallis, OR 97333
Brian.Bangs@oregonstate.edu, Paul.Scheerer@oregonstate.edu
541-757-4263

Background

Oregon chub (*Oregonichthys crameri*) are a small floodplain minnow native to the Willamette River system in western Oregon. Oregon chub are typically found in off channel habitats such as sloughs and oxbows that are periodically connected to each other and to mainstem rivers (Markle et al. 1991). Although Oregon chub were believed to be historically abundant throughout the Willamette Basin, surveys in the 1970s and 1980s indicated that the species distribution and abundance had severely declined (Bond 1974; Bond and Long 1984; Markle et al. 1991). Factors implicated in the species decline include the introduction of non-native predators as well as flood control activities that eliminated chub habitat and the historic connections between the floodplain and the mainstem Willamette (Markle et al. 1991; Scheerer 2002). As a result of severe declines, the species was listed as endangered in 1993 under the U.S. Endangered Species Act (U.S. Fish and Wildlife Service; USFWS 1993). Since then, a variety of conservation measures have been enacted to aid the recovery of Oregon chub, and the species has recovered to the point that it was removed from the Endangered Species List in 2015 (USFWS 2015).

A previous study found that Oregon chub were most abundant at isolated sites where non-native species were absent (Scheerer 2002). Managing populations of Oregon chub in isolation to avoid the threat of non-native species is counter to the species life history however; chub likely relied on connectivity between floodplains and mainstem rivers for dispersal to new sites and connectivity and genetic exchange among populations. In 2009, the Oregon Department of Fish and Wildlife (ODFW) initiated a study to examine the relationship between physical habitat characteristics and the fish community assemblages at several sites suitable for Oregon chub in the Dexter-Jasper reach of the Middle Fork (MF) Willamette River (Bangs et al. 2015). This area of the MF Willamette has several sites where Oregon chub are found with varying degrees of connectivity to the mainstem river. One objective of this study is to determine conditions that facilitate movement of Oregon chub among the different sites. As part of this study, ODFW has used physical tags and mark-recapture data to document movements of Oregon chub between several connected habitats in the Dexter-Jasper reach (Bangs et al. 2015). Despite limited evidence that shows Oregon chub move among sites in the Dexter-Jasper reach, the number, life stage, and timing of individuals moving between sites is mostly unknown.

Connectivity among populations/sites can be difficult to measure directly, especially with species such as Oregon chub that are difficult to tag and recapture, and when the time and life stages at which individuals migrate among sites are unknown. Genetic data offer an alternative and/or complementary means to examine movement and connectivity (Lowe and Allendorf 2010; Taylor et al. 2011). Unlike traditional methods for assessing connectivity that rely on physically marking individuals (e.g., radio tracking, PIT tagging), genetic methods do not necessarily require that individuals be handled multiple times and marked to determine if they are migrants from another population/site. Furthermore genetic data provide information from a broader time scale than traditional mark-recapture studies which are generally limited to the time over which individuals were sampled. When used in conjunction with traditional means of tracking individual movements, genetic data allow biologists not only to infer which individuals move among sites and when, but also whether or not individuals contribute genes to their new population and the frequency of genetic exchange among sites/populations.

Previously, DeHaan et al. (2012) conducted a range-wide analysis of genetic variation in Oregon chub populations that included several collections from the MF Willamette River. The authors found that each site contained a genetically unique population; however, there appeared to be a much greater level of genetic exchange among populations in the MF Willamette. This study included two collection sites within the Dexter-Jasper reach: Elijah Bristow Northeast Slough and Elijah Bristow Berry Slough. All other populations from the MF Willamette were located upstream of major flood control dams. Although these data suggest that Oregon chub do move among sites within the MF Willamette, most of the sites where chub are found in the Dexter-Jasper reach were not sampled as part of that study, and it's difficult to assess the degree of movement among sites and the level of genetic exchange based on data from only these two sites. Since that initial study, additional genetic sampling has occurred in the Dexter-Jasper reach as part of the ODFW floodplain study.

The objective of our study was to use genetic data to help determine the degree of connectivity among Oregon chub sites within the Dexter-Jasper reach. This information will complement the movement data currently being collected by ODFW and will help to determine what physical habitat conditions help facilitate connectivity and genetic exchange among sites. This information is important for maintaining healthy populations of Oregon chub now that the species has been removed from the Endangered Species List.

Methods

Oregon chub were collected from nine connected sites in the Dexter-Jasper Reach of the MF Willamette River (Table 1; Figure 1) following the methods outlined in Bangs et al. (2015). A small tissue sample was taken from up to 50 individuals from each location (see Table 1 for sample sizes) and preserved in 100% non-denatured ethanol. Total genomic DNA was extracted from all samples using Qiagen DNeasy 96 tissue extraction kits (Qiagen Inc., Valencia CA) following the manufacturer's instructions. All individuals were then genotyped at nine microsatellite loci following the methods outlined in DeHaan et al. (2012).

Individuals were grouped according to collection site for statistical analysis. We added fish collected in 2004 and 2005 from two sites in the Dexter-Jasper reach (Elijah Bristow Berry Slough and Elijah Bristow Northeast Slough) that were previously genotyped and analyzed to the dataset for statistical analysis. The addition of the 2004-2005 Elijah Bristow NE Slough collection allowed us to examine temporal variance in allele frequencies at this site. We also added fish collected from two sites outside the study area to our analysis for comparison purposes: EF Minnow Creek – a site in the MF Willamette upstream of Dexter Dam, and Big Island – a site in the McKenzie River (another Willamette subbasin). Collection and genotyping methods for these additional individuals can be found in DeHaan et al. (2012).

A variety of methods exist for examining the degree of genetic connectivity among populations (Lowe and Allendorf 2010). These include indirect methods which make inferences about the amount of gene flow among populations based on the level of differentiation among populations (i.e., F_{ST} ; Weir and Cockerham 1984) as well as direct methods which can be used to identify migrant individuals in a population (i.e., assignment testing; Manel et al. 2005) or quantify the level of contemporary gene flow among populations (Wilson and Rannala 2003). The application of these methods depends on the level of differentiation among populations; increasing genetic differentiation among sites/populations facilitates the use of more informative direct estimates (Lowe and Allendorf 2010). Because of this, we employed a variety of methods to first determine if the various collection sites represented independent populations or a single population and what the level of genetic variation among collection sites was.

Each collection was tested for conformance to Hardy-Weinberg equilibrium expectations (HWE) using exact tests implemented in the program GENEPOP v 4.1 (Raymond and Rousset

1995). We also tested collections for evidence of linkage disequilibrium (LD) using GENEPOP. Significance tests for HWE and LD tests were adjusted for multiple comparisons using a sequential Bonferroni adjustment (Rice 1989). After conducting HWE and LD tests for each collection site individually, we combined all of the sites in the Dexter-Jasper reach into a single population and conducted HWE and LD tests for that population. The purpose of combining all of the sites was to determine if all of the collection sites combined met expectations for a single, randomly mating population.

We used GENEPOP to determine the level of genetic variation among collection sites (i.e., pairwise F_{ST}). We also used GENEPOP to conduct contingency tests of allele frequency heterogeneity to determine if there were significant differences in allele frequencies among collection sites. In order to visualize the level of genetic variation among sites, we conducted a discriminant analysis of principal components (DAPC; Jombart et al. 2010) based on our allele frequency data. DAPC is similar to principal component analysis (PCA), but unlike PCA which maximizes the total variation in the dataset, DAPC maximizes the variation among different groups and minimizes variation within groups (Jombart et al. 2010). We conducted DAPC analysis using the *adegenet* package (Jombart 2008) for the R statistical environment (R Core Development Team 2015). We also used the Bayesian clustering methods implemented in the program STRUCTURE v2.3.4 (Pritchard et al. 2000) to determine the level of variation among sites and how many populations there were within our study area. STRUCTURE clusters individuals into a pre-defined number of populations (K) based on HWE expectations and allows the user to determine the most likely K for a dataset from a range of values. We tested K from 1 to 15 for our dataset. Each STRUCTURE analysis consisted of 20 replicate runs with 100,000 burn-in iterations followed by 300,000 data collection iterations. The optimal value of K was selected by examining posterior probability values as well as the graphical output from program.

Oftentimes fishes exhibit an isolation-by-distance (IBD) pattern of genetic variation where geographically proximate sites/populations show greater genetic similarity due to frequent genetic exchange. Previous work on Oregon chub observed this pattern for fish in the MF Willamette Basin (DeHaan et al. 2012). We used GENEPOP to conduct an IBD analysis for the Dexter-Jasper sites only by comparing the natural log of geographic distance in meters between collection sites to the pairwise genetic distance between sampling locations measured as $F_{ST}/(1 - F_{ST})$. We performed a Mantel test (1,000 permutations) to determine whether there was a

significant IBD relationship. Alternatively, genetic exchange among Oregon chub sites may be more of a function of habitat characteristics (e.g., level of connectivity, habitat size, habitat quality) than geographic distance. We also examined the relationship between habitat connectivity and genetic distance. We compared the mean number of days from 2012-2014 each site was hydrologically connected to the mainstem of the MF Willamette (see Bangs et al. 2015) to two measures of genetic distance: the mean pairwise F_{ST} for each collection site and the mean Cavalli-Sforza and Edwards' (1967) chord distance. We used R to calculate Pearson product moment correlation coefficients between mean connectivity and the two measures of genetic distance.

Results and Discussion

Following Bonferroni correction, none of the sites in the Dexter-Jasper reach showed any departures from HWE expectations. Big Island (located in the McKenzie River) deviated from HWE expectations at the locus *Ocr111* due to a heterozygote deficiency and EF Minnow Creek (located upstream of Dexter Dam) deviated from HWE expectations at the locus *Ocr105* also due to a heterozygote deficiency. Only the 2010 Elijah Bristow NE Slough collection showed evidence of linkage at a single locus pair. When all sites in the Dexter-Jasper reach were combined into a single population, the locus *Ocr109* deviated from HWE expectations due to a heterozygote deficit. Two pairs of loci (out of 36 total) showed evidence of linkage in the combined population.

Pairwise F_{ST} values among sites within the Dexter Jasper Reach ranged from -0.006 (essentially 0.0) for the comparison between Jasper Railroad Bridge and Dougren Slough to 0.023 between Elijah Bristow Berry Slough and Deep Muddy Slough and Jasper Railroad Bridge and the 2004-2005 Elijah Bristow NE Slough collection (Table 2). All pairwise F_{ST} values for the two populations outside of the study area were greater than those observed for Dexter-Jasper sites and ranged from 0.024 to 0.069 (Table 2). The number of loci (out of nine) that showed significant differences in allele frequencies between sites within the Dexter-Jasper Reach ranged from zero to three and from three to nine for the two populations outside of the study area (Table 2). Comparison of the two Elijah Bristow NE slough collections showed essentially no change in allele frequencies at this site over the course of five years.

DAPC plots for the Dexter-Jasper collection sites showed that individuals from the same collection site tended to cluster together, however there was considerable overlap among the different collection sites (Figure 2). Some of the Elijah Bristow sites (Northeast Slough, Berry Slough) were somewhat separated from the main group (Figure 2). When the samples from Big Island and EF Minnow Creek were added to the DAPC plots, the Big Island individuals clustered separately from the sites in the MF Willamette River (Figure 2). STRUCTURE analysis showed that K of 1 had the highest posterior probability. The graphical output for higher values of K (2-15) showed that individuals had essentially equal proportions of their genetic material associated with each cluster or population (Figure 3), often a sign that a lower K value is more appropriate. We did not observe a significant IBD relationship among the collection sites within the Dexter-Jasper reach (Spearman $r = 0.11$, Mantel test $P = 0.189$; Figure 4). We also did not observe a strong relationship between mean connectivity for each collection site and the two genetic distance metrics (Figure 5).

Direct tests to identify migratory individuals and quantify migration rates based on genetic data depend upon the ability to accurately differentiate populations. Tests for HWE and LD can be used to determine if collections of individuals are representative of a randomly mating population (Allendorf and Luikart 2007). Although none of the individual collection sites showed any deviations from HWE or LD, when all samples were combined into a single population, results of HWE and LD tests were consistent with a single population sample. A number of genetic methods exist to differentiate populations (Waples and Gaggiotti 2006) and we used several of these methods to determine if there were multiple populations of Oregon chub within the study area. Pairwise F_{ST} estimates among collection sites in the Dexter-Jasper reach were relatively low, with several estimates close to or equal to 0.0. We also observed relatively few significant differences in allele frequencies among collection sites. STRUCTURE analysis suggested that only a single population was present in our dataset. All of these data suggest that the Oregon chub in the Dexter-Jasper reach of the MF Willamette River represent a single population. Because collection sites were not genetically differentiated from one another, it was not possible to identify specific individuals as migrants from one population to the next. Instead we can conclude that there are relatively high rates of migration and genetic exchange among the different collection sites in the Dexter-Jasper reach. To date, mark-recapture data has documented relatively few individuals moving among sites (Bangs et al. 2015). Clearly these

two data sets in combination provide a much more complete picture of Oregon chub movements and gene flow in this section of the MF Willamette River.

Although our data suggest the presence of a single population, there did seem to be collection sites that were slightly differentiated from the others. The DAPC plots of just the Dexter-Jasper collection sites showed that some of the Elijah Bristow sites (e.g., Northeast Slough, Berry Slough) clustered somewhat separately from the rest of the collection sites. Although these sites are geographically separated from several other sites downstream, we did not observe a strong IBD relationship indicating that they are not differentiated because of their physical location. Some of the Elijah Bristow sites had relatively low connectivity (Bangs et al. 2015) which may restrict opportunities for genetic exchange at these sites. However, we did not observe a strong relationship between connectivity and genetic distance within the Dexter-Jasper reach. These data suggest that additional factors may influence genetic variation within the study area. For example, even though some sites were not frequently connected to the mainstem MF Willamette River, the times they are connected may coincide with the times that chub typically disperse from one site to another.

There are additional sites within the Dexter-Jasper reach where chub are found that we did not sample for this study (Figure 1). Collection of fish from these additional sites, as well as the addition of more fish from sites with relatively low sample sizes (e.g., Jasper Railroad Bridge, Dexter Dam Slough), and perhaps temporal replicate samples may be useful in the future for a more thorough examination of movement and genetic variation of Oregon chub within the Dexter-Jasper reach. Based on the data presented in this study, Oregon chub migrate extensively within the Dexter-Jasper reach and there is a high degree of genetic exchange among different sites chub are found.

Acknowledgements

Funding for this study was provided by the US Army Corps of Engineers Cooperative Agreement numbers W9127N-09-2-0007 and 14-2-0009-0001 and administered through the Oregon Department of Fish and Wildlife. We greatly appreciate the efforts of all the ODFW technicians who collected Oregon chub tissue samples for this study and Kieslana Wing for providing laboratory assistance. Patty Crandell and Christian Smith provided useful comments on an earlier version of this report. All data from this study have been archived in the AFTC

Conservation Genetics Lab Progeny database. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

References

- Allendorf, F. W., and G. K. Luikart. 2007. Conservation and the Genetics of Populations, 1st edition. Blackwell Publishing, Oxford.
- Bangs, B. L., P. D. Scheerer, and S. Clements. 2015. Effects of U.S. Army Corps of Engineers Willamette project operations on Oregon chub and other floodplain fishes (2009-2012). Oregon Department of Fish and Wildlife, Corvallis, OR. Final Report, 306pp.
- Bond, C. E. 1974. Endangered plants and animals of Oregon: I. fishes., Agricultural Experiment Station, Oregon State University.
- Bond, D. D., and J. J. Long. 1984. Status of the Oregon chub, *Hybopsis crameri*. Pages 483-486 in Proceedings of the Annual Conference of Western Fish and Wildlife Agencies.
- Cavalli-Sforza, L.L. & Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 21: 550-570.
- DeHaan, P. W., P. D. Scheerer, R. Rhew, and W. R. Ardren. 2012. Analyses of genetic variation in Oregon chub, a threatened floodplain minnow in a highly altered environment. *Transactions of the American Fisheries Society* 141(2):533-549.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403-1405.
- Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *Bmc Genetics* 11.
- Lowe, W. H., and F. W. Allendorf. 2010. What can genetics tell us about population connectivity? *Molecular Ecology* 19(15):3038-3051.
- Manel, S., O. E. Gaggiotti, and R. S. Waples. 2005. Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution* 20(3):136-142.
- Markle, D. F., T. N. Pearsons, and D. T. Bills. 1991. Natural history of *Oregonichthys* (Pices: Cyprinidae) with a description of a new species from the Umpqua River of Oregon, USA. *Copeia* 1991(2):277-293.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-959.
- Raymond, M., and F. Rousset. 1995. GENEPOP (Version-1.2) - Population-genetics software for exact tests and ecumenicism. *Journal of Heredity* 86(3):248-249.

- R. Core Development Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43(1):223-225.
- Scheerer, P. D. 2002. Implications of floodplain isolation and connectivity on the conservation of an endangered minnow, Oregon chub, in the Willamette River, Oregon. *Transactions of the American Fisheries Society* 131(6):1070-1080.
- Taylor, E. B., J. L. Gow, J. Witt, and R. Zemplak. 2011. Connectivity among populations of pygmy whitefish (*Prosopium coulterii*) in northwestern North America inferred from microsatellite DNA analyses. *Canadian Journal of Zoology* 89(4):255-266.
- USFWS. 1993. Endangered and threatened wildlife and plants: determination of endangered status for the Oregon chub. *Federal Register* 58(18 October 1993):53800-53804.
- USFWS. 2015. Endangered and threatened wildlife and plants; removing the Oregon chub from the list of endangered and threatened wildlife. *Federal Register* 80: (19 February 2015):9126-9150.
- Waples, R. S., and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15(6):1419-1439.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38(6):1358-1370.
- Wilson, G. A., and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163(3):1177-1191.

Table 1. Oregon chub collection sites in the Dexter-Jasper reach of the MF Willamette River. Sample size (n) represents the number of individuals genotyped from each site for analysis.

Collection Location	Date Collected	n	Zone	Easting	Northing
Pengra Oxbow Slough	9/2/2010	38	10T	511070	4867078
Jasper Railroad Bridge Slough	9/28/2010	6	10T	508890	4869217
Elijah Bristow Island Pond	9/15/2010	47	10T	514043	4864334
Pengra Island Slough	9/2/2010	20	10T	510618	4868035
Elijah Bristow South Slough	9/15/2010	41	10T	510780	4867977
Dougren Slough	9/9/2010	32	10T	510579	4868037
Dexter Dam Slough	8/31/2010	14	10T	514707	4683601
Deep Muddy Slough	9/9/2010	34	10T	508927	4869191
Elijah Bristow NE Slough	8/31/2010	43	10T	514931	4864608
Elijah Bristow NE Slough	2004-2005	42	10T	514931	4864608
Elijah Bristow Berry Slough	2004-2005	47	10T	513061	4865107

Table 2. Genetic variation among Oregon chub collection sites. Values below the diagonal represent pairwise F_{ST} estimates and values above the diagonal represent the number of loci (out of 9) that showed significant differences in allele frequencies. Big Island and EF Minnow Creek were sites located outside of the Dexter-Jasper reach.

	Deep Muddy Slough	Dexter Dam Slough	Dougren Slough	EB Island Slough	EB NE Slough (2010)	EB South Slough	EB Berry Slough	EB NE Slough (04-05)	Jasper Railroad Bridge	Pengra Island Slough	Pengra Oxbow	Big Island	EF Minnow Cr
Deep Muddy Slough	-	0	0	1	3	3	2	3	0	1	1	8	5
Dexter Dam Slough	0.011	-	0	0	0	0	1	2	0	0	0	4	3
Dougren Slough	0.005	0.004	-	1	0	0	1	1	0	0	0	7	4
EB Island Slough	0.015	-0.005	0.010	-	2	1	1	2	0	0	1	8	6
EB NE Slough (2010)	0.016	0.010	0.008	0.013	-	0	2	2	0	0	4	7	4
EB South Slough	0.020	-0.002	0.011	0.004	0.005	-	1	2	0	0	1	8	5
EB Berry Slough	0.023	0.017	0.016	0.021	0.015	0.015	-	0	0	1	2	9	2
EB NE Slough (04-05)	0.020	0.018	0.013	0.017	0.008	0.018	0.009	-	1	2	2	7	5
Jasper Railroad Bridge	0.003	0.003	-0.006	0.006	0.006	0.007	0.022	0.023	-	0	0	4	1
Pengra Island Slough	0.015	0.009	0.006	0.007	0.008	0.008	0.021	0.017	-0.003	-	0	6	3
Pengra Oxbow	0.008	0.005	0.006	0.011	0.012	0.011	0.014	0.016	0.004	0.003	-	7	5
Big Island	0.067	0.045	0.058	0.067	0.059	0.062	0.057	0.059	0.073	0.069	0.063	-	8
EF Minnow Creek	0.038	0.027	0.035	0.044	0.028	0.033	0.024	0.032	0.051	0.045	0.030	0.056	-

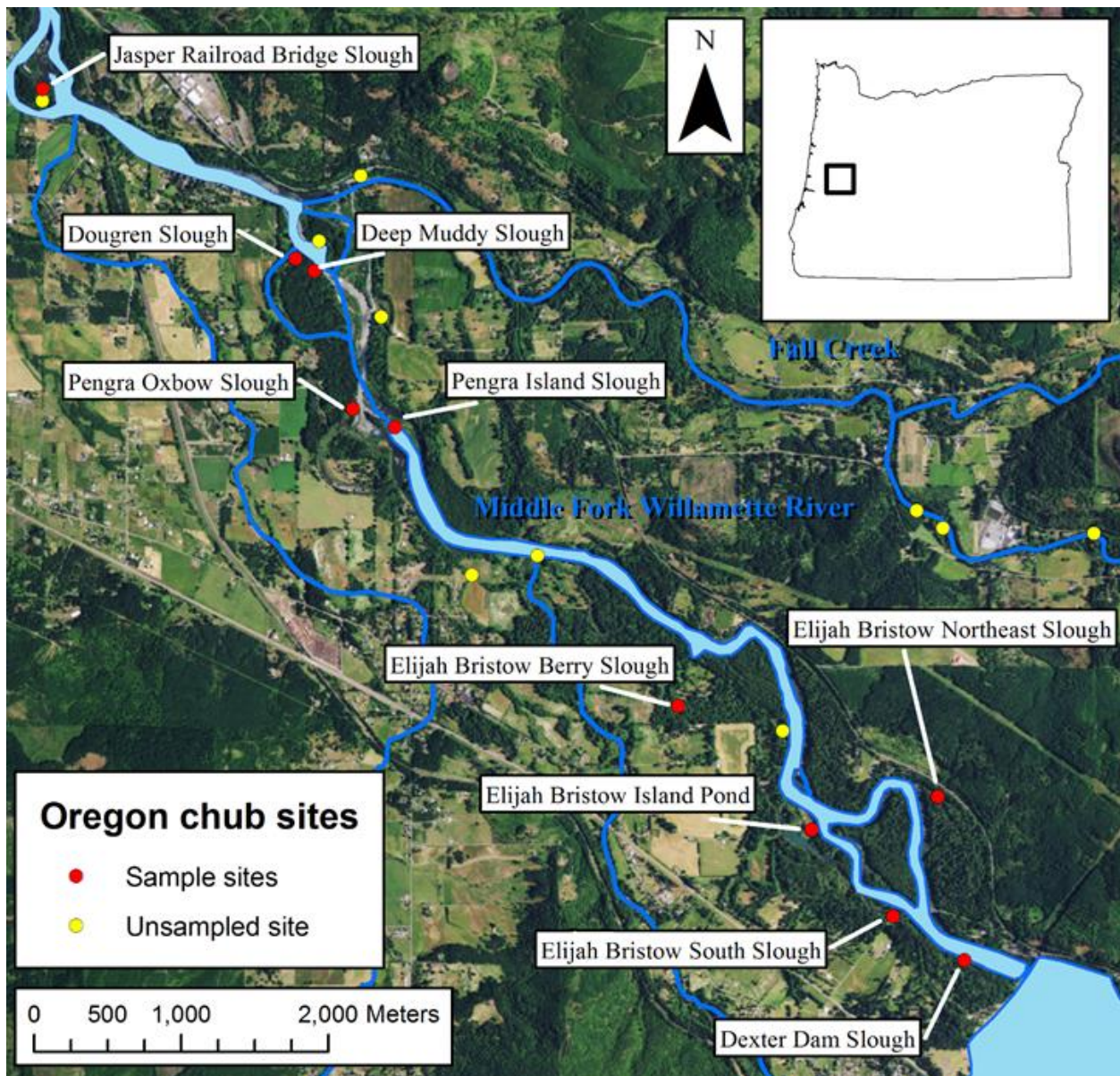


Figure 1. Locations in the Dexter-Jasper reach of the MF Willamette River where Oregon chub were sampled for this study.

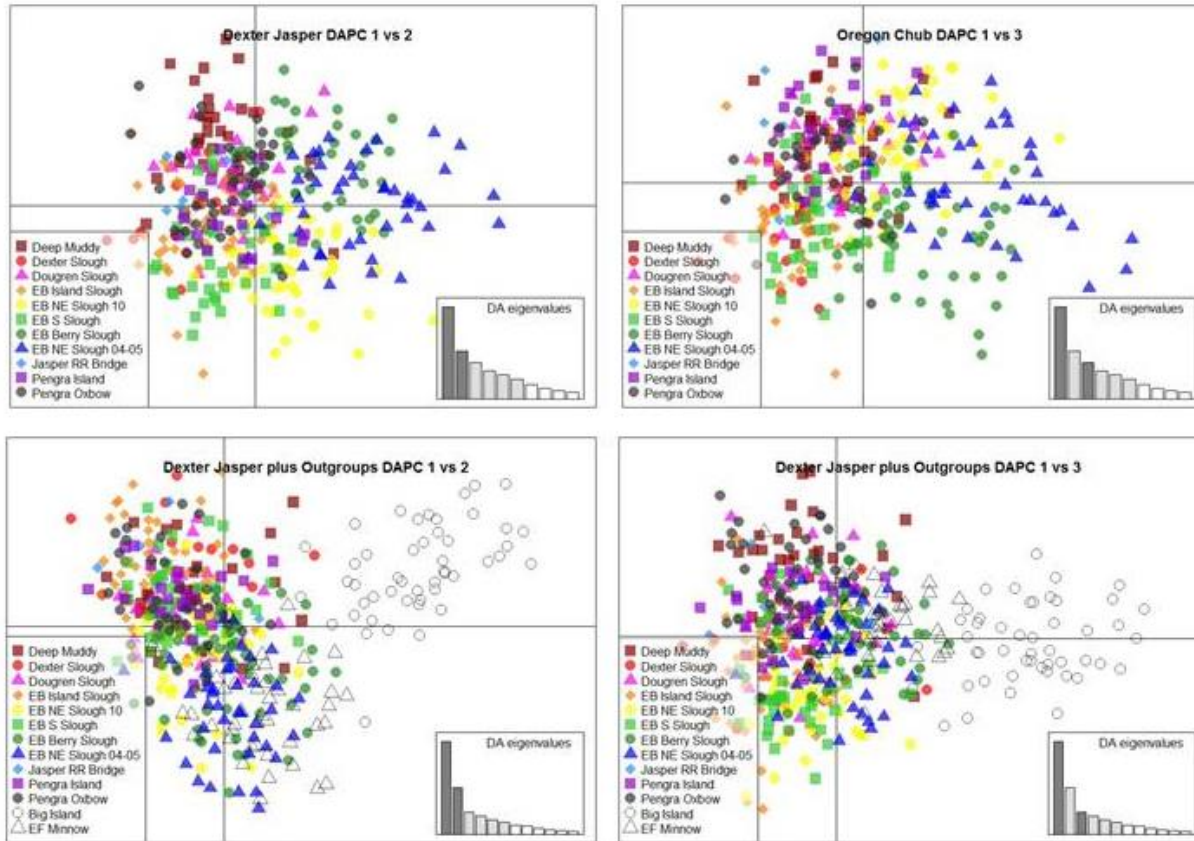


Figure 2. Discriminant analysis of principal components (DAPC) plots based on allele frequency data. Each point on the graph represents an individual fish and colors and symbols correspond to collection locations. The top two plots represent collection sites within the Dexter-Jasper reach and the bottom two plots represent the Dexter-Jasper sites plus two populations outside the study area. Plots on the left represent the first and second principle components and plots on the right represent the first and third principle components.

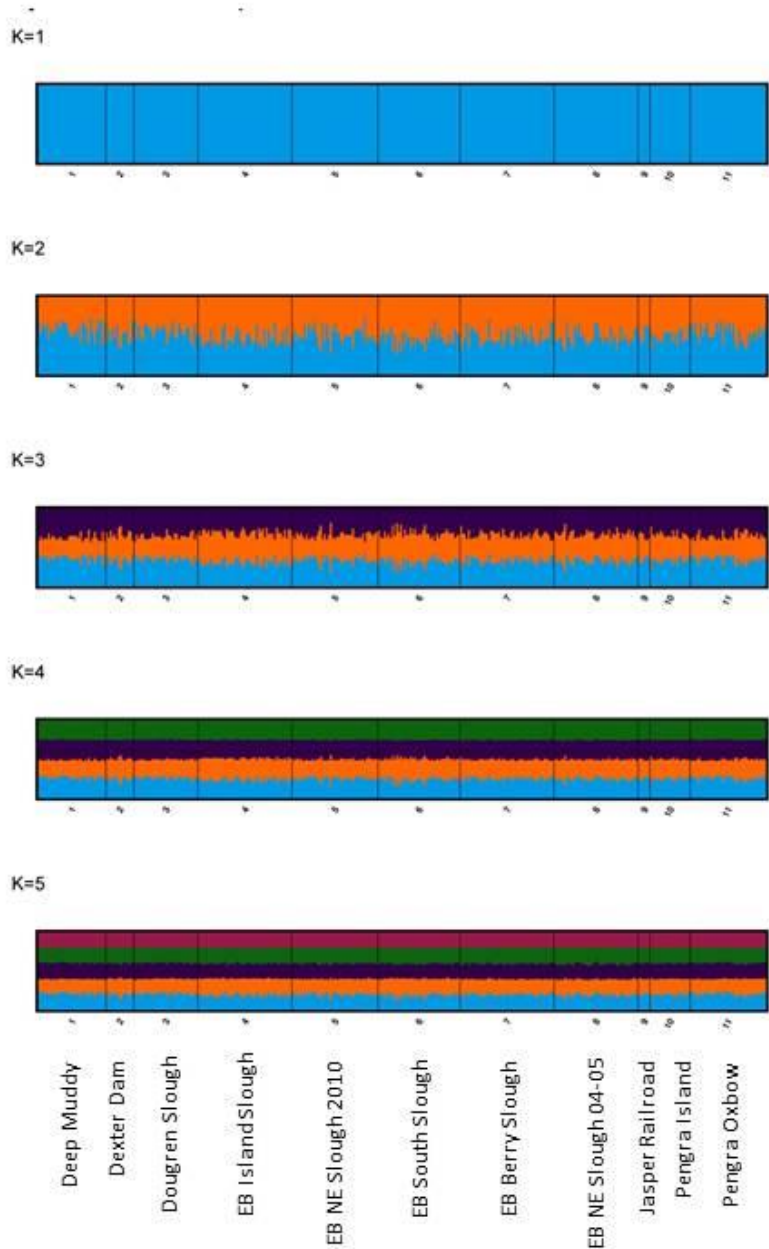


Figure 3. STRUCTURE results for Oregon chub collection sites in the Dexter-Jasper reach. Each vertical bar on the graph represents an individual fish and the shading represents the proportion of that individual's genetic material that corresponds to each genetic cluster. Results for $K = 1$ to 5 populations are shown. Additional values of K up to 15 all showed the same pattern, where each individual had roughly equal proportions of its genetic material attributed to each cluster.

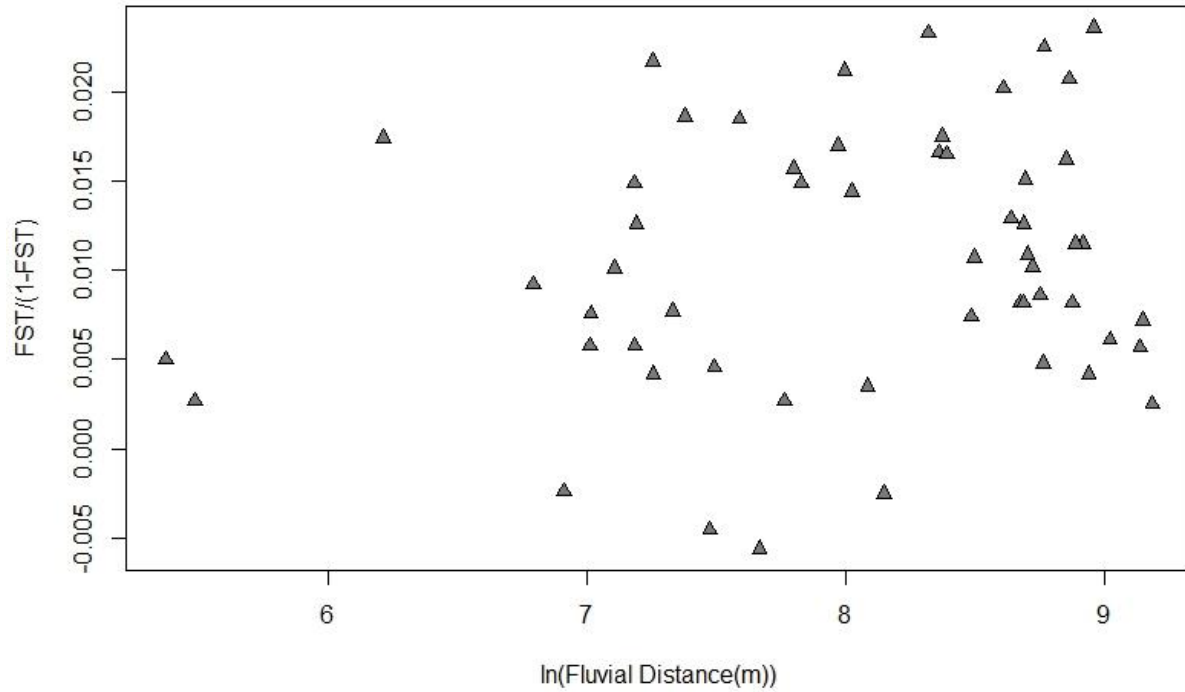


Figure 4. Relationship between geographic distance and genetic distance among 11 Oregon chub sites in the Dexter-Jasper reach. Geographic distance (x-axis) is represented by the natural log of fluvial distances among collection sites and genetic distance (y-axis) is represented by $F_{ST}/(1 - F_{ST})$. A Mantel test showed there was no significant relationship ($P = 0.189$).

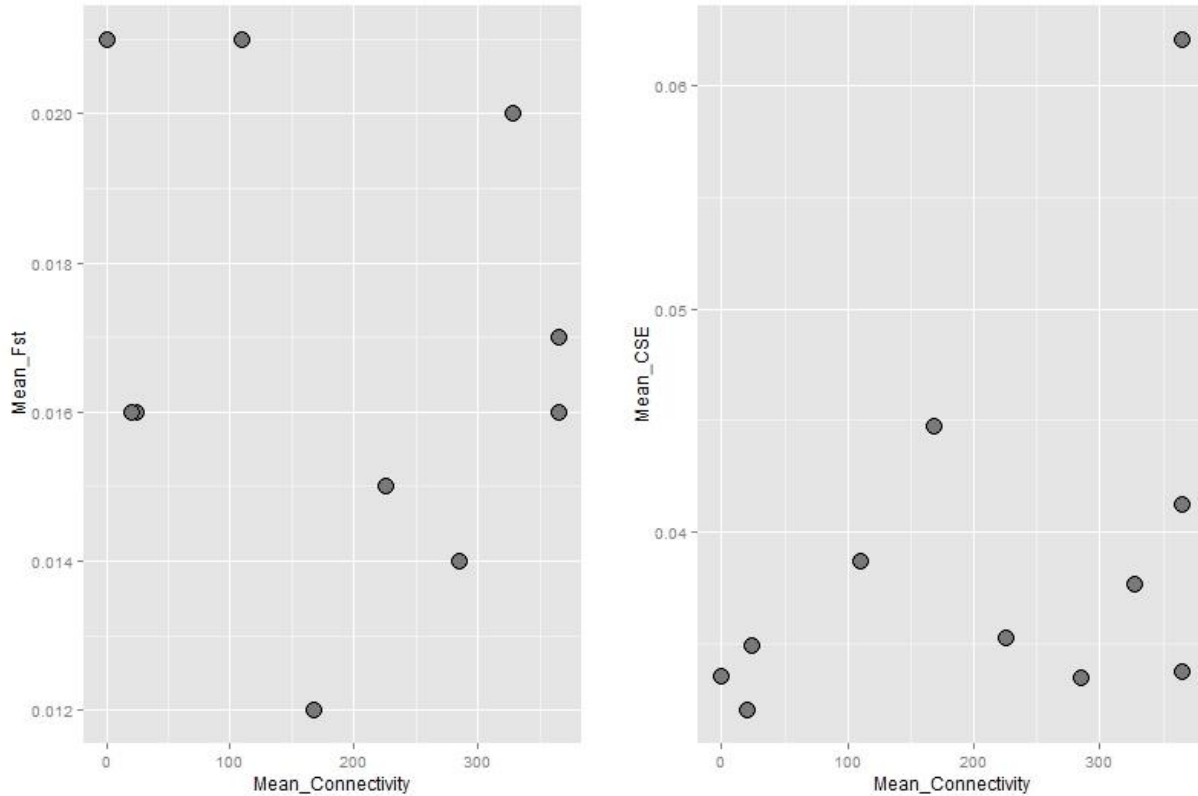


Figure 5. Relationship between habitat connectivity and two genetic distance metrics. Habitat connectivity was measured as the mean number of days from 2012-2014 that a site was hydrologically connected to the mainstem MF Willamette River. The plot on the left represents the relationship between connectivity and the mean pairwise F_{ST} for a collection site (Pearson $r = -0.17$, $P = 0.615$). The plot on the right represents the relationship between connectivity and the mean Cavalli-Sforza and Edwards chord distance for a collection site (Pearson $r = 0.43$, $P = 0.183$).