

# Genetics

## Annotated Bibliography

**Anderson, E.C. and J.C. Garza.** 2005. A description of full parental genotyping. Report submitted to the Pacific Salmon Commission, Vancouver, British Columbia. 11 p.

<https://swfsc.noaa.gov/publications/FED/00675.pdf>

How full parental genotyping (FPG) could be implemented to “tag” hatchery salmon is described. The broodstock at hatcheries would be genotyped every year and would then be kept in a database (the parent database). The genotypes of sampled fish (in fisheries, at hatcheries, etc.) could then be compared to the parent database. The parents of any fish sampled could be determined given that its parents are in the database. Knowledge of the parent pair provides the hatchery of origin and the cohort of a recovered fish. The parent database could readily be integrated into a GSI database that shares molecular markers and would thus provide information about stock of origin (and sometimes cohort of origin) for all sampled fish (whether marked or unmarked).

**Anderson, E.C. and J.C. Garza.** 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics* 172: 2567-2582.

Importance-sampling algorithms are provided which permit likelihood-based inference of parentage in large studies with many potential parents and offspring. The computation of tail probabilities is accelerated when using SNPs for parentage inference with the importance-sampling algorithms. Simulations using the importance-sampling algorithms to calculate the power of SNPs for large-scale parentage studies show that 60-100 SNPs may allow accurate pedigree reconstruction, even in situations involving thousands of potential parents and offspring.

**Banks, M.A., D.P. Jacobson, I. Meusnier, C.A. Greig, V.K. Rashbrook, W.R. Ardren, C.T. Smith, J. Bernier-Latmani, J. Van Sickle, and K.G. O'Malley.** 2014. Testing advances in molecular discrimination among Chinook salmon life histories: evidence from a blind test. *Animal Genetics* 45(3): 412-420.

Simulation and blind tests were conducted for 623 known-origin Chinook salmon to compare and contrast the accuracy of different population sampling baselines and microsatellite loci panels. The identification of fall-run with great accuracy was presented. The overall correct assignment prediction from simulations was 98% while the blind test success rate was 84%. Fewer of winter run, spring from Butte Creek, and spring from Deer and Mill creek's assignments were correct than predicted.

**Banks, M.A., V.K. Rashbrook, M.J. Calavetta, C.A. Dean, and D. Hedgecock.** 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of Chinook salmon (*Oncorhynchus tshawytscha*) in California's Central Valley. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 915-927.

Ten microsatellite DNA markers were used to assess genetic diversity of the four runs (winter, spring, fall, and late fall) of Chinook salmon in the Central Valley. There was substantial genetic divergence among runs. Two distinct lineages of spring run were discovered; Butte Creek and Mill/Deer creeks. Genetic structure was observed to accord with the diverse Chinook salmon life histories seen in the Central Valley and provides a means for discrimination of protected populations.

**Garza, J.C. and E.D. Crandall.** 2013. Genetic analysis of Chinook salmon from the Napa River, California. Fisheries Ecology Division, Southwest Fisheries Science Center, NOAA Fisheries and the Institute of Marine Sciences, University of California, Santa Cruz.

<http://naparcd.org/documents/NapaRiverChinookReport2013.pdf>

Chinook salmon were absent from the Napa River in the 1980s and 1990s when California Chinook salmon ESUs were being defined. Genetic analysis was performed on samples taken from juveniles caught in downstream migrant trapping and from adult Chinook salmon carcasses in the Napa River watershed. The most likely origin of these fish, or their ancestors, was evaluated. The number of spawners (effective population size) that contributed to the juveniles sampled was estimated. All Chinook salmon sampled assigned to the Central Valley fall/late fall run except for four juveniles in 2010 that assigned to the Klamath River and were identified as full siblings. The number of estimated breeders in 2010 was 9 while in 2011 it was 74.

**Garza, J.C., S.M. Blankenship, C. Lemaire, and G. Charrier.** 2008. Genetic population structure of Chinook salmon (*Oncorhynchus tshawytscha*) in California's Central Valley. Final Report for CalFed Project "Comprehensive Evaluation of Population Structure and Diversity for Central Valley Chinook Salmon". University of California, Santa Cruz, Institute of Marine Sciences, and NOAA Southwest Fisheries Science Center, Santa Cruz, California. 80 p.

<https://swfsc.noaa.gov/publications/CR/2008/2008Garza4.pdf>

Genetic data from 20 microsatellite loci was used to compare population samples of adult Chinook salmon (N=2613) from 13 Central Valley locations, as well as the Klamath River as an outgroup. There was widespread population homogeneity within the fall-run which is evidence of substantial recent gene flow between all sampled fall-run populations in the Central Valley. There was subtle but significant differentiation between the fall-run and the late fall-run in the Upper Sacramento River and Battle Creek. There was also substantial differentiation of the Merced River Hatchery fall-run stock. The two primary remaining spring-run lineages (Butte Creek and Deer/Mill creeks) are monophyletic. The "spring-run" from the Feather River are genetically fall-run, although there is subtle, but significant, differentiation between the two groups at the hatchery. There was significant linkage disequilibrium in this population, but no others, supporting the hypothesis that it is a remnant of the ancestral Feather River spring-run that has been heavily introgressed with fall-run genes. There is clear genetic differentiation between spring-run and fall-run spawning in the same stream for Deer, Mill, and Butte creeks, and no evidence of gene

flow or hybridization between them. The data indicates that run timing is more important than geography for describing genetic structure in the Central Valley, with spring-run populations from different basins more closely related than are spring-run and fall-run populations from the same basin.

**Harvey, B. and C. Stroble.** 2013. Comparison of genetic versus Delta Model Length-at-Date race assignments for juvenile Chinook salmon at state and federal south Delta salvage facilities. California Department of Water Resources, Sacramento, California.

[http://www.water.ca.gov/aes/docs/1-ChinookGenetics\\_Final%20Report%20for\\_publication\\_2013-10-01\\_SERIF.pdf](http://www.water.ca.gov/aes/docs/1-ChinookGenetics_Final%20Report%20for_publication_2013-10-01_SERIF.pdf)

Summary of results for genetic identification of late-fall, fall, winter, and spring-run juvenile Chinook salmon salvaged at CVP and SWP pumping facilities are presented. Used genetic race assignments to evaluate the accuracy of assignments based on the Length-at-Date (Delta Model size criteria) approach. Nearly half of genetic assignments differed from corresponding Length-at-Date assignments. The Length-at-Date approach is most consistent for Winter-run. Post 2005 yearly genetic loss averaged 60% of Length-at-Date loss. Length-at-Date loss is most similar to genetic loss between mid-February and mid-March (genetic to Length-at-Date average 71%) which is also when the greatest density of genetic winter-run salvage occurs.

**Hedgecock, D.** 2002. Microsatellite DNA for the management and protection of California's Central Valley Chinook salmon (*Oncorhynchus tshawytscha*). Report prepared for the Department of Water Resources, Sacramento, CA.

[http://wwwdwr.water.ca.gov/iep/docs/Hedgecock\\_2002.pdf](http://wwwdwr.water.ca.gov/iep/docs/Hedgecock_2002.pdf)

Of the 6752 juvenile Chinook salmon genotyped over six years of salvage sampling, 711 were assigned to winter-run using the most accurate method. Of the fish assigned to winter-run, 95.5% were found within the winter growth bounds. However, 50.7% of the juvenile salmon within the predicted winter-run growth lines are not assignable to the winter-run (proportions are highly variable across years). In Delta monitoring samples, 14% of juvenile Chinook salmon captured assigned to winter-run (317 out of 2280). Feather River spring run Chinook salmon are distinct from Central Valley fall run (based on 12 microsatellite loci) but are genetically much closer to fall run than to either Deer-Mill Creek spring or Butte Creek spring.

**Hedgecock, D., M.A. Banks, V.K. Rashbrook, C.A. Dean, and S.M. Blankenship.** 2001.

Applications of population genetics to conservation of Chinook salmon diversity in the Central Valley. In: *Contributions to the Biology of Central Valley Salmonids* (Ed. By R.L. Brown). State of California Resources Agency Department of Fish and Game. Fishery Bulletin 179: 45-70.

Winter run, and to a lesser degree Butte Creek spring run, show lower levels of allelic diversity than other runs, suggesting that these populations experienced past reductions in size (bottlenecks) which may explain a part of their divergence from other runs. For juvenile Chinook salmon salvaged at the pumping facilities, mixed stock analysis establishes the prior probability for the runs and, using a Bayesian statistical approach, corrects the individual population assignment for unequal relative frequencies of sub-populations. Growth curves overestimate the losses of winter-run in the Delta. In addition, winter-run juveniles are caught at similar size throughout the emigration season in contrast to the growth curves.

**Hess, J.E., A.P. Matala, and S.R. Narum.** 2011. Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River Basin. *Molecular Ecology Resources* 11 (Suppl. 1): 137-149.

Performance of 13 microsatellites and 92 SNP loci was tested for fine-scale GSI of Chinook salmon across the Columbia River Basin. Microsatellites outperformed the SNPs in resolving fine-scale relationships but the combination of all 105 markers provided the greatest GSI power. Between 100 and 200 highly informative SNP loci are estimated to be needed to resolve stocks in finer-scale GSI applications (correct assignment > 90%).

**Larson, W.A., L.W. Seeb, M.V. Everett, R.K. Waples, W.D. Templin, and J.E. Seeb.** 2014.

Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications* 7: 355-369.

Identification of 10,944 SNPs was done using restriction-site-associated DNA (RAD) sequencing to examine population structure, demography, and adaptive divergence in five populations of Chinook salmon from western Alaska. The ability to assign individuals back to region of origin was greatly improved. Estimates of effective size were generally above 1000 and were biased downward when physically linked loci were not removed. Identification of 733 loci and three genomic regions under putative selection was accomplished using outlier tests based on genetic differentiation. The markers and genomic regions identified are good candidates for future research and can be used to create high-resolution panels for genetic monitoring and population assignment.

**Larson, W.A., J.E. Seeb, C.E. Pascal, W.D. Templin, and L.W. Seeb.** 2014. Single-nucleotide polymorphisms (SNPs) identified through genotyping-by-sequencing improve genetic stock identification of Chinook salmon (*Oncorhynchus tshawytscha*) from western Alaska. *Canadian Journal of Fisheries and Aquatic Sciences* 71: 698-708.

Genetic stock identification (GSI) can be difficult when populations are closely related. Genotyping was performed on 11,850 single-nucleotide polymorphisms (SNPs) from five closely related populations of Chinook salmon in western Alaska. A subset of these SNPs (96) displaying high differentiation were converted into high-throughput genotyping assays. These 96 new SNPs (RAD 96) and 191 previously developed SNPs (CTC191) were screened in 28 western Alaskan populations. Regional assignment power was evaluated for five different SNP panels with the resulting assignment tests indicating that SNPs in the RAD96 were more useful for GSI than those in CTC191. In addition, increasing the number of reporting groups in western Alaska from one to three was feasible with the  $F_{ST}$ 96. The method presented is an efficient way to discover SNPs for GSI.

**Pyper, B. T. Garrison, S. Cramer, P. Brandes, D.P. Jacobson, and M.A. Banks.** 2013. Absolute abundance estimates of juvenile spring-run and winter-run Chinook salmon at Chipps Island. Cramer Fish Sciences, U.S. Fish and Wildlife Service, and Oregon State University prepared for the Delta Stewardship Council, Sacramento, CA. 89 pp.

[http://deltacouncil.ca.gov/sites/default/files/documents/files/Final%20Chipps\\_DNA\\_Abundance\\_Report\\_7\\_2\\_13\\_2.pdf](http://deltacouncil.ca.gov/sites/default/files/documents/files/Final%20Chipps_DNA_Abundance_Report_7_2_13_2.pdf)

Juvenile Chinook salmon abundance by run was estimated at Chipps Island using trawl data. Run assignments were made based on genetic markers for trawl samples collected from October, 2007 to June, 2011. DNA assignments indicated that during the study period fall run composed between 84.0% and 92.8 % of the annual juvenile abundance, late-fall run composed 1.9% to 4.4%, and Butte Creek spring run composed 3.9% and 9.0%. Mill-Deer creek spring run and winter run each composed less than 3% of the total abundance during the study period. Estimates of DNA assignments were highly uncertain for late-fall run and Mill-Deer spring run. Run composition based on DNA often were substantially different from those based on length-at-date criteria. More fish were assigned to fall run and fewer assigned to spring run and winter runs based on DNA compared to length criteria. Winter-run DNA assignments fit the closest to their expected length-at-date range. However, relatively large numbers of fall, late-fall, and spring run fish overlapped with the length criteria for winter-run. Large numbers of fall run (based on DNA) overlapped with the length criteria for spring-run. DNA assignments provided more accurate and reduced annual estimates of, run composition for the spring and winter runs (one half to one sixth of run compositions based on length criteria). The ranges in annual abundances from 2008 to 2011 (based on the midrange of empirical efficiency estimates) were: 1.4 million to 7.5 million for fall run, 71 thousand to 186 thousand for late-fall run, 67 thousand to 331 thousand for Butte Creek spring run, 36 thousand to 92 thousand for Mill-Deer creek spring run, and 45 thousand to 63 thousand for winter run. Fall run abundance estimates were the most precise, with abundance estimates for Butte Creek spring run and winter run being relatively precise. Abundance estimates for late-fall run and Mill-Deer creek spring run were very imprecise.

**Rawding, D.J., C.S. Sharpe, and S.M. Blankenship.** 2014. Genetic-based estimates of adult Chinook salmon spawner abundance from carcass surveys and juvenile out-migrant traps. *Transactions of the American Fisheries Society* 143: 55-67.

Genetic approaches to estimate the spawner abundance of a population of Chinook salmon was investigated using genetic mark-recapture and rarefaction curves. The marks were the genotyped carcasses collected from the spawning area during the carcass survey. The parents that assigned to the juveniles (collected from a downstream migrant trap below the spawning area) through parentage analysis were considered the recaptures, which was a subset of the genotypes captured. The Petersen estimator was used with the binomial and hypergeometric models to estimate the genetic mark-recapture spawner abundance. The results were in agreement with spawner abundance estimates based on redd counts, area-under-the-curve methods, and carcass tagging based on the Jolly-Seber model. Using a rarefaction curve approach (based only on the juvenile offspring sample) the spawner abundance estimate was similar, although lower, than the genetic mark-recapture approach.

**Steele, C.A., E.C. Anderson, M.W. Ackerman, M.A. Hess, N.R. Campbell, S.R. Narum, and M.R. Campbell.** 2013. A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. *Canadian Journal of Fisheries and Aquatic Sciences* 70: 1046-1054.

It was empirically confirmed that fewer than 100 SNPs were needed to accurately conduct parentage-based tagging. Comparable accuracy was demonstrated between the selected panel of SNPs and a panel of microsatellites. Stock assignments made with the SNPs panel matched those made using CWTs. The study also demonstrated that an estimated PBT rate for the offspring could be predicted with fewer than 100 SNPs when sampling of spawners was incomplete.