

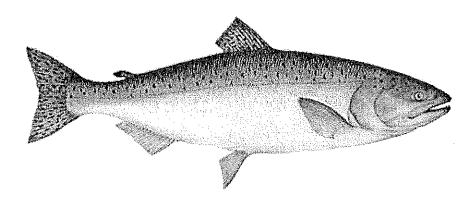
Hatfield Marine
Science Center
2030 SE Marine Science
Drive
Newport, Oregon
97365





# MOLECULAR GENETIC TOOLS FOR DISCRIMINATION AMONG RUNS OF THE CALIFORNIA'S ENDANGERED CENTRAL VALLEY CHINOOK SALMON (Oncorhynchus

tshawytsha)



Chinook Salmon

Researcher: Christian Andresen

**Mentors:** Dr. Isabelle Meusnier Dr. Michael Banks

Research Experience for Undergraduates Program

Marine Genetics Laboratory

## INTRODUCTION

Chinook or king salmon, Oncorhynchus tshawytsha is the largest of five species that inhabit the eastern Pacific coast, reaching lengths of 150 cm and weights over 15 kilos. They are the least abundant among Oncorhynchus across the whole Pacific Northwest, but the most numerous in southern latitudes of United States. Oncorhynchus species are anadromous, meaning that they migrate from the sea to fresh water to spawn. Thus, they face challenges both in the ocean and river systems as a result of environment degradation and over harvesting. This project focuses on the California's central valley Chinook salmon, which are characterized by different runs named after the time when adults return to fresh water for spawning; spring, fall, late fall, and winter (see figure 1) (Fisher, 1994). Recent genetic characterization using microsatellites supports these subpopulations, and has also shown that spring run is comprised of two discrete sub-populations; Butte Creek spring run is distinct from Mill and Deer Creeks spring run (Banks et al. 2000). The Sacramento River winter-run Chinook was the first anadromous salmon run listed under the Endangered Species Act in 1989. Also, the central valley Chinook spring runs were declared as threatened by the state of California Department of fish and Game in February of 1999, and eight months later, declared threatened by the Endangered Species Act (California Dep. 2004).

Chinook salmon have survived enormous human environmental impacts, from the Gold Rush in the mid 1800's, which greatly contaminated river systems with high levels of heavy metals, to the rise of hydropower and the construction of dams and water diversion systems across the Sacramento and San Joaquin rivers (which conform the California's Central Valley river system). Also, enormous water demands of agriculture and California's population have made a significant impact on all salmon species.

Management of this species is complicated because endangered populations coexist with stable populations in the river and ocean, where they are impacted by various factors such as water diversion and ocean fisheries affecting their survival (Banks, 2005). Moreover, it is impossible to assign fish to a particular run using morphological features. Fortunately, heritable genetic markers can provide an economic, practical, and powerful solution for distinguishing different stocks because of the accuracy of genetic characterization. This can only be applied if sufficient reproductive isolation exists among communities, and characterization of accumulated genetic differences is possible among populations.

ί.

A study conducted by Dr. Michael Banks using microsatellites, simple sequence repeats. demonstrated high power to discriminate winter run. However, microsatellites had less power to discriminate among spring, fall, and late fall runs. Thus, to increase discriminatory power we chose to focus on molecular genetic analysis of two circadian genes that are associated with physiology and behavior. These two circadian genes: Cryptochrome (CRY) and Brain and Muscle ARNT-like (BMAL) genes are linked to the timing and regulation of fish's development and daily rhythms. We chose circadian genes because reproductive timing in salmon is believed to be under circannual control (occurring in yearly periods), and the temporal isolation of Chinook spawning runs must be regulated by an internal timing mechanism such as circadian genes. The selected circadian genes, CRY and BMAL, are an important regulatory mechanism of maturation and migration timing in Chinook, which should allow us to distinguish between different runs. In addition, recent research by Dr. Isabelle Meusnier, found one microsatellite on the circadian genes that discriminated the winter run (Meusnier Personal Com.). Therefore, the purpose of this project is to continue to resolve genetic differences among alternate life history types in Chinook salmon by analyzing the intra and interindividual genetic variation on California's Central Valley individuals from five different runs (Butte Creek, and Deer and Mills Creeks in the spring, Fall, Late-Fall, and Winter). This would allow us to

discriminate different stocks, so that conservation and resource management strategies can be focused on particular runs of Chinook salmon.

# **METHODS**

Samples Collection and Gene Source

The DNA collection was comprised of samples collected by volunteers and the California Department of Fish and Game between the years 2003 and 2004, and information of collection location, date, and sex was recorded (see figure 1). Tissue was obtained from fin clips and liver. Samples were preserved in 95 % ethanol. Genomic DNA was extracted through DNEasy (Qiagen Inc.) or Chelex/proteinase K (Estoup *et al*, 1996) protocols.

Dr. Isabelle Meusnier had found and sequenced 5 CRY genes and four BMAL genes. Primers were designed to amplify the five Cry gene copies obtained from Dr. Meusnier in order to sequence the gene of individuals of each run in smaller fragments (± 1000 bp) (see appendix 1).

PCR conditions

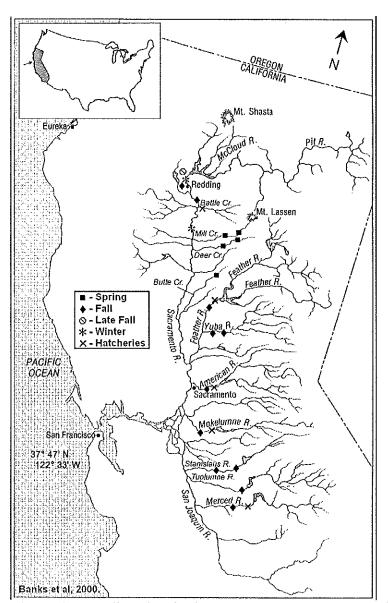


Figure 1. Sampling sites in the Sacramento and San Joaquin rivers and their tributaries

On the PCR (Polymerase Chain Reaction) optimization of primers, amplifications were performed in 15 µl volumes on a MJ Research thermocycler containing: 1X PCR buffer (*Promega*), 25 mM MgCl<sub>2</sub>, 10mM of dNTPs, 10mM of Primers, and 5 U/µl of *Promega* Taq polymerase. Amplification was accomplished with an initial denaturation at 94° C for 3minutes, then 30 cycles of the following steps; 94° C, 30 sec; gradient PCR 45° C to 60° C, 1 min; 72° C, 2 min; then, a final elongation of 3 minutes. Of these 15 µl, 8 µl of amplified products were run in a 1.5% agarose gel (see appendix for mix and PCR program details). The PCR volume for sequencing was 60 µl with the same mix as the optimization mix, but different annealing temperature (see appendix 2). Eight microliters were used to test the products in a 1.5% agarose gel.

#### PCR Products Cleaning

After testing all the primers at different temperatures, the ones that successfully bound the wanted fragment of the CRY gene were tested with more Chinook individuals of each run (see appendix 2). Then, the samples were cleaned using an Ethanol precipitation Method (See appendix for method details). The samples that resulted in multibands were loaded in a gel (electrophoresis at 70 volts) and then cut and cleaned following the QIAquick Gel Extraction Kit Protocol (see appendix 2).

#### Locus Variation Analysis

The CRY gene fragments were sequenced on the ABI 3730XL Genotyper machine, and exported and aligned using *BioEdit* v7.0.0 software, apparent variation observed among runs was revised with the Chromatogram and then corrected. Also, MICAS (Microsatellite Analysis Server) Software (Sreenu, *et al.*, 2004) was used to find loci with microsatellites in CRY and BMAL genes sequences. Next, new primers were designed utilizing fluorescently labeled M13-Tailed modify primer (5'-GGATAATTTCACACAGGTCCGCTTAT-3', Boutin, *et al.* 2001) to bind the fragments of variation, and more DNA samples of each run were tested (see appendix 7). These

smaller fragments were separated via capillary electrophoresis on a ABI Genotyper and analyzed with GeneMapper Software version 3.7, because mutations among individuals are only characterized by capillary electrophoresis due to their small fragment size. Candidate loci were tested using the computer application WHICHLOCI, a computer program for determining relative discriminatory power among candidate genetic loci (Banks et al, 2003).

## RESULTS

Around half of the initial primers sets designed bond successfully (see appendix 2). Variation like single nucleotide polymorphisms (SNPs), insertion/deletion polymorphisms, and microsatellites were found when the sequences of one individual of each run were aligned in BioEdit (see appendix 3). However, when this variation was sequenced in six individuals of each run, there was not consistent variation among individuals of the same run. These results show significant evidence of genetic variation among individuals of the same run on the CRY genes.

As a result of inconsistent variation among sequence data for run individuals, we focused on characterizing microsatellites and minisatellites. From within CRY and BMAL genes sequences and analyzed them with MICAS software (Sreenu, 2004) and BioEdit (see appendix 3). The results were as follows: three minisatellites, and five microsatellites on the CRY gene copies sequence, and one minisatellite and two microsatellites on the BMAL gene copies sequence. We designed M13-tailed modify fluorescent primers to bind these fragments and determine their readability and usability in microsatellite analysis (see appendix 6). After sequencing the M13-tailed modify primers that bond the specific fragments with variation, we decided to focused on three loci that showed clear variation when ran in an electrophoresis gel (see appendix 4).

The variation found (microsatellites) was tested on 96 individuals of each run from the base line samples via capillary electrophoresis and analyzed in order to verify its power. Results for candidate locus was tested using the computer application WHICHLOCI.

The locus Cry-TB-1m showed five alleles from 172 to 184 bp, but no significant variation among individuals of different runs. BMAL-TD-1m showed ten alleles from 309 to 326 bp, but no stable variation among the runs. And Cry-THP-7m locus showed significant variation of Fall-run from the rest of the runs. Cry-THP-7m locus had the ability to best assign individuals to Fall run with 15 alleles ranging from 363 to 411 bp.

It is important to mention that due to time reasons, we where only able to discriminate one of the five runs, considering that the winter run has been characterized already using circadian genes by Dr. Meusnier, and there are more loci to be tested, and further characterization for the rest of the runs (Butte Creek, and Deer and Mills Creeks in the spring, and Fall).

# CONCLUSIONS AND FUTURE PROSPECTS

Throughout this study, Chinook salmon circadian genes have demonstrated complex structure and diversity. This genetic diversity was our first obstacle in variation analysis. However, we were able to identifyrun specific differences among this variation. Nevertheless, this evidence supports that California's central valley Chinook populations are genetically diverse; a fact that promotes evolutionary adaptation to environmental changes, underlying a positive future for Chinook populations. It is important to understand that this diversity we characterized is expressed in the circadian genes, and not necessarily phenotypic variation among the runs.

Our data suggests that circadian genes are linked with life history of Chinook salmon and we have been able to discriminate successfully Fall run with one microsatellite locus. In addition, these study supports that reproductive isolation exists among populations and characterization of accumulated genetic differences of circadian genes is feasible among two (Winter and Fall) of the five Chinook runs of the Sacramento and San Joaquin river basin. This finding represents the new crafted tool for accurate stock identification for Fall run.

Moreover, this molecular genetics advancement not only will help management and protection of California's Chinook salmon, but will set a new window for discriminatory tools for population genetics on *Oncorhynchus*, and other andronomous species populations that have reproductive timing differences. Therefore, it is important to continue to understand the Chinook genome because significant answers like this one are waiting to be found.

## REFERENCES

- Banks, M.A., Rashbrook, V.K., Calavetta, M.J., Dean, C.A. and Hedgecock, D. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of chinook salmon in California's Central Valley. *Canadian Journal of Fisheries and Aquatic Sciences*. 57:915-927.
- Banks, M. A., W. Eichert, and J. B. Olsen. 2003. Which genetic loci have greater population assignment power? *Bioinformatics*. **19**:1436–1438.
- Banks, M. A. 2005. Stock Identification for Conservation of Threatened or Endangered Species. *Stock Identifications Methods*. **29**: 609-629.
- Boutin, I., Raposo, M., Raymond, M. and Deschepper, C. 2001. M13-tailed Primers Improve the Readeability and Usability of Microsatellite Analyses Performed with Two Different Allelesizing Methods. *BioTechniques*. 31:24-28
- National Marine Fisheries Service (NMFS). 1994. Endangered and threatened species; status of Sacramento River winter-run chinook salmon. Fed. Register **59**(No. 2): 440–450.
- Estoup, A., Presa, P., Kreig, F., Vaiman, D., Guyomard, R. 1993. (CT)n and (GT)n Microsatellites: a new class of genetic markers for *salmo trutta*. *Heredity*, 71: 488-496
- Fisher, F. W., 1994. Past and present status of central valley Chinook salmon. Conserv. Biol. 8: 870–873.
- Sreenu, V.B., Ranjitkumar, G., Swaminathan, S., Priya, S., Bose, B., Narendra Pavan, M., Nagaraju, J., Nagarajaram, H.A. 2004. MICAS Database of Prokaryotic Microsatellites. *Nucleic Acids Research*. Web accessed: July 2006. http://210.212.212.7/MIC/index.html

# **Appendix**

CRY copy	Primers	Length bp	Forward	Reverse
TA	1	+ 1249	5' - CGCTAGTTTGAGCCATCAAT- 3'	5' - GTCTTCAACTCTACCATGACATTAT
TA	2	1374	5' - TTTCAGTCCCATTTGACAGT- 3'	5' - CACTAGATCACTACAACTAACCACT
TA	3	+1264	5' - TCAAGAAGTGAATTGGAAGTCC- 3'	5' - CCTGACACCATGTATGATTGAA - 3'
TA	4	1350	5' - ATGGTCAGAGGGCATCAGTC- 3'	5' - GGAGGAGTACCTGTGGATGT - 3'
TA	5	1391	5' - GGCACTGAGTTTGAAGGTGG- 3'	5' - TGAGGTGAGGACAGAGAGAG - 3'
TA	6	1382	5' - ATCACGCAAGGTTGTTTCTA- 3'	5' - GAGCACTACAAAGAGAGGGA - 3'
TA	7a	600	5' - GTTGTGACGATGCTCTGTTT- 3'	5' - AGGAAAGGAGAGTGGGAGAG - 3'
TA	7b	707	5' - GTTGTGACGATGCTCTGTTT- 3'	5' - TCCAGGTCGTGGTCTAACTC - 3'
TA	8	+ 1300	5' - GTGAGCACACTCTAACTCCTC- 3'	5' - ACACATTCACCCCTGACTTA - 3'
TA	9	411	5' - AATAAACAGGTAAGTCAGGG- 3'	5' - ACAGAACAAAGGACACATTA - 3'
TE	1a	907	5' - CAGTTATGCAACTGTCTTGTTC- 3'	5' - CTCAGTCTCCAATTAGTTTGAC - 3'
7.6	1b	1200	5' - CACCACAAAACTACGACTGTT- 3'	5' - GTACTTTAGGGAATGGGTCA - 3'
TE	2	1144	5' - GTCAAACTAATTGGAGACTGAG- 3'	5' - CGCAAAGAGACTCATGTTTA - 3'
ΤE	3	1159	5' - AGTAGTAAACATGAGTCTCTTTGCG- 3'	5' - AACGCTCGCTCACCATAAAA - 3'
116	4	802	5' - CTGGGTAAGACAACTTTTCTATTGG- 3'	5' - TGAGGAAAACAGGAGCAATC - 3'
TE S	5	1391	5' - ACAAGAAGGTACAGTATCTCACAC- 3'	5' - AAGAGAGTGGGCAGGTGATA - 3'
TÉ.	6	901	5' - CTCACTTTGTATTGAACTCTCTCTC- 3'	5' - ATGCTTTGTGACATCACCCC - 3'
TE	7	1398	5' - AACCCACACTCCACCTCTCC- 3'	5' - AAATAAACATGGTGAAAGAACAAC -
T/B	1	864	5' - AAATGGTATTTTCCAGAACCCC-3'	5' - TGCAACACACTGACAAGGAG - 3'
7/8	2	1258	5' - CTGGAGAGAAAGGTGAGGAG- 3'	5' - CTGGAAGAACCAAACAGAGA - 3'
7(8)	3	1265	5' - GACTACATAAGGTGAGAAGACCC- 3'	5' - CATGTCCTCTGCCTGAATAA - 3'
TB.	4	1437	5' - TATTCAGGCAGAGGACATGC- 3'	5' - TCAAGGTCTCAGTCCCCTCG - 3'
TCD	1	1185	5' - TGTCAGGTTGGATGGGGAGC- 3'	5' - CACATGGACAGAGATAGCGG - 3'
TCD	2	730	5' - AGTGAGTGAATCTTACAGATACACG- 3'	5' - ACGCCTGGAACATAGAGGGT - 3'
TCD	3	1208	5' - AGAGAAAGGTGGGAAGGAGT- 3'	5' - CAAATGGAGAAAAGACCGTT - 3'
TCD	4	+ 934	5' - CACTTAGTTGTCCACTGCTAAT- 3'	5' - GATTGTTACCTGCTTGAGCG - 3'
TCD	5	1328	5' - TCTCCATTTCACAGGGCTGC- 3'	5' - CACTGCTCCAAAACCTCCAT - 3'
TCD	6	+ 900	5' - GTTTCATCAGACCAGAGGAC- 3'	5' - ACAGAGTCAGGGAAGGAGAA - 3'
TCD	7	+?846	5' - AAACTCTCCAGCAGAATTGA- 3'	5' - GTCAGTCTTCACACAAAATAAC - 3'
THP	1	782	5' - GGAGGAGGCAAGATCAGATG- 3'	5' - TCACAACCATGATTCCAAGC - 3'
THP	2	+376	5' - TTGGAATCATGGTTGTGAAT- 3'	5' - CGATCCTGATAACAGTAGAAAC - 3'
THP	3	932	5' - GTTTCTACTGTTATCAGGATCG- 3'	5' - AGCATCGGTTAGTGGTGGTA - 3'
THP	4	+1002	5' - CCACACTATCTACCTAGAGTTCTC- 3'	5' - AAATAGCCCTTACAAAGCCT - 3'
THP	5	+1215	5' - CACACTTAACCAGCATGGCA- 3'	5' - TCTTCCGCAACTGAGTGGGA - 3'
THP	6	1362	5' - ACAACTGCTGATGTAAGAAGTG- 3'	5' - GTGTGTTGACAGTCTGTATTCAGT -
THP	7	1373	5' - TGGATAAACCACAGCCTGTC- 3'	5' - TAACCCAAAGGTAAGTGCTT - 3'
THP	8	+989	5' - TTGCCCAGCACAAATCCACT- 3'	5' - CCTTGTCTTCCAGGTCGGTT - 3'
THP	9	856	5' - CAACAACGCACTCCTCCTGT-3'	5' - TGAACCAACTAAGCAGCAAC - 3'
THP	10	839	5' - GCACCTGTGGTATGTTTGGA-3'	5' - AGTGTTCCTGAAGCAAGCGG - 3'
THP	11	1173	5' - TTCAATGCTCTTATTTGGGG - 3'	5' - CTGATCTGGAGTTGCTGTTATTCTT

Appendix 1. Initial Primer sequences used to try to amplify California's Central Valley Chinook salmon Cry gene.

Primers selected for Sequencing					
Loci	ıs	Length bp	Aneling Temp		
TA	7b	707	45-60		
TA	9	411	50.9		
TE	1a	907	54		
-TE	3	1159	60		
TE	6	901	58		
TCD	2	730	58		
TCD	3	1208	51		
TCD	4	934	58		
TCD	7	846	58		
THP	1	782	58		
THP	3	932	58		
THP	9	856	58		
THP	10	839	58		
THP	11	1173	58		
To Cu	t (Mu	ıltibanded)			
TB	2	1258	54		
ТВ	3	1265	60		
TCD	1	1185	58		
TCD	6	900	60		
THP	5	1215	45		

Appendix 2. Length and annealing
temperature of the primers that
successfully bond.

	Type of			
Locus	Variation	Location		
		A T-001		
Cry_TE_P6	SNPs	91, 434, 808 bp		
Cry_TE_P3	Microsatellite			
Cry_THpal_P1	SNPs	190, 235, 435, 666 bp		
Cry_TA_P7b	ins/del			
Cry_THpal_P3	SNPs	700, 703 bp		
Cry_THpal_P5	SNPs	124 bp		
Cry_TB_P1	Microsatellite			
Cry_TCD_P4	SNPs	600, 820, 1219 bp		
Cry_TCD_P7	SNPs	703 bp		

Appendix 3. Variation found when the sequence of one individual of each run was aligned.

CRY Locus		Length bp	Aneling T°C	Repeat Motif	Alleles per locus
ТВ	1m	136	54 - 60	microsat "CA"	5
THP	7m	378	49 - 54	microsat "CA" impure	15
BMAL Locus					
TD	1m	300	54 - 60	microsat "CA"	10

Appendix 4. Candidate locus for discrimination among

### Appendix 5. Ethanol precipitation method for purifying PCR products

- \* Ethanol precipitation method for purifying PCR products \*
- 1. For each PCR reaction (50 µL) prepare a 1.5mL microcentrifuge tube (label them) containing the following:
- 100 μL of 95% ethanol (EtOH)
- 5 μL of 3M sodium acetate (NaOAc), pH 4.6
- 2. Pipet (use filter tips) the entire contents of each PCR reaction into a tube of sodium acetate/ethanol mixture. Mix throughly.
- 3. Vortex the tubes and leave at -20°C (in the freezer) for 30 to 40 min to precipitate the PCR products.
- 4. Spin the tubes in a microcentrifuge for 20 minutes at maximum speed (make sure that the tubes are set in the centrifuge in a way that you will know which part of the tube will have the pellet precipitated). Then take them out carefully trying not to disrupt the pellet.
- 5. Carefully aspirate the supernatant with a regular tip and discard (use 200 or 300  $\mu L$  pipet).
- 6. Rinse the pellet with 300µL of 70% ethanol.
- 7. Tap gently the samples.
- 8. Spin for 5 minutes in a microcentrifuge at maximum speed (remember to set them in position to know where the pellet is). Again, carefully aspirate or decant the supernatant and discard. The pellet does not stick very well to the wall of the tube so be careful not to lose it!
- 9. Dry the pellet for 20 minutes. Do not over-dry. Resuspend in 30 µL of sterile water, tap gently and let it sit for 1 hr.

Appendix 7. PCR mixes and programs for optimization, sequencing and Tailed primers

# MIX for PCR optimization

Reactive	Initi	al conc	final conc		15	μl Rxn Vol
H₂O		Mar Ann Ann Inn and				9.025
Buffer 10X	10	Х	1	Х		1.500
MgCl₂	25	mM	1.5	mM		0.900
dNTPs	10	mM	0.2	mM		0.300
primer F	10	μΜ	0.4	μM		0.600
primer R	10	μM	0.4	μΜ		0.600
Taq	5	U/µI	0.025	U/µI		0.075
Total Volume						13.000
					2	ul DNA

### PCR Program:

T°C	Time	
94	3'	
94	30"	
45-60	1'	X 3
72	2'	
72	7'	
25	1'	

# PCR MIX for sequencing

Reactive	Initi	al conc	nc final conc		60	μl Rxn Vol
H <sub>2</sub> O						38.100
Buffer 10X	10	X	1	Х		6.000
MgCl₂	25	mM	1.5	mΜ		3.600
dNTPs	10	mM	0.2	mМ		1.200
primer F	10	μΜ	0.4	μM		2.400
primer R	10	μM	0.4	μM		2.400
Taq	5	U/µI	0.025	U/µI		0.300
Total Volume						54.000
					6	μί DNA

### PCR Program:

T°C	Time	
94	3'	
94	30"	
Anealing °C	1'	X 30
72	2'	
72	7'	
25	1'	
	<u></u>	

# PCR MIX for tailed primers

Reactive	In	itial conc	fina	al conc	
H <sub>2</sub> O					
Buffer 10X	10	Х	1	Х	
MgCl₂	25	mM	1.5	mM	
dNTPs	10	mM	0.2	mΜ	
primer F	10	μM	0.4	μM	
primer Rt	10	μΜ	0.08	μM	
primer M13Rmod	10	μM	0.4	μM	
Taq	5	U/µI	0.025	U/µI	

### PCR Program:

T°C	Time	
94	3'	
94 Anealing	30"	
T°C	1'	X 40
72	1'	
		7
72	30'	
		7
25	1'	]