Update Report

Period 2/1/2013 - 1/31/2014

Project R/LME-1 - Linking genetic variation, selection and adaptation in Chinook salmon next generation sequencing and an oligonucleotide microarray with complete genome coverage

STUDENTS SUPPORTED

Brieuc, Marine, mbrieuc@uw.edu, University of Washington, School of Aquatic and Fishery Sciences, status cont, field of study Genetics, advisor Kerry Naish, degree type PhD, degree date 2013-06-01, degree completed this period Yes
Student Project Title Locating the genomic regions underlying adaptive divergence in Chinook salmon life history traits

Involvement with Sea Grant This Period grad student

Post-Graduation Plans Postdoctoral researcher

Larson, Wes, wlarson1988@gmail.com, University of Washington, Aquatic and Fishery Sciences, status new, no field of study, no advisor, degree type PhD, degree date 2015-06-01, degree completed this period No
Student Project Title none

Involvement with Sea Grant This Period Assisted with Task 1 sequencing and applied Sea Grant derived map

Post-Graduation Plans none

CONFERENCES / PRESENTATIONS

Discover Science Weekend - Seattle Aquarium, public/profession presentation, 400 attendees, 2013-11-09
Naish KA (2013) The allure of more, better, faster four stories of genomic diversity, selection, and fitness in Pacific salmon. Department seminar, School of Aquatic and Fishery Sciences, Seattle, October 03, 2013, public/profession presentation, 100 attendees, 2013-10-03
Brieuc MSO, CD Waters, M Kodama, JE Seeb, KA. Naish (2014) A dense linkage map for Chinook salmon (Oncorhynchus tshawytscha) and comparative mapping reveal variable chromosomal divergence in salmonids following an ancestral whole genome duplication event. Oral Presentation by MSO Brieuc at the Plant and Animal Genome XXII conference, San Diego, California, January 11, 2014, public/profession presentation, 100 attendees, 2014-01-11
Seeb JE (2013) Genotyping by sequencing in Chinook salmon. Invited keynote, ConGen 2013 Population Genomic Data Analysis, 3-8 September, 2013., public/profession presentation, 45 attendees, 2013-09-03

Seeb JE (2013) RAD sequencing in non-model organisms. Invited departmental seminar, Hatfield Marine Science Center, March 8, 2013., public/profession presentation, 60 attendees, 2013-03-08

ADDITIONAL METRICS

K-12 Students Reached Marine Brieuc and Kerry Naish participated at the Discover Science Weekend at the Seattle Aquarium, November 9 2013.	400	Acres of degraded ecosystems restored as a result of Sea Grant activities	0
Curricula Developed Jim Seeb Worked with 14-instructor team to mentor 31 international graduate students at ConGen 2013 at University of Montana.	1	Resource Managers who use Ecosystem-Based Approaches to Management	0
Volunteer Hours	0	HACCP - Number of people with new certifications	0
Cumulative Clean Marina Program - certifications	0		

PATENTS AND ECONOMIC BENEFITS

Description		Patents	Economic Benefit ($)	Businesses Created	Businesses Retained	Jobs Created	Jobs Retained
The economic outputs of the maps of the Chinook genome are not directly measureable, but its	Actual (2/1/2013 - 1/31/2014)	0	0	0	0	0	0
	Anticipated (2/1/2014 - 1/31/2015)	0	0	0	0	0	0

importance cannot be overstated. The genes on the map will be used by Washington Department of Fish and Wildlife to help prioritize populations for conservation and to develop markers for stock identificatio n.

TOOLS, TECH, AND INFORMATION SERVICES

Description		Developed	Used	Names of Managers	Number of Managers
Use of high-throughput genotyping by sequencing (GBS) to detect and map thousands of SNPs in haploid Chinook salmon. R/LME-1	Actual (2/1/2013 - 1/31/2014)	1	1		0
	Anticipated (2/1/2014 - 1/31/2015)	0	0		
Chinook genome map to enhance genetic improvement of aquaculture populations	Actual (2/1/2013 - 1/31/2014)	0	1	Alaska Departmen t of Fish and Game	1
	Anticipated (2/1/2014 - 1/31/2015)	0	0		

and serve as a framework for other species genome maps, noteably coho.
R/LME-1

HAZARD RESILIENCE IN COASTAL COMMUNITIES

No Communities Reported This Period

ADDITIONAL MEASURES

Safe and sustainable seafood

Number of stakeholders modifying practices
Actual (2/1/2013 - 1/31/2014)
Anticipated (2/1/2014 - 1/31/2015)

Number of fishers using new techniques
Actual (2/1/2013 - 1/31/2014)
Anticipated (2/1/2014 - 1/31/2015)

Sustainable Coastal Development
Actual (2/1/2013 - 1/31/2014)
Anticipated (2/1/2014 - 1/31/2015)

Coastal Ecosystems
Actual (2/1/2013 - 1/31/2014)
Anticipated (2/1/2014 - 1/31/2015)

PARTNERS

Partner Name Alaska Department of Fish and Game, type Government, scale Regional
Partner Name Alaska Sustainable Salmon Fund, type Other, scale Regional
Partner Name Centre for Aquatic Biotechnology Regulatory Research (CABRR), Centre for Aquaculture and Environmental Research, Fisheries and Oceans Canada (DFO)
Partner Name Department of Molecular Biology and Biochemistry, Simon Fraser University (SFU)
Partner Name University of Montana
Partner Name Washington State Department of Fish and Wildlife

IMPACTS AND ACCOMPLISHMENTS

Title Washington Sea Grant research develops a genetic toolkit to survey the Chinook salmon genome and strengthen conservation

Type impact

Relevance, Response, Results Relevance Nine populations of Chinook salmon are currently protected under the Endangered Species Act along the West Coast, and five are found in Washington waters. Genetic tools can provide important information for the conservation,

recovery and protection of this iconic species, making it possible to track catches of protected fish in fisheries where healthy and depleted populations comingle; guide hatchery management to protect population diversity and fitness; survey the Chinook genome for adaptations that support recovery; and set more effective conservation priorities. Response Washington Sea Grant-supported research has led to the mapping of a significant fraction of the Chinook salmon genome. Results The maps have significantly increased understanding of genetic variation within Chinook salmon, identifying thousands of gene loci, including regions associated with thermal tolerance and growth, two qualities of particular interest to managers. The research showed that the salmon genome underwent a recent duplication and has twice as many chromosomes as an ancestral species did. The genetic markers identified can be developed for a range of applications. Resource agencies are already using the maps to identify stocks, calculate effective population sizes, and identify the parts of the genome responsible for adaptive differentiation.

Recap Washington Sea Grant-supported research mapped a significant portion of the Chinook salmon genome, identifying markers for key survival and adaptation factors and providing powerful tools for protection and recovery of this important species.

Comments Primary Focus Area LME (SSSS) Secondary Focus Area LME (HCE) State Goals Support conservation and sustainable use of living marine resources through effective and responsible approaches, tools, models and information for harvesting wild and cultured stocks and preserving protected species (SSSS Supply). Support conservation and sustainable use of living marine resources through effective and responsible approaches, tools, models and information for harvesting wild and cultured stocks and preserving protected species (HCE Science).

Related Partners

PUBLICATIONS

Title A dense linkage map for Chinook salmon (Oncorhynchus tshawytscha) reveals variable chromosomal divergence after an ancestral whole genome duplication event.

Type Reprints from Peer-Reviewed Journals, Books, Proceedings and Other Documents Publication Year 2014 Uploaded File none URL http //www.g3journal.org/content/4/3/447.full.pdf+html

Abstract Comparisons between the genomes of salmon species reveal that they underwent extensive chromosomal rearrangements following whole genome duplication that occurred in their lineage 58 to 63 million years ago. Extant salmonids are diploid, but occasional pairing between homeologous chromosomes exists in males. The consequences of re-diploidization can be characterized by mapping the position of duplicated loci in such species. Linkage maps are also a valuable tool for genome-wide applications such as genome-wide association studies, quantitative trait loci mapping or genome scans. Here, we investigated chromosomal evolution in Chinook salmon (Oncorhynchus tshawytscha) following genome duplication by mapping 7146 Restriction-site Associated DNA (RAD) loci in gynogenetic haploid, gynogenetic diploid and diploid crosses. In the process we developed a reference database of

RAD loci for Chinook salmon comprising 48528 non-duplicated loci and 6409 known duplicated loci, which will facilitate locus identification and data sharing. We created a very dense linkage map anchored to all 34 chromosomes for the species and all arms were identified through centromere mapping. The map positions of 799 duplicated loci revealed that homeologous pairs have diverged at different rates following whole genome duplication, and that degree of differentiation along arms was variable. Many of the homeologous pairs with high numbers of duplicated markers appear conserved with other salmon species, suggesting that retention of conserved homeologous pairing in some arms preceded species divergence. As chromosome arms are highly conserved across species, the major resources developed for Chinook salmon in this study are also relevant for other related species.

Citation Brieuc, M. S. O., C. D. Waters, J. E. Seeb, and K. Naish. 2014. A dense linkage map for Chinook salmon (Oncorhynchus tshawytscha) reveals variable chromosomal divergence after an ancestral whole gdenome duplication event Genes Genomes Genetics 4 447-460.

Copyright Restrictions + Other Notes

Journal Title Genes Genomes Genetics

OTHER DOCUMENTS

No Documents Reported This Period

LEVERAGED FUNDS

Type influenced Period 2013-01-01 2013-11-01Amount $300000

Purpose RAD sequencing on Alaska Chinook salmon

Source Alaska Sustainable Salmon Fund

UPDATE NARRATIVE

Uploaded File Naish_5225_update_narr....3.pdf

Narrative (must be submitted online as a separate document, 2-3 pages)

R/LME-1 Progress Report 2/1/2013-1/31/2014

Linking genetic variation, selection and adaptation in Chinook salmon: next generation sequencing and an oligonucleotide microarray with complete genome coverage,

Kerry Naish & Jim Seeb, University of Washington, School of Aquatic and Fishery Sciences

1. Activities carried out

The aim of our study is to develop molecular markers that span the genome of Chinook salmon so that these markers can be used in standardized genome-wide population surveys. This aim will be attained by sequencing DNA markers spaced throughout the genome in individuals that represent the range of Chinook salmon (Task 1), identifying polymorphic sites (Task 2), creating assays for these markers (Task 3) and mapping these markers to the Chinook map (Task 4). Our activities to date for each of these tasks are as follows:

Task 1: We used RAD sequencing, a targeted approach that identifies thousands of DNA loci simultaneously, to sequence 11 populations from the Columbia Basin and 3 from Puget Sound. Five populations from Alaska, important for this effort, were sequenced on leveraged funds from the Alaska Sustainable Salmon Fund.

Task 2: Bioinformatic approaches for identifying polymorphic DNA sites across the Chinook genome were developed and implemented. Reference databases comprising all sequences from Task 1 were created and checked for quality and consistency. This reference database is a key resource for aligning all sequences generated by future projects, and has been made publicly available.

Task 3 We provided transcriptome sequence to the University of Montana who provided a 10,000 feature array for our testing. Tests are still underway, although the Montana approach appears expensive and uncertain. We continue the tests in 2014; however, we are evaluating alternative amplicon sequencing strategies that appear to have more appeal to the agencies who are the end-point users.

Task 4: We sequenced three haploid families, three gynogenetic diploid families and one diploid family previously mapped with microsatellite markers. Variable DNA sequences were mapped to all Chinook chromosomes using the haploid and gynogenetic diploid families. Microsatellites in the diploid family were used to identify chromosome arms and align the map with the rainbow trout chromosomes. We also mapped the SNP markers that have previously developed for the Chinook Coastwide SNP database. We also mapped genes responsible for thermal tolerance and growth.

Participants

We collaborated with the Cle Elum Supplementation and Research Facilities, managed by the Yakama Nations, and the Washington Department of Fish and Wildlife to create the haploid families for genome mapping. Use of these populations facilitated collaborations with these agencies and with NOAA on related work, researching the effects of domestication selection in Chinook hatchery lines.

We also collaborated with Washington Department of Fish and Wildlife, Alaska Department of Fisheries, and the Northwest Fisheries Science Center who are using SNPs developed by this project for studies of stock identification and reproductive success.

Finally, scientists from the University of Montana developed, from our data and for our joint use, a 10,000 feature array that we are using to test Chinook salmon from this project as well as other salmonids of interest in the Pacific Northwest.

Results

Task 1: We successfully sequenced 31 to 48 individuals per population. Our libraries comprised 36 individuals per sequencing lane and yielded an average of 80 million high quality reads per library. We identified 62,449 loci with high representation (80% or greater) across all populations and with a depth greater than 10 reads per individual. All of these loci were screened for quality and aligned, where possible, with the existing rainbow trout data (Miller *et al.* 2012). Loci passing this initial screen were added to a reference sequence database.

Task 2: We genotyped individuals using a bioinformatic pipeline based on PYTHON scripts and STACKS (Catchen *et al.* 2011). We identified 12,464 polymorphic loci across all populations with a minor allele frequency greater than 5% in at least one of the populations. These loci were screened using BLAST searches and haploid mapping was used (Task 4) to determine whether they were duplicated across chromosomes.

Task 3: With the help of a sister lab at the University of Montana we created and are initiating testing of a 10 000 feature array in haploid Chinook salmon (and other salmonid species).

Task 4: A substantial fraction of the polymorphic loci identified in task 2 were also variable across all of our mapping families. The three haploid families were used to map and identify duplicated loci. The salmon genome underwent a recent duplication – twice the number of chromosomes compared to the ancestral species. These loci can be used to identify the two matching chromosomes, but can complicate population genetic analyses. These haploid maps were then compared to the diploid map with microsatellite markers to identify the chromosome arms. Finally, we used the gynogenetic diploid crosses to identify the centromere location for each chromosome. Recombination probability increases as the distance from the centromere increases. Therefore the number of heterozygote offspring from a gynogenetic diploid cross is about 0 on or close to the centromere and close to 1 at the telomeres. Identifying the centromere and the chromosome arms is essential for ensuring complete coverage, but also gives us a means of comparing the salmonid genomes

as each species is sequenced. Out of the 62,449 loci identified in Task 1, 9,100 appeared to be duplicated. We mapped a total of 7107 non duplicated RAD loci and 149 of the previously identified SNP loci. Notably, in 2013 we successfully identified genomic regions associated with thermal tolerance and growth (Figure 1).

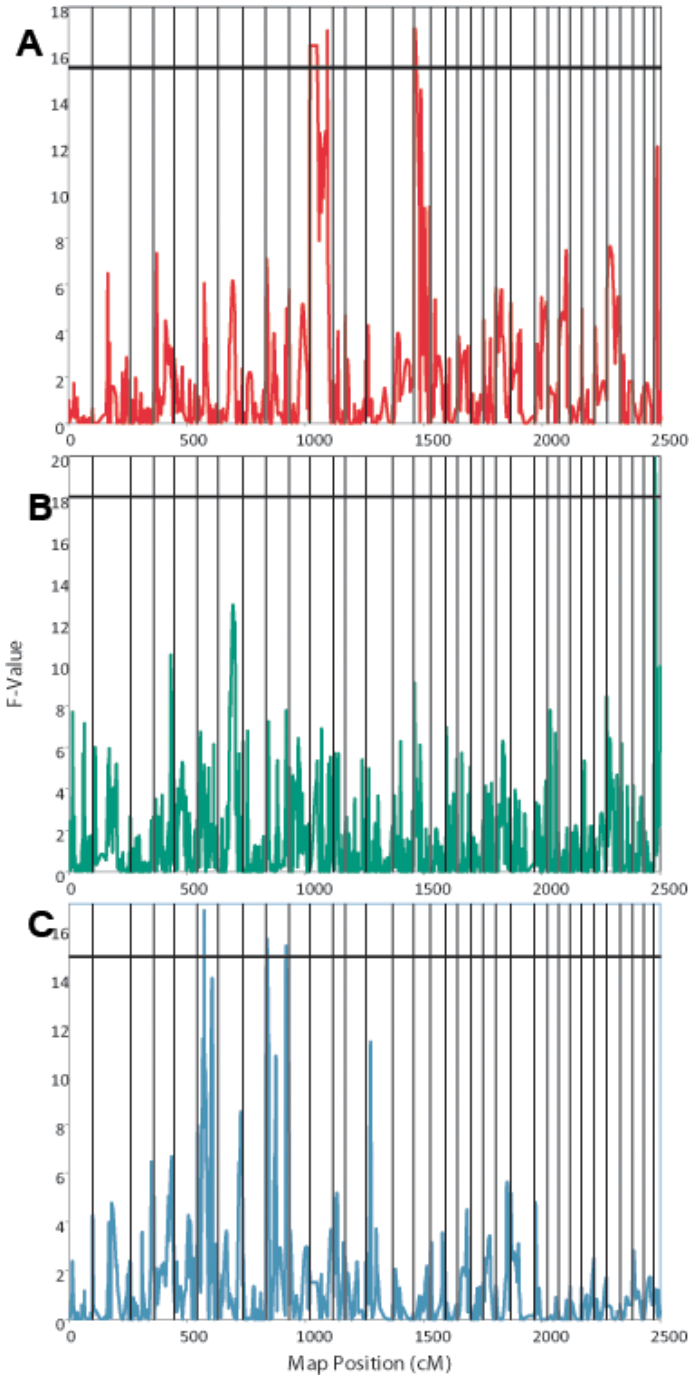

Figure 1 – Identification of genomic regions responsible for thermal tolerance (A and B) and growth (C). The distribution of F-values containing QTL significant at the experiment-wide level in the male (M7) and female (F14) parents. (A) Fvalues for thermotolerance in the male. (B) F-values for thermotolerance in the female F14. (C) F-values for body weight in the male. The dark horizontal line in all three figures is the experiment-wide ($P < 0.05$) significance threshold determined by a 10 000 permutation test (Churchill and Doerge 1994). Vertical lines designate individual linkage groups. Linkage group 11 on plot A contains two peaks, our model specified one QTL per linkage group, so only the higher of the two peaks was determined to be a QTL. QTL, Quantitative trait loci.
.

Challenges encountered

One challenge was developing the relevant bioinformatics tools for analyzing the data. However, we have established a full and reliable approach to attaining our goals. We have also assisted two laboratories in data analysis, and anticipate hosting International researchers in the coming year.

Changes in project directions

The objectives and deliverables remain intact, although the custom microarray originally envisioned was to be built for only 1 536 SNPs. Currently, we are describing many thousands of SNPs using RAD sequencing, and in 2013 we started testing 10,000 feature array developed by the University of Montana from our data (instead of the 1 536 array originally envisioned). Even as these tests are underway we find that end-point users are considering a less expensive approach termed amplicon sequencing. Our focus in 2014 will be to complete testing of the 10,000 feature array as well as running parallel tests on 200-500 locus amplicon panels identified through collaborations with Washington Department of Fish and Wildlife and Alaska Department of Fish and Game. These project upgrades came at a steep price that was paid by collateral research (KN funded by NOAA and JS funded by Moore Foundation and the Alaska Sustainable Salmon Fund).