Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1CW23

Title: Hakai Institute Juvenile Salmon Program Time Series

Creator: Brett Johnson - **ORCID:** 0000-0001-9317-0364

Affiliation: Hakai Institute

Principal Investigator: Brett Johnson, Brian Hunt

Data Manager: Brett Johnson, Tim van der Stap, Krystal Bachen

Funder: Tula Foundation

Template: Hakai Institute Data Management Plan

Project abstract:

The Hakai Institute Juvenile Salmon program is an ongoing initiative that was established in 2015 in partnership with the University of British Columbia, University of Toronto, Simon Fraser University and Salmon Coast Field Station. This program researches the early life history of juvenile salmon in coastal British Columbia. Primary research objectives are determining: 1) Migration timing rates and routes; 2) Migration habitat, including physical and chemical oceanographic conditions, and availability of plankton prey; 3) The impacts of prey phenology, quantity and quality on juvenile salmon growth and condition; 4) Species and stock-specific feeding biology and competitive interactions; 5) Pathogen and parasite infection dynamics; and 6) Mortality estimates. The program targets Fraser River sockeye, and pink and chum salmon, but additionally provides information on coho, chinook, and herring through incidental capture. The field program operates between May and July during the peak of the juvenile sockeye outward migration. Purse seine and oceanographic sampling are conducted in the northern Strait of Georgia / Discovery Islands region (~ 220 km from the Fraser River mouth).

Start date: 05-12-2015

Last modified: 01-29-2024

Copyright information:

language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Hakai Institute Juvenile Salmon Program Time Series

Data Overview

The core of the Hakai JSP are the observations made at sea during seining operations, and the observations and measurements made in the lab during fish dissections. Many samples are produced from fish that are collected in the lab, and those data may be addressed in their own Data Management Plans.

Data from the field and lab dissections are lumped into the Hakai JSP Time Series at www.github.com/hakaiinstitute/jsp-data and includes:

- Catch abundance by species
- Fork lengths and weights
- Sea lice counts
- Visual transect surface activity counts
- Survey metadata
- CTD data
- YSI Data
- Genetic Stock ID results

Laboratory samples' unique identifiers are also stored in the GitHub jsp-data repository, while their physical location is tracked in eLab software. This includes:

- Muscle tissue for fatty acid analysis, isotope ratios, RNA:DNA ratios
- Fin clips for genetic stock ID
- Liver, kidney, heart, brain, gill tissues for histology and pathogen analysis
- Otoliths for ageing and microchemistry analysis
- Zooplankton samples

Data Storage and QA/QC Procedures

Most data are collected on physical data sheets in the field and in the lab, with the exception of CTD data which is recorded on the instrument and metadata are collected using digital Forms on a tablet. Physical spreadsheets from the field and the lab are transcribed as soon as possible by entering data into Google Sheets spreadsheets. These spreadsheets have numerous quality assurance features including conditional formatting for accidentally duplicated fields, expected ranges of values, drop-down selection of categorical variables, and index matches to lookup values from a reference table to auto-populate some data entry fields. Every row of data that is entered is double-checked for accuracy and transcription errors by visually cross-referencing the physical data sheet with the google sheet by someone who didn't enter the data.

We use a new Googlesheets workbook for each field and lab season. At the end of each field season data from the annual spreadsheets are copied into the Master JSP Data Tables spreadsheet, or the Master JSP Sample Inventory which contains all years of data. These spreadsheets are then read into an data-integration-and-QC.R script which conducts Quality Checks in part by building a relational data model that checks the validity of primary and foreign keys ie. whether every sample has a corresponding fish, and every fish has a corresponding

seine, and every seine belongs to a survey. The structure and contents of the JSP data model can be seen in this Entity Relationship Diagram. This script also checks for outliers in various variables, and makes sure that all locations are not on land and more.

For complete documentation of the data processing steps see our jsp-data README file and read through the comments in the various scripts.

Data Sharing Policy and Timeline

Data produced in the field and the lab dissections carried out by Hakai staff will be made public within 6 months of collection. Being made public refers to the data being uploaded to the Hakai JSP GitHub repository which can be accessed by anyone. The data are available immediately to internal collaborators, including UBC graduate students in the Pelagic Ecosystems (Hunt) Lab in their raw format via googlesheets in the JSP Team Drive.

Data are released under CC BY 4.0 and have no ethical or legal restrictions on their use. We do however expect researchers using these data to contact us for collaboration opportunities.

Data produced from secondary sample processing, such as fatty acids, eDNA, stomach contents and so are embargoed for a period of up to three years so that students have first-access rights. Projects that mostly use data collected at Hakai or if the student/Post doc position is funded by Hakai are encouraged to create a standalone Data Management Plan.

Data Publishing Plan

Core observations made in the field and from lab dissections conducted by Hakai staff including species counts, fork lengths, and weights are published to the Ocean Biodiversity Information System here which also is automatically registered with the Global Biodiversity Information Facility here. These records will be updated annually. Both of these records link back to the central Hakai Institute Data Catalogue record for the Juvenile Salmon Program Time Series which is referenced with a Digital Object Identifier at https://doi.org/10.21966/1.566666

We also publish an annual report as a North Pacific Anadromous Fish Commission Document.

Secondary sample processing results and or derived/synthesis data sets using Hakai JSP data or created by Hakai funded student/postdoc positions should create a metadata record and submit a data package by filling out a metadata intake form and sending a data package to catalogue-team@hakai.org. Data packages must include a plain text (.txt, .csv) data table, a data dictionary defining every variable in dataset, protocols, a changelog to track changes between versions, and readme file that provides additional context not provided in supplementary material or metadata abstract.

Roles and Responsibilities

Brett Johnson will publish annual updates to the JSP Time Series as well as update the OBIS record. Tim van der Stap and Krystal Bachen are responsible for quality checking each years annual field and lab data. Krystal Bachen is responsible for sample management, including updating eLab as needed. Any papers authored using the JSP Time Series data must result in Brett, Tim, Krystal, and Brian being offered co-authorship opportunity. Otherwise the principles outlined in the CRediT authorshipship guidelines: https://www.elsevier.com/authors/policies-and-guidelines/credit-author-statement. New datasets published related to the JSP will be added to this DMP by Brett Johnson.

Planned Research Outputs

Dataset - "Fatty acids from juvenile salmon"

This dataset will include fatty acid data derived from juvenile salmon tissue samples.

Alicia Andersen is going to provide this dataset.

Dataset - "Otolith Microchemistry from juvenile sockeye"

Microchemistry data as obtained from otolith analysis. Otoliths were obtained from juvenile salmon as part of the Juvenile Salmon Program (JSP). The responsible party for this dataset is Yuliya Kuzmenko (UBC).

The Paper has been published at https://doi.org/10.26428/1606-9919-2021-201-669-685 but we need to reach out to Yuliya to collect data.

Dataset - "Sealice microbiome viral and bacterial data"

The responsible parties for this dataset are Tianyi Chang and Curtis Suttle.

Need an update from Tianyi

Dataset - "Stable isotope ratios from juvenile salmon muscle tissue"

This dataset will contain isotope data obtained from juvenile salmon caught during the JSP. The responsible party for this dataset is Brian Hunt.

Dataset - "Zooplankton Isotopes from JSP surface horizontal tows"

This dataset will contain isotope data obtained from the zooplankton caught as part of the Juvenile Salmon Program (JSP). Brian Hunt is the responsible party for this dataset.

Dataset - "eDNA from JSP seines"

Environmental DNA (eDNA) is collected as part of the Juvenile Salmon Program field operations, starting in 2020. The responsible party for this dataset is Natalie Benoit (UBC).

Dataset - "Zooplankton Taxonomy from JSP and Oceanography horizontal surface tows"

An integration of Sam James, Vanessa Fladmark, and Natalie Mahara's zooplankton taxonomy from horizontal surface tows in the Discovery Islands and Johnstone Strait.

Dataset - "RNA pathogen and gene expression data: Fitchip data from sockeye"

Dataset from fish tissue samples sent to Kristi Millers lab. Currently 2015 sockeye, pink and chum dataset is published as part of the Hakai JSP data package.

Katherine Medcalf' (UoT) project will produce fit chip data from Krisiti Miller's lab for 2016 sockeye and will be used in conjunction with sea lice count data and the 2015 fit chip data. Katherine is working on a separate DMP for that.

Dataset - "Genetic stock identification of sockeye"

Fin clips and liver samples are sent to the Pacific Biological Station for stock identification. Currently, we have results from 2015–2019.

These data could be published in a stand alone data package, with an accompanying data paper. Currently, they reside in the JSP GitHub repo along with the JSP Time-series dataset.

Dataset - "Scales collected from sockeye, pink, chum, coho, chinook and herring"

Scales are collected for every salmon and herring and archived. No analysis has been completed at this stage nor is any planned.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Fatty acids from juvenile salmon	Dataset	Unspecified	Open	Hakai EIMS		Creative Commons Attribution 4.0 International	ISO 19115	No	No
Otolith Microchemistry from juvenile sockeye	Dataset	Unspecified	Open	Hakai EIMS		Creative Commons Attribution 4.0 International	None specified	No	No
Sealice microbiome viral and bacterial data	Dataset	Unspecified	Open	None specified		Creative Commons Attribution 4.0 International	None specified	No	No
Stable isotope ratios from juvenile salmon muscle ...	Dataset	Unspecified	Open	Hakai EIMS		Creative Commons Attribution 4.0 International	ISO 19115	No	No
Zooplankton Isotopes from JSP surface horizontal t ...	Dataset	Unspecified	Open	None specified		Creative Commons Attribution 4.0 International	None specified	No	No

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
eDNA from JSP seines	Dataset	2024-04-13	Open	Ocean Biodiversity Information System		Creative Commons Attribution 4.0 International	ISO 19115 North American Profile ISO 19115	No	No
Zooplankton Taxonomy from JSP and Oceanography hor ...	Dataset	2022-03-30	Open	Hakai EIMS Ocean Biogeographic Information System		Creative Commons Attribution 4.0 International	None specified	No	No
RNA pathogen and gene expression data: Fitchip dat ...	Dataset	Unspecified	Open	Hakai EIMS		Creative Commons Attribution 4.0 International	None specified	No	No
Genetic stock identification of sockeye	Dataset	Unspecified	Open	Ocean Biogeographic Information System		None specified	None specified	No	No
Scales collected from sockeye, pink, chum, coho, c ...	Dataset	Unspecified	Open	None specified		None specified	None specified	No	No