Spring-Run Chinook Salmon JPE Run Identification Program Research and Initial Monitoring Plan

Updated: November, 2023

Prepared for the California Department of Water Resources by:

Mallory Boro, Melinda Baerwald, Sarah Brown, Nicole Kwan, and Brett Harvey

California Department of Water Resources

Noble Hendrix

QEDA Consulting

Sean Canfield

UC Davis

Jeff Rodzen and Sheena Holley

California Department of Fish and Wildlife

Contents

Figures

Tables

Introduction

Central Valley spring-run Chinook salmon (*Oncorhynchus tshawytscha*; "spring-run") are listed as threatened under both the California Endangered Species Act and the federal Endangered Species Act. In March 2020, the California Department of Fish and Wildlife (CDFW) issued Incidental Take Permit No. 2081-2019-066-00 (ITP) to the California Department of Water Resources (DWR) for the operation of the State Water Project (SWP), which describes the necessary conditions to avoid, minimize, and mitigate for impacts of the SWP on spring-run Chinook salmon, among other covered species. Condition of Approval 7.5.2 of the ITP requires DWR to convene an interagency team (JPE Team) to support the development and implementation of an annual spring-run Juvenile Production Estimate (JPE). In December of 2020, DWR submitted the Incidental Take Permit Spring-Run Chinook Salmon Juvenile Production Estimate Science Plan (JPE Science Plan 2020) for years 2020 to 2024 to CDFW, which outlined the components necessary for the development of a spring-run JPE. The plan has been reviewed by CDFW and was approved in February 2021.

A primary element of the JPE Science Plan is the development of a Chinook salmon Run ID program to identify juvenile spring-run individuals at key monitoring locations throughout their known range within the Central Valley. In the Central Valley, the juvenile spring-run co-occurs with three other Chinook salmon runs (i.e., winter, fall, and late-fall) and, to the naked eye, are morphologically indistinguishable from these other runs. Current non-genetic run identification methods rely on two different Length-At-Date (LAD) models: the River Model and the Delta Model. The accuracy of these models has been shown to be questionable, leading to frequent run misidentification (see PLAD section below). Therefore, some means of identifying juvenile spring-run from the other salmon runs migrating through the Sacramento-San Joaquin Delta (Delta) will be critical to calculating a spring-run JPE. The purpose of this document is to provide an updated outline of the research and monitoring DWR has and will continue to implement to meet the goal of developing a genetically supported spring-run identification program (referred to as the Run ID Plan for the rest of the document) ready for full implementation in 2025. The original version of the Run ID Plan was written, approved, and implemented in October 2021. This current version of the plan provides updates to the recommended sample sizes for each monitoring location.

Figure 1 Fork Length Distribution of Genetically Assigned Chinook Salmon and Delta Model Length-at-Date Size Ranges for Fish Sampled at Salvage Facilities from 2004 to 2010 (Harvey et al. 2014)
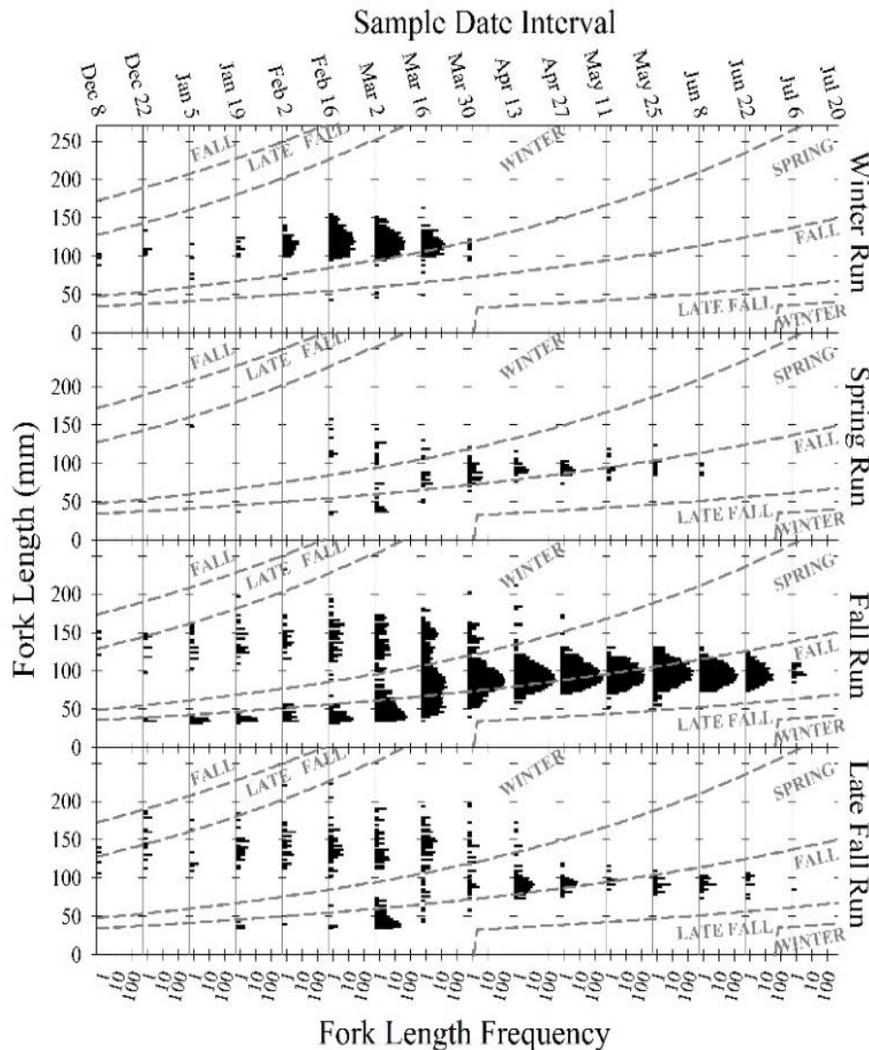
Notes: Horizontal bars indicate fork length distribution of genetically assigned Chinook Salmon. Gray dash lines indicate Delta Model LAD size ranges.

Multiple genetic tools exist or are in development that will be applied in combination with newly developed Probabilistic Length-at-Date (PLAD) models to achieve a cost-effective means of reaching the level of identification accuracy necessary (and logistically feasible) as determined through modeling and the JPE structured decision-making process (more

details regarding this process are described in the JPE Science Plan). Given the anticipated need for large-scale run identification for juvenile Chinook salmon in the Sacramento River, its tributaries, and at salvage facilities, it is worthwhile to optimize identification methods. Widely used current methods have either low power of run identification (i.e., the LAD approaches discussed below) or could benefit from efficiency improvements (e.g., genetic approaches).

Figure 2 Proportion of Genetic Run Sampled within Each Delta Model Length-at-Date Range at the State and Federal Salvage Facilities, 2004–2010 (Harvey and Stroble 2013)
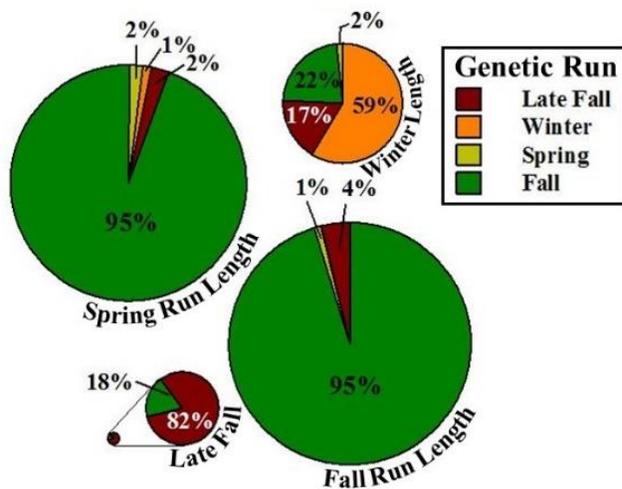
Note: Pie size represents relative number of sampled fish, N = 11,609.

The goal of the Run ID Program described in this plan is to improve run identification accuracy, simplicity, and speed at minimal cost by harnessing new technologies. The two aspects of this plan, development of PLAD models and application of new genetic technology for key sampling locations, will work synergistically and iteratively to inform one another, thus creating a feedback loop of constant improvement.

What is PLAD?

PLAD is a Bayesian approach for assigning run identification probabilities for Central Valley Chinook salmon. Development of the approach was initiated in 2017 and originally targeted identification of winter-run Chinook salmon, but the approach is also applicable to identification of other Chinook salmon runs, including spring-run (Noble Hendrix, Qeda Consulting, personal communication). The PLAD is an extension of the LAD approach for assigning

run identification to a juvenile salmon based on capture date and fork length (Figure 1). The LAD approach was originally proposed in 1989, concurrent with the federal listing of winter-run Chinook salmon under the Endangered Species Act, as a tool to assess take of juvenile winter-run Chinook salmon by the State and federal water projects (Harvey 2011). Currently, there are two different LAD criteria applied in the Central Valley. The *Delta Model* is used for fish sampled at the State and federal water project salvage facilities and in the DWR Yolo Bypass Fish Monitoring Program. The *River Model* is used for most other locations and sampling programs. The LAD approach relies on two major assumptions: (1) juvenile salmon of different runs hatch during distinct periods of the calendar year; and (2) all juvenile salmon grow at a constant rate. Genetic analyses show that neither of these assumptions are entirely true, and there is large overlap in size distributions between runs (Harvey and Stroble 2013; Harvey et al. 2014). Fall-run juveniles have considerable size overlap with spring-run juveniles for the Delta Model (Figure 1, third panel), and both fall-run and winter-run overlap considerably with spring-run for the River Model. Because of the large abundance of fall-run relative to spring-run, this overlap can lead to a high number of false positive spring-run assignments by LAD (Figure 2, green slice of upper left pie). Nonetheless, the LAD approach continues to be used for run assignment in many, if not most, CV monitoring programs, ostensibly because its speed and simplicity is useful for "real-time" management and because its application has minimal cost.

Figure 3 Conceptual Depiction of Probabilistic Length-at-Date Size Ranges for Two Runs
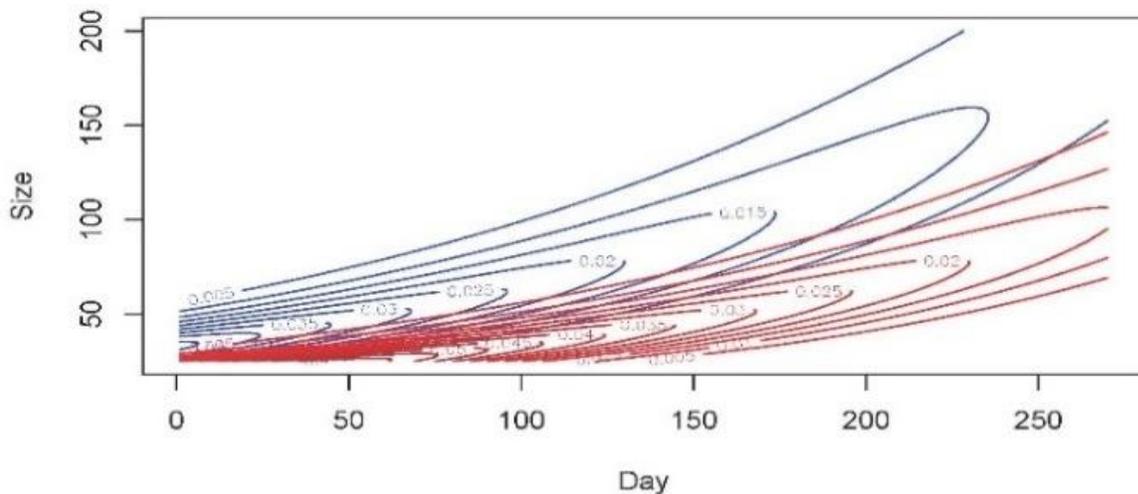
Source: Noble Hendrix, unpublished.

The probabilistic approach of PLAD has a similar construct to the original deterministic approach of LAD in that it relies on the fork length and sample date of a juvenile salmon to assign a run. Unlike LAD, the PLAD may assign more than one run for a given juvenile salmon, along with a probability for each run assignment (Figure 3). The assignment probabilities are based on genetic identification of catch from the preceding years of various monitoring programs. In addition to genetic information, variables such as geographic area, flow, and temperature may be incorporated as predictive variables into PLAD models. The long-term plan is to post model predictions on an internet platform such as SacPAS or possibly make the models available as an R application, which will be updated throughout the juvenile migration season as ongoing genetic and other calibration data become available. Field crews will access the models to determine assignment probabilities and, based on a predetermined cut-off for an unacceptable level of uncertainty in PLAD identification (i.e., low assignment probability), sampled juveniles will undergo tissue sampling for subsequent genetic run identification. Genetic results will subsequently be fed back into PLAD model calibrations for continual model improvement.

PLAD models under current development are focused on identification of winter-run versus non-winter-run, and assignment probabilities are calibrated with coupled genetic and fork length data from key sampling locations along winter-run migration routes. Spring-run PLAD models will require genetic, fork length, and environmental data specific to sampling locations where spring-run PLAD models will be used. PLAD assignment accuracy for a given sampling location will depend on the site-specificity, with size distinctions between runs becoming more blurred over the migration season and as juveniles move downstream. This is a byproduct of variable migration and growth rates, along with the mixing of populations from streams with different environmental conditions. For all PLAD models, but especially for these downstream locations, accuracy will depend on the data used to calibrate the models, and in particular, the accuracy of the genetic run identification.

Genetic tools for identifying spring-run Chinook salmon

During the initial research and development phase, we will likely continue to use established genetic tools and techniques to identify run, which will then be used to inform and improve PLAD curves. There are several sets (panels) of Chinook salmon SNP loci available and in use for genetic stock

identification and run assignment. An established baseline constructed from thousands of individuals along the western United States using a panel of 96 SNP loci was developed by NOAA (Clemento et al. 2014) and is used for identification at DWR salvage facilities. Another SNP panel comprised of 80 SNPs identified from Chinook salmon sampled at all major CV tributaries is also used for stock identification (Meek et al. 2016). The genetic test used for a specific location often depends on the expected Chinook salmon runs present and the degree of genetic resolution required to answer particular scientific questions. Spring-run and fall-run Chinook salmon releases from the Feather River Fish Hatchery will also be taken into account because of the relatively large number of hatchery smolts compared to natural origin smolts, and because of introgression between spring-run and fall-run, which is most prevalent in the Feather River Fish Hatchery and Feather River natural origin fish, though introgression is observed in other spring populations as well.

Individual fish can be genetically typed at various SNP panels using several different sequencing systems. The two approaches most used are Fluidigm's microfluidic genotyping chips and Genotyping-in-thousands by sequencing (GT-seq) (Campbell et al. 2015) by the direct sequencing of PCR amplicons on an Illumina sequencer. Using Fluidigm's technology, DNA extract from each individual salmon is kept separate and genetically typed using up to 96 individual SNP genotyping assays. Typically, 96 samples (or slightly fewer when controls are also included) and 96 different SNP assays are run simultaneously on a microfluidic chip, yielding 9,216 genotypes per chip. It takes about six hours to obtain genotypes from a DNA sample plate. The GT-seq approach uses massive parallel sequencing (e.g., Illumina sequencing) to collect data on potentially hundreds of genetic markers on a few hundred individual fish simultaneously. Microhaplotyping is a variation of the GT-seq approach wherein multiple multiallelic markers (as opposed to individual biallelic SNPs) are genetically typed using massively parallel sequencing. Currently, SNP panels used for run assignment frequently allow for some identification to tributaries of origin, such as spring-run individuals to Deer/Mill creeks and Butte Creek (Meek et al. 2016).

The adult migration genetic type of an individual fish does not necessarily align with the genetic stock assignment obtained from SNP or GT-seq panels because of the introgressive hybridization that occurs between Chinook salmon runs when they interbreed. For example, a Chinook salmon may

assign with > 90 percent probability to the fall genetic stock using SNP panels, yet it may be a spring-returning fish (i.e., a spring-run). This is a problematic situation for JPE-related run identification since protection under both federal and State ESAs is based on the future expected phenotypic expression of spring-running behavior for juveniles (if they were to survive to adulthood). Therefore, improved genetic identification of juveniles likely to exhibit spring-running behavior will be critical to an accurate JPE. (Note that the spring-run San Joaquin River population is experimental and not included under these protections.) To address this problem, we expect to use genetic markers from the GREB1L/ROCK region of Chromosome 28 (Prince et al. 2017). This genomic region was recently found to have a high statistical correlation with the migration timing variation of adult salmon, which is the primary phenotypic characteristic used to distinguish between salmon runs. We can use GREB1L markers to determine "early" or "late" migration genotypes, which will help with identification in streams where there is overlap during spawning between spring-run and fall-run salmon resulting in juveniles of both runs and potential hybrid juveniles appearing simultaneously in monitoring samples. Genetic typing assays that target this genomic region are available and in use for both Fluidigm and GT-seq genotyping approaches. Genotyping at these markers will also provide insight into the rate of hybridization (i.e., numbers of fish that carry both an "early" and "late" genetic variant) and the distribution of migratory behavior among hybrids, which may display early or late returning adult migration timing, or intermediate timing.

A rapid, accurate, sensitive, and cost-effective alternative to the current common genetic typing methods like those described above (Fluidigm and GT-seq genotyping approaches) is innovative CRISPR technology. SHERLOCK (Specific High sensitivity Enzymatic Reporter unLOCKing) is a CRISPR-based system originally designed as a diagnostic tool for human pathogen detection and is capable of being conducted in the field with minimal training or equipment (Gootenberg et al. 2017). SHERLOCK assays have been successfully developed to distinguish three different smelt species, Delta Smelt (*Hypomesus transpacificus*), Longfin Smelt (*Spirinchus thaleichthys*), and Wakasagi (*Hypomesus nipponensis*) in the San Francisco Estuary (Baerwald et al. 2020). Recently, SHERLOCK assays were also developed to distinguish Chinook salmon runs that co-occur in the Central Valley (Baerwald et al. 2023). To distinguish Chinook salmon runs using SHERLOCK, we took a tiered/hierarchical approach. First, we identified early

run types (winter-run and spring-run) from late run types (fall-run and late fall-run). Then, if an individual is identified as an early run type, we use a second set of SHERLOCK assays to distinguish between spring-run and winter-run (Figure 4). The use of SHERLOCK will reduce processing and analysis time substantially, because prior research has shown that DNA extraction is not necessary prior to SHERLOCK when using mucus samples (Baerwald et al. 2020), and individual molecular reactions take less than an hour to obtain results. The use of mucus sampled with a swab from the external surface of the fish may make fin clip samples unnecessary for some sampling situations, reducing the invasiveness and potential stress of genetic sampling on fish populations (Tilley et al. 2020). SHERLOCK assays are also significantly less expensive to process than current genetic methods used in the Central Valley. Although SHERLOCK has many advantages, it is relatively untested in comparison to Fluidigm and GT-seq approaches for Chinook salmon run assignment. Therefore, during the JPE research and development phase, we are conducting side-by-side testing of SHERLOCK with Fluidigm and/or GT-seq approaches on duplicate samples from the same individuals collected from all the JPE monitoring locations and will assess the correlation between these methods. Ultimately, we will determine the best long-term genetic methods or combination of methods for JPE-related run assignment based on the needs of the program and on cost and time savings.

Figure 4 Diagram illustrating the workflow from collecting a sample to making a run type call using the four SHERLOCK assays being used in this study
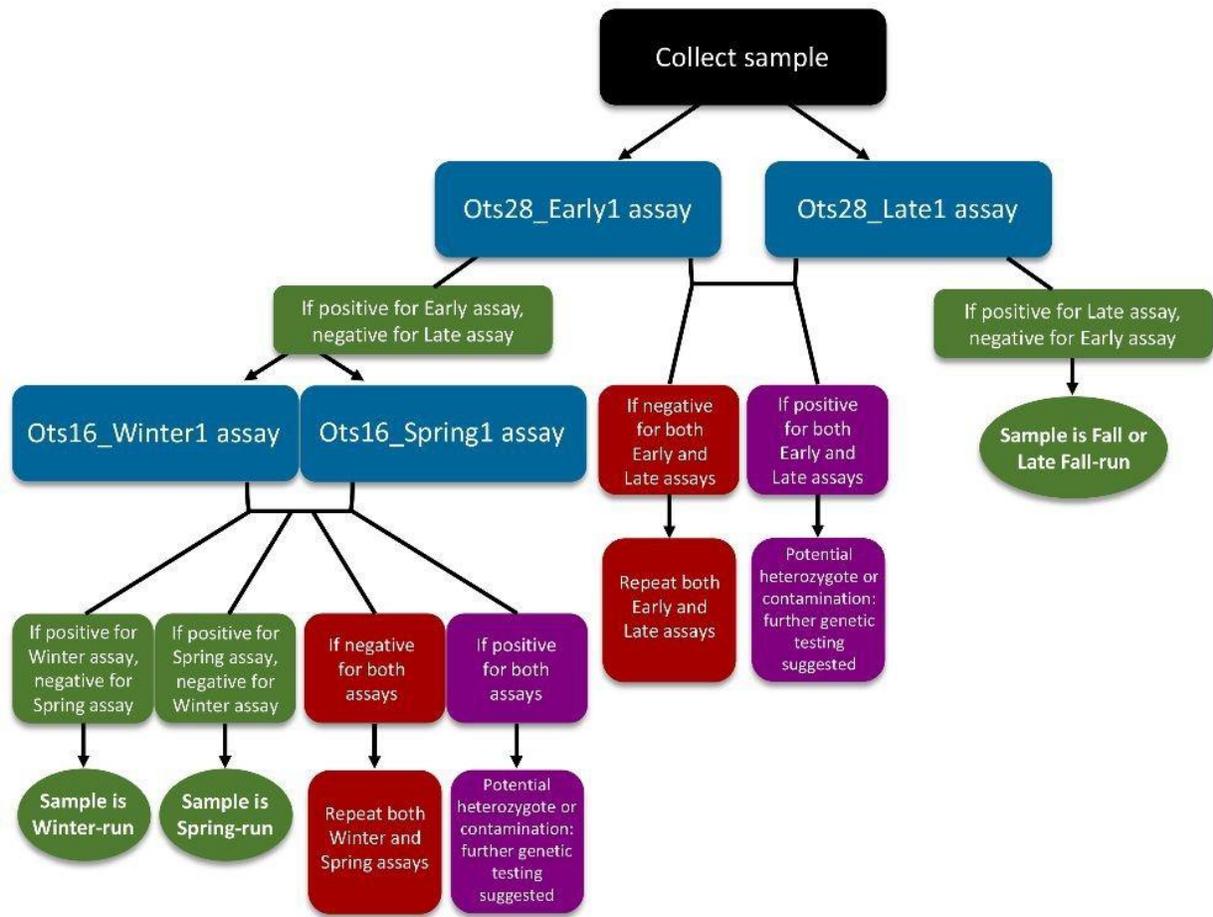
Current Monitoring

Currently, genetic sampling is being conducted at numerous locations within and downstream of the tributaries and rivers that flow into the San Francisco Estuary. Known surveys include Chipps Island trawl, midwater and Kodiak trawl at Sherwood Harbor on the Sacramento River, beach seines throughout the Delta, and screw traps at the Red Bluff Diversion Dam, Lower Stanislaus River, Lower American River, Battle Creek, and Upper Clear Creek, as well as adult carcass and snorkel surveys on the San Joaquin River, Battle Creek, and Lower Clear Creek. While not all these sites and life stages will be included in this genetic research and initial monitoring plan, they do represent a wide-ranging effort to collect genetic samples that could

contribute to our understanding of Chinook salmon population size and distribution.

PLAD LOCATION SPECIFIC MODEL DEVELOPMENT

PLAD model

The PLAD run assignment model is a type of finite mixture model in which there are multiple runs, each with a potentially different size distribution given their spawning timing and development rate prior to the date of sampling. This model is used to estimate the probability of run assignment given the fork length of the individual and the date of the sample.

$$R_{i,t} \sim Cat(p_{i,t,1:4})$$

The run of individual i at time t (R_i,t) is a categorial random variable (Cat), and the probability of run assignment for individual i in each of the four runs is p_i,t,1:4. Note that Cat is equivalent to a multinomial random variable with a single observation, i.e., $Cat(p_i,t,1:4) = multinomial(1, p_i,t,1:4)$. The probabilities of run assignment are computed from the fork length of the individual FL_i,t at the time of sampling t given the predicted distribution of fork lengths for run j FL_j,t.

$$p_{i,j} = Pr(FL_{i,t} \mid FL_{j,t}) \; / \; S \; j{=}1{:}4 \; Pr(FL_{i,t} \mid FL_{j,t})$$

We can work with the unnormalized probabilities of run assignment for individual i (U_i,j), where $U_{i,j,t} = Pr(FL_{i,t} \mid FL_{j,t})$.

The predicted distribution of fork lengths for run j, FL_j,t, is modeled as

$$FL_{j,t} \sim lognormal(m_{j,t}, s2_j)$$

where m_j,t is the mean of the distribution and $s2_j$ is the variance. The mean is modeled using a simple log-linear growth model

$$log(m_{j,t}) = a_j + b_j \, t$$

where a_j is the parameter that relates the timing of entry into the sampling population for run j; high values of a_j are associated with runs that are earlier to emerge, whereas low values of a_j are associated with runs that are

later to emerge. The parameter bj defines the growth rate of run j; larger values indicate faster growing runs, whereas lower values indicate slower growth runs. In order to prevent shrinking fork lengths of fish, bj > 0.

We note that this structure is flexible and other growth models can be used besides the log-linear model described here to define the distribution of sizes predicted for each run j at each sampling time t. Other growth models that could also be incorporated include the Von Bertalanffy or the Ratkowski growth models.

Estimating the parameters of the growth model

In order to make predictions of the fork lengths for each run, estimates of the parameters (α_j, β_j, σ^2_j) of the growth model are needed. We use Bayesian estimation to obtain posterior distributions of each of these parameters, which allows us to make predictions of the probability of run assignment and incorporate estimates of uncertainty with the run assignment, if they are provided. For example, under this modeling framework, the following (hypothetical) statement can be made, "The run assignment of the observed fish has the highest probability of being winter-run with a mean estimate of 0.73, and the 95 percent probability interval on this estimate is (0.56, 0.93)."

Updating the model parameters

The statistical model that underlies the PLAD predictions of run assignment can be updated whenever new genetic identification data become available. This process is relatively straightforward: (1) update the set of observations of fork lengths from juveniles with genetic identification; (2) refit the statistical models with the updated data to obtain updated posterior distributions of the model parameters (aj, bj, s2j); (3) use the updated estimates of the model parameters to make updated predictions on run assignment (Figure 5).

The genetic identification data can come through regular sampling via a standard monitoring protocol in which specific numbers of fish are sampled for genetic testing at regular intervals (e.g., biweekly). This will be the initial approach employed to calibrate the initial PLAD models (described in more detail in the Genetic Approach Initial Studies section). However, as soon as is practical, the PLAD models will be used to help direct genetic testing effort through an adaptive, iterative process which will maximize the value of

genetic information toward reducing uncertainty of spring-run JPE predictions. The adaptive approach (described in more detail in the Field Implementation section) will use PLAD models to direct field genetic sampling toward juvenile salmon that will yield the highest value information for improving PLAD model accuracy or for targeting juveniles for which the PLAD models are least effective at distinguishing between runs. As this new information is used to re-parameterize PLAD models, PLAD predictions will become more accurate at both run identification and at directing genetic sampling toward high-value samples. This iterative adaptive approach will facilitate both identification accuracy and increased efficiency of the valley-wide Run ID Program.

Figure 5 Schematic of Anticipated General Steps That Will Be Taken in the Run Identification Process

(N. Hendrix)
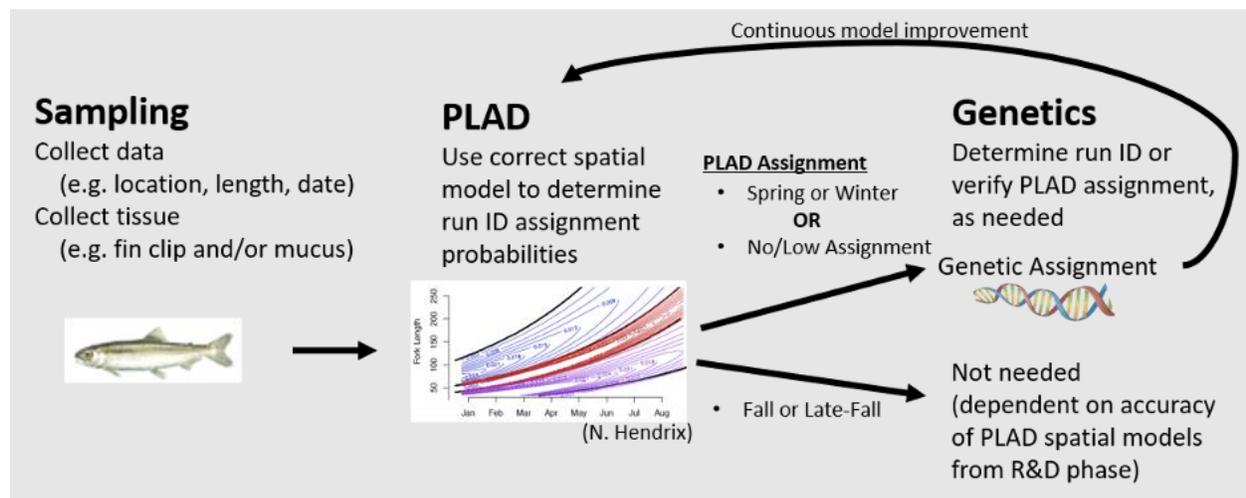
Model predictive accuracy, management decisions, and genetic sampling

There may be situations where managers will want to know the proportions of each run to a higher degree of accuracy than can be provided by the PLAD model. For example, this situation may occur when the sizes of spring-run and fall-run strongly overlap, but managers want a specific estimate of the proportion of spring-run in the mixture. The PLAD model predictions provide the bounds on the estimated probability of run assignment to facilitate these types of decisions. For example, the PLAD model could hypothetically predict the probability of run assignment for spring-run as 0.48 (0.42, 0.55) and the probability of fall-run as 0.52 (0.35, 0.63) for a fish in hand with a fork length of 75mm. This result indicates that the PLAD model is fairly confident

that there is an approximately 50 percent chance that the fish in hand is spring-run and a 50 percent chance that it is fall-run. If a management decision required knowing the proportion of spring-run at this sampling location and time, then this fish and many others could be sampled for genetic analysis to estimate the proportion spring-run from those samples using a binomial estimator. Of course, once the genetic testing was completed, these data would also be used to update the PLAD model as described above.

FIELD IMPLEMENTATION

Tissue samples for genetic analysis do not need to be collected from all juvenile salmon sampled in monitoring programs that support the JPE development. Instead, PLAD models will be used to locate where and when low probability spring-run assignments are anticipated, so that genetic testing can be done on samples from specific times and locations that would derive the most benefit from genetic assignment. The ability to target genetic testing is expected to improve with each iteration of PLAD using new genetic information. At these locations and times, tissue for genetic analysis will be collected from a subset of juvenile salmon encountered in monitoring programs. The number of genetic samples collected will be specific to each monitoring location depending on daily catch, fork length, expected proportions of spring-run in the sample, and the genetic test applied to samples from that location (Table 1). As described above, the decision to genetically sample an individual, and how many individuals to sample, will be based on the PLAD output for each location. The precise PLAD uncertainty thresholds used to guide genetic sampling will vary by site and depend on a balance between the value of additional information toward reducing JPE uncertainty and the ability of the Run ID program to process those samples. As more genetic data is collected and the PLAD curves are refined, we expect fewer genetic samples will need to be taken, as curves become more accurate over time and PLAD estimates are better able to guide genetic sampling to the most informative locations and times for reducing JPE uncertainty (Figure 5). Implementation of sampling in the field will consider life stage, agency familiarity with genetic sampling techniques, and archiving needs. As it is difficult to fin clip smaller fish (FL < 55 mm), the eventual goal is to have juvenile salmon genetically sampled using a swab.

It is necessary that we can identify fish of San Joaquin origin at the fish salvage facilities located at the State and federal water projects' south Delta pump intakes, because the San Joaquin population is experimental and not included in take calculations. Distinguishing the San Joaquin population from Sacramento River origin spring-run using genetics may be problematic because of the current practice of sourcing the San Joaquin River stock using crosses from the Feather River Hatchery. Potential alternative tools to distinguish San Joaquin origin juveniles include otoliths, CWT, acoustic tagging, photonic tagging (which has been done previously by SJRRP) (Hutcherson et al. 2020), or additional fin clips. Initial assessments may find that the San Joaquin contribution to take is not large enough to warrant further action. Currently, the number of juveniles produced from the San Joaquin population is orders of magnitude less than those produced from the Feather River. However, the San Joaquin population will need to be assessed regularly to track a possible increase in contribution to take.

GENETIC APPROACH INITIAL STUDIES

We will initially analyze samples using the GREB1L region to identify early vs. late migrating individuals. If a sample is identified as an early migrator (i.e., homozygous for early alleles or heterozygous), it will be further analyzed using one of the SNP panels to distinguish spring-run from winter-run. We will also evaluate the best genetic approaches for JPE monitoring, including how many SNPs are needed to adequately define a juvenile's run (which may be hybrid), and for some applications, further identify tributary origin. Once the initial testing and optimizing of the sampling and analysis process are complete, sampling will scale up to accommodate the sample numbers needed for PLAD and JPE estimates.

Mucus sample collection was introduced in this study as a result of research suggesting that swabbing the external surface of the fish reduced fish stress and was less invasive than fin clipping (Tilley et al. 2020). Both mucus and fin clip samples were collected in the 2021–2022 SR-JPE field study to validate both the SHERLOCK method and the use of mucus. Based on field observations of mortalities after handling and genetic sampling for the spring-run JPE, it has become imperative to determine which method of sampling will be the least harmful to small juvenile salmon.

We will conduct an enclosure study of juvenile fall-run salmon to assess general welfare and latent mortality following handling and genetic sampling.

Our study will consist of four treatments: (1) control group (netted and handled but no genetic sample taken), (2) fin clip only, (3) swab only, and (4) swab and fin clip collected. Cages stocked with small juveniles (>50 mm) will be placed in the Feather River. Following genetic sampling, fish fork length will be measured, and fish will be returned to their enclosures. Fish will be maintained in the enclosures for 72 hours to assess for latent mortality. Ultimately, we will determine the best long-term genetic methods or combination of methods for JPE-related run assignment based on the needs of the program and fish welfare. The data from this study will be used to inform future sampling efforts for the SR-JPE and other mandated projects necessitating rapid genetic protocols.

We will simultaneously develop SHERLOCK-based sampling and analysis methods which we expect will eventually replace current genotyping approaches, such as SNP panels, used for Central Valley Chinook salmon run identification. A duplicate mucus swab will be taken from a subset of juveniles for this purpose. The SHERLOCK swab samples will be analyzed using the previously described hierarchical approach of first using a pair of assays in the GREB1L region capable of distinguishing early versus late run type, and then analyzing those samples that test positive for early run type using a separate pair of assays capable of distinguishing spring-run from winter-run.

A subset of the fin clip and swab samples not consumed for genetic analysis may be archived at the CDFW Central Valley Tissue Archive.

TIMELINE

Since initial PLAD curves can be developed using genetic data that has already been collected, work will begin on these models prior to samples being collected as part of the new spring-run JPE monitoring program. These initial curves will also be used to refine the estimated number of needed genetic samples listed in Table 1. Genetic sampling as part of the new monitoring program began in January 2022, with swab and fin clip samples taken to begin developing SHERLOCK-based identification. Over the course of the 4-year research and development period prior to selection of a JPE approach (2021–2024) (Figure 6), PLAD curves will be refined with continued genetic sampling, which will in turn be used to update estimated genetic sampling requirements. Genetic sampling methods and plans will be evaluated and improved routinely over the research and development

period, after a JPE approach is established, and as new information and techniques become available (Figure 7).

Figure 6 Overall timeline of run identification research and initial monitoring plan before full run identification implementation in 2025
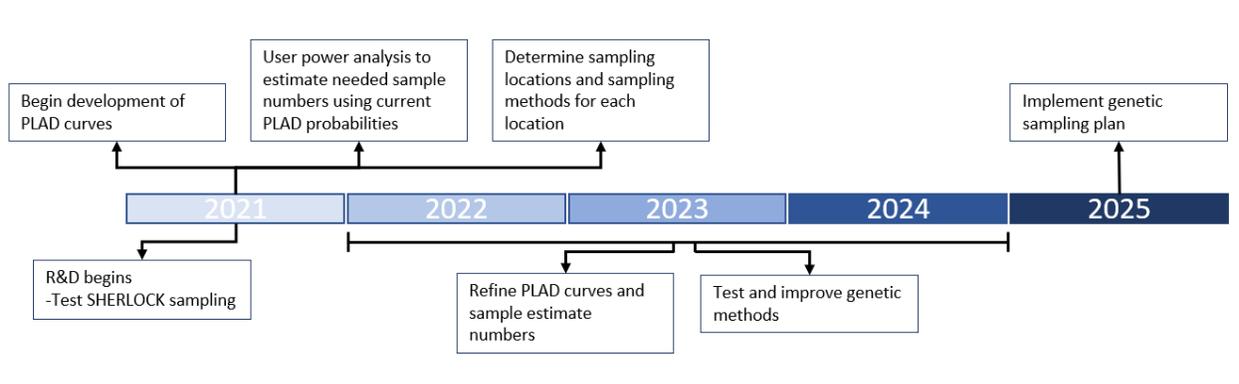
Table 1 Estimated annual genetic samples needed to build PLAD models and estimate spring-run abundance based on expected catch and run composition at JPE juvenile monitoring locations

Stream	Collection Type & Location	Max Annual Catch	Min Annual Catch	Max Daily Catch	Runs in Catch	% SR in Catch	Fin Clip Min FL	Sampling Agency	Genetic Tissue Sample	Annual Authorized Take (Genetic Sampling)	Total Sample Numbers for 2022 WY	Total Sample Numbers for 2023 WY
Mill	RST at RM 2.9 (above Ward Dam)	15,000	1,000	200	SR, FR, LFR	25–50	55mm	CDFW	Fin and/or Swab	400	62	67
Deer	RST at RM 4.9 (above SVRICDD)	15,000	1,000	200	SR, FR, LFR	25–50	55mm	CDFW	Fin and/or Swab	500	NA	58
Clear	RSTs at RM 1.77 (lower Clear Creek) and RM 8.4 (upper Clear Creek)	919,611	25,394	106,642	SR, FR, LFR	1–25	None; Yolk sac must be fully absorbed	USFWS	Fin and/or Swab	750	65	238
Battle	RST at RM 6.2 (above CNFH)	8,899	650	1,780	SR, FR, LFR, WR	25–50	None; Yolk sac must be fully absorbed	USFWS	Fin and/or Swab	190,000	50	105
Butte	RST at RM 44 (PPDD)	400,000	10,000	1,500	SR, FR	95	None	CDFW		100 fin or opercule	86	86

Stream	Collection Type & Location	Max Annual Catch	Min Annual Catch	Max Daily Catch	Runs in Catch	% SR in Catch	Fin Clip Min FL	Sampling Agency	Genetic Tissue Sample	Annual Authorized Take (Genetic Sampling)	Total Sample Numbers for 2022 WY	Total Sample Numbers for 2023 WY
									Fin and/or Swab	plus 4,000 fin clips		
Yuba	RST at RM 7.5 (Hallwood Blvd)	780,000	200,000	4,000	SR, FR	25–50	None	CDFW	Fin and/or Swab	200 LAD Spring Run; 3,000 Fall Run (limit shared with Feather)	NA	270
Feather-RM17	RST at RM 17 (lower Feather River)	3,000	2,500	350	SR, FR, LFR, WR	11	All	CDFW	Fin and/or Swab	Natural: 5,000 juveniles, 300 smolts. Hatchery: 1,500 juveniles, 200 smolts.	251	167
Feather-RM61	RST at RM 61	600,000	70,000	60,000	SR, FR, LFR	4	None	DWR	Fin and/or Swab	300 LAD spring-run; 3,000 Fall Run (limit shared with Yuba)	1,025	784

Stream	Collection Type & Location	Max Annual Catch	Min Annual Catch	Max Daily Catch	Runs in Catch	% SR in Catch	Fin Clip Min FL	Sampling Agency	Genetic Tissue Sample	Annual Authorized Take (Genetic Sampling)	Total Sample Numbers for 2022 WY	Total Sample Numbers for 2023 WY
Sacramento	RST at RM 119 (Tisdale Weir)	100,000	5,000	15,000	SR, FR, LFR, WR	39	None	CDFW	Fin and/or Swab	1,500 juveniles, 200 smolts	77	203
Sacramento	RST at RM 88.5 (Knights Landing)	100,000	5,000	15,000	SR, FR, LFR, WR	59	None	CDFW	Fin and/or Swab	1,500 juveniles, 200 smolts	31	165
Sacramento	RST at RM 75 (Delta Entry)	2,000	1,000	300	SR, FR, LFR, WR	20	All	CDFW	Fin and/or Swab	Natural: 5,000 juveniles, 300 smolts	48	66

Note: Initial samples numbers were calculated based on an initial power analysis scaled up using the maximum % SR catch and multiplied by 10 for each of the planned sampling occasions over the course of the 5-month sampling period. RM = river mile; UNK = unknown; SVRICDD = Stanford Vina Ranch Irrigation Company Diversion Dam; CNFH = Coleman National Fish Hatchery; PPDD = Parrot-Phelon Dam; LFC & HFC = Low & High Flow Channel.

Figure 7 Gantt chart detailing timeline of individual components of genetic sampling plan

2021												
Task	January	February	March	April	May	June	July	August	September	October	November	December
Draft initial sampling plan			■	■	■	■	■					
Set up PLAD contact through SWC				■	■	■	■	■	■	■		
Determine initial sample numbers needed at each site for PLAD curves					■	■	■					
Continue development of SHERLOCK assays								■	■	■	■	■
Develop sampling SOPs									■	■	■	■
Begin setting up contract for genetic analysis									■	■	■	■
Conduct training based on sampling SOPs										■	■	■
Develop site specific thresholds to determine entry into genetic sampling pipeline according to PLAD and resource capacity										■	■	■
Develop initial PLAD curves for chosen monitoring sites											■	■

2022												
Task	January	February	March	April	May	June	July	August	September	October	November	December
Begin sampling using PLAD-updated sample numbers	■	■	■	■	■							
Develop data management pipeline for receiving, QCing, and distributing genetic data	■	■	■	■	■	■	■	■	■	■	■	■
Genetic analysis of 2022 samples		■	■	■	■	■	■					
Test SHERLOCK assays with subset of 2022 samples		■	■	■	■	■	■					
Update PLAD curves for each site				■	■	■	■	■	■	■	■	■
Refine sample numbers for each site based on PLAD curves					■	■	■	■	■	■	■	■
Update SOPs for SHERLOCK samples					■	■	■	■	■	■	■	■

2023												
Task	January	February	March	April	May	June	July	August	September	October	November	December
Begin sampling using PLAD-updated sample numbers	■	■	■	■	■							
Finish optimization of SHERLOCK assays	■	■	■	■	■	■	■					
Refine data management pipeline for receiving, QCing, and distributing genetic data	■	■	■	■	■	■	■	■	■	■	■	■
Genetic analysis of 2023 samples			■	■	■	■	■	■				
Update PLAD curves for each site for 2024			■	■	■	■	■	■	■	■	■	■
Refine sample numbers for each site based on PLAD curves				■	■	■	■	■	■	■	■	■

2024												
Task	January	February	March	April	May	June	July	August	September	October	November	December
Begin sampling using PLAD-updated sample numbers	■	■	■	■	■							
Deploy SHERLOCK assays for genetic analysis	■	■	■	■	■	■	■					
Maintain data management pipeline for receiving, QCing, and distributing genetic data	■	■	■	■	■	■	■	■	■	■	■	■
Initial development of race identification plan and procedures for JPE implementation	■	■	■	■	■	■	■	■	■	■	■	■
Genetic analysis of 2024 samples	■	■	■	■	■	■	■	■				
Update PLAD curves for each site			■	■	■	■	■	■	■	■	■	■
Refine sample numbers for each site based on PLAD curves				■	■	■	■	■	■	■	■	■

REFERENCES

Baerwald MR, Goodbla AM, Nagarajan RP, Gootenberg JS, Abudayyeh OO, Zhang F, and Schreier AD. 2020. "Rapid and accurate species identification for ecological studies and monitoring using CRISPR-based SHERLOCK." Molecular Ecology Resources 20:4, 961-970. doi: 10.1111/1755-0998.13186.

Baerwald MR, Funk EC, Goodbla AM, Campbell MA, Thompson T, Meek MH, and Schreier AD. 2023. "Rapid CRISPR-Cas13a genetic identification enables new opportunities for listed Chinook salmon management." Molecular Ecology Resources 00, 1-13. Doi: 10.1111/1755-0998.13777.

Campbell NR, Harmon SA, and Narum SR. 2015. "Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing." Molecular Ecology Resources 15:4, 85-67. doi: 10.1111/1755-0998.12357.

Clemento AJ, Crandall ED, Garza JC, and Anderson EC. 2014. "Evaluation of a single nucleotide polymorphism baseline for genetic stock identification of Chinook Salmon (*Oncorhynchus tshawytscha*) in the California Current large marine ecosystem." Fishery Bulletin 112, 112–130.

Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Nina Donghia N, Daringer NM, Freije CA, Myhrvold C, Bhattacharyya RP, Livny J, Regev A, Koonin EV, Hung DT, Sabeti PC, Collins JJ, and Zhang F. 2017. "Nucleic acid detection with CRISPR-Cas13a/C2c2." Science 356:6336, 438-442. doi: 10.1126/science.aam9321.

Harvey BN. 2011. "Length-at-date criteria to classify juvenile Chinook Salmon in the California Central Valley: Development and implementation history." Interagency Ecological Program Newsletter 24:3.

Harvey BN and Carol 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." Technical Report 88 for the Interagency Ecological Program for the San Francisco Bay/Delta Estuary. March.

Harvey BN, Jacobson DP, and Banks MA. 2014. "Quantifying the Uncertainty of a Juvenile Chinook Salmon Race Identification Method for a Mixed-Race Stock." North American Journal of Fisheries Management, 34:6 1177-1186. doi: 10.1080/02755947.2014.951804.

Hutcherson J, Sutphin Z, Ferguson P, Grill M, Garza JC, and Clemento A. 2020. "Juvenile spring-run Chinook Salmon production and emigration in the San Joaquin River Restoration Area." 2017–18 Monitoring and Analysis Plan. San Joaquin River Restoration Program.

JPE Science Plan. 2020. *Incidental Take Permit Spring-Run Chinook Salmon Juvenile Production Estimate Science Plan 2020–2024*. Available at: water.ca.gov/-/media/DWR-Website/Web-Pages/Programs/State-Water-Project/Files/ITP/ITP-Spring-run-Chinook-Salmon-JPE-Science-plan-final-approved_Final_PDF_04-05-22.pdf.

Meek MH, Baerwald MR, Stephens MR, Goodbla A, Miller MR, Tomalty KMH, and May B. 2016. "Sequencing improves our ability to study threatened migratory species: Genetic population assignment in California's Central Valley Chinook Salmon." Ecology and Evolution, 2045-7758. doi: 10.1002/ece3.2493.

Prince DJ, O'Rourke SM, Thompson TQ, Ali OA, Lyman HS, Saglam IK, Hotaling TJ, Spidle AP, and Miller MR. 2017. "The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation." Science Advances 3:8. doi: 10.1126/sciadv.1603198.

Tilley CA, Gutierrez HC, Sebire M, Obasaju O, Reichmann F, Katsiadaki I, Barber I, and Norton WHJ. 2020. "Skin swabbing is a refined technique to collect DNA from model fish species." Scientific Reports 10:18212. doi: 10.1038/s41598-020-75304-1.