PACIFIC SALMON FOUNDATION

GENETICS AND 'OMICS' TECHNOLOGIES REVIEW FOR SALMON HATCHERIES

Prepared For:

Pacific Salmon Foundation
1682 W 7th Ave, Vancouver, BC, V6J 4S6

Prepared By:

EDI Environmental Dynamics Inc.
208A – 2520 Bowen Road, Nanaimo, BC, V9T 3L3

2020

EXECUTIVE SUMMARY

Salmon hatcheries have a long history in the management of Pacific salmon by increasing freshwater juvenile survival from a relatively small sample of adults that potentially translate to increased adult abundance and enhanced fisheries. As such, certain hatcheries are also used as tools to conserve threatened or endangered natural populations. Currently, greater than six billion juvenile salmon are released from hatcheries each year around the Pacific Rim (Waples *et al*. 2020). However, the effectiveness of increasing adult salmon abundance (for fisheries and/or spawners) varies over time and differs among specific hatcheries. Furthermore, hatchery salmon typically have lower performance than wild salmon (based on relative survival of smolts within release years), and there are concerns about potential negative impacts of hatchery-produced salmon interacting with naturally reproducing populations. And recently, epigenetic effects on gene expression, despite similar genetic background amongst salmon, have also been implied as a mechanism contributing to the poorer relative survival of some hatchery-produced salmon (Fraser 2008, Tave and Hutson 2019).

Four 'omics'[1] technologies now provide the opportunity to greatly advance the study salmon genetics and performance including the potential to improve hatchery performance. For example, existing biomarkers associated with important salmon traits (e.g. readiness for smoltification) provide valuable screening information on salmon health and condition that may help explain the survival patterns over time and among hatcheries. The recently developed technologies for environmental DNA (eDNA) and RNA (eRNA) in water samples are also considered within the scope of 'omics' described in this report (Cristescu 2019). Although there are reviews of 'omics' technology applications for fisheries and aquaculture, we are not aware of any such review specific to hatcheries. As such, a review of the four 'omics' technologies and their potential applicability to salmon hatchery research follows.

Broad-sense genomics (expansion of traditional genetics analyses), mainly using parentage-based tagging (PBT), covers fishery and escapement contributions (e.g. % hatchery and relative survival of families), other hatchery metrics (e.g. % straying and group performance), trait architecture (e.g. heritability), and genetic connections (e.g. hatchery genetic introgression). Narrow-sense genomics (novel analyses not possible before) covers gene biomarkers (sub-stock definition, e.g. run timing), trait epigenetic programming (e.g. age-at-maturity), and species distribution and abundance (e.g. rough salmon count using eDNA). Transcriptomics covers gene expression biomarkers (e.g. smoltification), as well as potential theoretical living and dead components (e.g. live salmon count using eRNA) and environmental gene expression (e.g. non-invasive biomarkers using mRNA derived from eRNA). Proteomics and metabolomics cover protein expression biomarkers (e.g. egg quality) and metabolite expression biomarkers (e.g. disease diagnostics).

Opportunities for biomarker development in association with salmon traits important to hatcheries is common for all four technologies. There are also opportunities for hatchery and wild salmon differences or other group differences within the hatchery setting, i.e. epigenetic programming differences and other expression differences, using the four 'omics' technologies to examine mechanistic explanations. This review has found that a breeding design (i.e. full factorial) with PBT can provide estimates of genetic and maternal effects explaining salmon traits important to hatcheries, e.g. fishery or escapement contributions. Large genetic effects can be examined further using an 'omics' technology to identify the specific targets increasing offspring quality, e.g. gene biomarkers. PBT is repeatedly identified as a useful tool because of the amount of new information that can be provided, especially for the adult life stage. The review also identified four knowledge gaps: (1) group performance, (2) offspring quality, (3) health and condition biomarkers, and (4) hatchery-wild differences. Filling in the four knowledge gaps using 'omics' technologies can provide guidance on how to improve hatchery performance and assess effects on natural populations.

Citation for this publication:

EDI (Environmental Dynamics Inc.). 2020. Genetics and 'Omics' technologies review for salmon hatcheries. 56p. Pacific Salmon Foundation, Vancouver, B.C. doi.org/10.48689/a2f1a2ee-5db1-4d76-a307-a95be0aa64c1

1. The term 'omics' refers to a field of study in biology ending in -omics, such as genomics, transcriptomics, proteomics, or metabolomics.

Cover photos by: top: Eiko Jones, left, centre and right: Sam James

ACKNOWLEDGMENTS

I would like to thank Ruth Withler, Terry Beacham, Kristi Miller, and Karia Kaukinen from Fisheries and Oceans Canada (DFO), as well as Isobel Pearsall and Brian Riddell from the Pacific Salmon Foundation for input on the scope of the review. I would also like to thank Shawn Garner from the University of Western Ontario for providing the scans of relevant chapters from a quantitative genetics' methods book, as well as Jeff Meggs from EDI and Cam West from DFO (retired manager of Salmon Enhancement Program) for providing comments on an earlier version of the review. Funding for this review was provided by the Pacific Salmon Foundation through the BC Salmon Restoration and Innovations Fund.

AUTHORSHIP

This report was prepared by EDI Environmental Dynamics Inc. Staff who contributed to this project include:

Aimee Lee Houde, *Ph.D.*, R.P.Bio. — Primary Author

Certification

Aimee Lee Houde, R.P.Bio. #3130

7 July 2020

Original Stamped Version on File

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES


Photo by: Mitch Miller

1. INTRODUCTION

Juvenile salmon[2] survival has decreased in the early marine environment for certain populations. Such decreases in early marine survival (first few months or even weeks after ocean entry) are strongly associated with decreases in salmon productivity (Beamish *et al*. 2009). In particular, coho salmon (*Oncorhynchus kisutch*) and Chinook salmon (*O. tshawytscha*) production has decreased in southern British Columbia (BC). For example, early marine survival has been < 5% for juvenile coho salmon entering the Strait of Georgia (Beamish *et al*. 2010). In efforts to increase salmon abundance, federally-managed hatcheries are used to produce coho salmon, Chinook salmon, sockeye salmon (*O. nerka*), chum salmon (*O. keta*), and pink salmon (*O. gorbuscha*) throughout British Columbia (BC) and Yukon (DFO 2013). Certain federal hatcheries also produce steelhead trout (*O. mykiss*, same species as rainbow trout) and cutthroat trout (*O. clarkii*) in partnership with the Province of BC . Overall, greater than six billion juvenile salmon are released from hatcheries each year around the Pacific Rim (Waples *et al*. 2020). Such hatchery programs have generally met the objective of increasing adult salmon abundance (Araki and Schmid 2010). However, hatchery performance (e.g. contribution to a fishery and escapement or adult returns) changes over time and differs among hatcheries.

Furthermore, controversy exists about the potential negative impacts of hatchery salmon on wild salmon (Lynch and O'Hely 2001, Frankham 2008). From an ecological perspective, wild salmon may be impacted by hatchery salmon because of increased competition, predation, disease transfer, and parasite loads (Waples *et al*. 2020). From a genetic perspective, population fitness may decrease because of hatchery salmon breeding and interbreeding with wild salmon in the natural environment, such that traits that are maladaptive in the natural environment are introduced (Araki and Schmid 2010). In Canada, the *Wild Salmon Policy* states that the conservation of wild salmon, their habitats, and their genetic diversity is the highest priority for resource management decision making (DFO 2005). Washington state has a similar commitment to conserve the genetic integrity of wild salmon populations (Larson *et al*. 2018) or minimize hatchery salmon impacts to natural populations (Anderson *et al*. 2020). There is thus a need for managing salmon hatcheries in a manner that is consistent with the *Wild Salmon Policy*. To this end, the salmon enhancement program through the use of hatcheries has the objectives of minimizing negative environmental and ecological impacts on other fish stocks (MacKinlay *et al*. 2004).

Hatchery salmon generally have lower survival than wild salmon (Fraser 2008, Christie *et al*. 2014). For example, hatchery salmon survival appears to be about half that of wild salmon based on the rate of abundance decline for coho salmon in the Strait of Georgia, i.e. 130,000 vs. 60,000 fish/year (Beamish *et al*. 2010). Early marine survival is linked with smolt status as they transition to saltwater environments, for example, pre-smolts and de-smolts experience higher mortality rates compared with smolts in salinity challenge trials. Hatchery salmon generally have lower seawater tolerance than wild salmon, possibly because the physiological smolt window is altered in the hatchery environment (e.g. Shrimpton *et al*. 1994, Chittenden *et al*. 2008). Thus, there is a need to improve hatchery salmon survival in the natural environment. Indeed, the salmon enhancement program has the objective of optimizing hatchery salmon survival (MacKinlay *et al*. 2004).

Tools to help increase hatchery salmon survival, as well as explain changes in hatchery performance over time or by differences among hatcheries would be valuable to managers. Although genetics or genomics technologies have been reviewed for fisheries, aquaculture, and biosecurity applications (e.g. Bernatchez *et al*. 2017), they have not been extensively reviewed for use by hatcheries. In particular, the value of genomics technologies appears to be underestimated for resource management, e.g. collection of powerful data to inform management (Bernatchez *et al*. 2017), and one hurdle may be the need for greater communication of the applications or deliverables to user groups (Garner *et al*. 2016). Yet, a survey of hatchery personnel indicated a general desire to incorporate genetic technologies into hatcheries (Fisch *et al*. 2015).

Here, I provide a review of four 'omics' technologies, i.e. genomics, transcriptomics, proteomics, and metabolomics, to see how they may be used to achieve hatchery deliverables, e.g. improvements in hatchery performance and hatchery salmon health and condition. Focussing on salmonids (i.e. salmon, trout, and char species), I searched the scientific literature of the past 10 years with broad key words (i.e. genetics, genomics, transcriptomics, proteomics, and metabolomics) and narrow key words (i.e. parentage-based tagging, transcription, gene expression, qPCR, eDNA, eRNA, and environmental) using the Web of Science and Google Scholar databases. In the remainder of the Introduction below, I provide historical and other background information on different types of hatcheries and their objectives, hatchery genetics-lessons learned, the transition from genetics to genomics, the rise of epigenetics, and an overview of the four 'omics' technologies. A more detailed review of the 'omics' technologies and their applications to hatcheries is provided in Sections 2 to 6. Furthermore, I provide a detailed section (Section 7) on knowledge gaps, highlighting experimental designs using the 'omics' technology with the potential to increase our understanding of how to best improve hatchery performance. Finally, in the last section (Section 8), I provide summary lists of the hatchery deliverables and knowledge gaps.

2. Throughout this review for simplicity, I refer to salmon, trout, and char species collectively as 'salmon'.

1.1 > HATCHERY TYPES

There are three general types of hatchery in operation in BC with differing goals (Table 1); albeit, the three types may represent more of a continuum than discrete groups (Fraser 2008).

Table 1. Summary of three hatchery types.

	AUGMENTATION	SUPPLEMENTATION	CAPTIVE BREEDING
Primary objective	Increase salmon abundance for commercial and recreational fishing opportunities.	Improve the status of an existing population by intentional demographic integration of hatchery and wild salmon.	Prevent imminent extinction. Salmon are unable to survive in the natural environment for at least a portion of their life cycle.
Population status	Stable.	Threatened.	Endangered.
Reproduction in the natural environment	Not a goal *per se* (Araki and Schmid 2010).	A goal.	A goal.
Other goals	Minimize harvest of wild salmon (Bernatchez *et al.* 2017).	-	Maintain genetic diversity and fitness for reintroduction into the natural environment.

These three hatchery types possess differences in their underlying genetic goals and information needs. For example, there can be family-specific contributions to the fishery and escapement, suggesting a genetic basis (Beacham *et al.* 2019b). Augmentation hatcheries may benefit from understanding the genetic basis of higher contribution to the fishery, whereas supplementation and captive breeding hatcheries may benefit from understanding the genetic basis of higher contribution to the escapement or reproductive success. Captive breeding hatcheries are typically dealing with smaller population sizes than supplemental hatcheries based on population status, i.e. endangered vs. threatened. Small populations can lose fitness faster than large populations because of inbreeding depression and the loss of genetic diversity (Frankham 2005), such that genetic factors may be more important for captive breeding hatcheries than supplemental hatcheries.

1.2 > GENETICS LESSONS-LEARNED

Some foundational population genetic principles have been incorporated into current hatchery operations, i.e. using local salmon, avoiding inbreeding depression, and maintaining genetic diversity. However, further improvements in hatchery performance may be made if additional genetics lessons-learned were incorporated into hatchery operations, such as allowing mate choice (i.e. offspring genetic quality) and greater resemblance to the natural rearing environment (i.e. semi-natural rearing techniques).

Non-local hatchery salmon consistently reproduce poorly in the local natural environment, whereas local hatchery salmon generally reproduce better (Araki *et al.* 2008). With non-local hatchery salmon, there is also the risk of outbreeding depression from interbreeding with local salmon (Christie *et al.* 2014). That is, there may be genetic incompatibilities or the dilution of locally adapted genes such that the population fitness may decrease. There is a similar situation with the straying of hatchery salmon into non-local environments (Bernatchez *et al.* 2017). Even with local salmon in hatcheries, reproductive isolation from the local natural population may decrease fitness; this decrease in fitness may be slowed with integration with wild salmon (Frankham 2008). Hatcheries have moved towards local and wild integration for the brood stock and have implemented measures to decrease straying to improve performance and limit population fitness decreases. That is, local salmon are used in at least 95% of the BC hatcheries (MacKinlay *et al.* 2004).

Additionally, inbreeding depression and lower genetic diversity may decrease population fitness (Frankham 2005). Inbreeding depression is caused by the exposure of two recessive alleles associated with deleterious traits, a situation which is more likely to arise from the interbreeding of closely related individuals. Evolutionary potential (i.e. ability to adapt to a changing environment) has a positive relationship to measures of genetic diversity, such as heritability. Albeit, there are ties between inbreeding depression and genetic diversity, i.e. small populations have a higher risk of both inbreeding depression and the loss of genetic diversity than large populations. Hence, inbreeding depression may be more likely for captive breeding hatcheries (Fraser 2008). Although hatcheries generally avoid small founding populations (MacKinlay *et al*. 2004), it may also occur in supplementation or augmentation hatcheries with few founding individuals. Certain captive breeding hatcheries (e.g. winter-run Chinook salmon in California and Inner Bay of Fundy Atlantic salmon, *Salmo salar*, in Nova Scotia) have moved towards genetic measures of relatedness and inbreeding, as well as the utilization of pedigrees to avoid kin breeding (Fraser 2008). Hatcheries have also moved towards breeding techniques, such as single pair matings of females and males for large hatcheries, to limit the loss of genetic diversity. An earlier science review using genetic tools available at the time suggested that most hatcheries maintained genetic diversity (Fraser 2008). However, hatcheries may lose genetic diversity with inbreeding, smaller population sizes, higher variance in family size, and lower gene flow (Anderson *et al*. 2020). BC hatcheries use mating procedures to limit the loss of genetic diversity, e.g. matrix spawning for small groups (MacKinlay *et al*. 2004).

The mating procedures used to maintain such genetic diversity may be at the expense of offspring genetic quality (Neff *et al*. 2011). That is, the prevention of natural mate choice may inhibit sexual selection and the corresponding genetic benefits, e.g. increased immunity from major histocompatibility complex (MHC) combinations (Fraser 2008, Fisch *et al*. 2015). In more detail, trait genetic architecture has two components: 'good genes' and 'compatible genes' (Figure 1). A good gene contains an allele that increases offspring fitness independent of the genome, such that a good gene displays additive genetic variance, which is heritable (Neff and Pitcher 2005). For example, males with a good gene will produce offspring with higher fitness than males without the good gene; thus, the mate choice of all females should be for the male with the good gene. In contrast, a 'compatible gene' contains alleles that increases offspring fitness dependent on the genome, i.e. a specific genotype, such that a compatible gene displays non-additive genetic variance. For instance, males with compatible genes will produce offspring with either higher or lower average fitness dependent on the female pairing; thus, mate choice of each female should be for a different male, such that the compatible genes are in the direction of higher offspring fitness. The trait genetic architecture of good genes (additive genetic variance) and compatible genes (non-additive genetic variance) of salmon offspring can be determined using a full factorial breeding design. By allowing the mating of low-quality individuals ('poor genes') or the mating of unsuited individuals ('incompatible genes') hatcheries may be producing offspring with lower fitness that would have been minimized by sexual selection (Neff *et al*. 2011). Although there may be logistical constraints, hatcheries may benefit from enabling mate choice, e.g. allowing females to select from multiple males in a spawning channel, to increase offspring genetic quality.

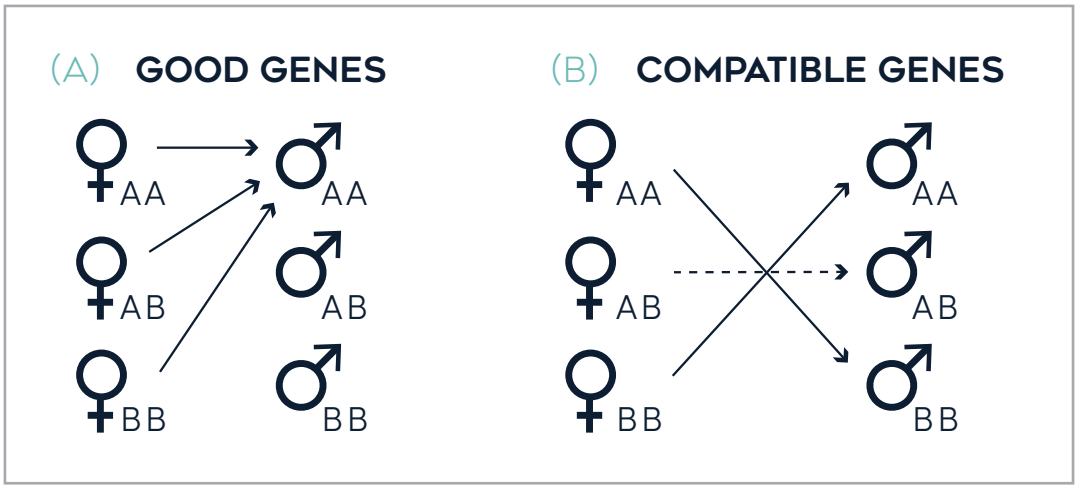

Figure 1.
Genetic quality as either good genes or compatible genes (source Neff and Pitcher 2005).

Regardless of the hatchery efforts to limit population fitness decreases, there remains a debate as to whether the boost in population size outweighs any other potential genetic consequence (Lynch and O'Hely 2001, Frankham 2008). Populations may have detrimental genetic change because of antagonistic selection, i.e. rare alleles that are detrimental in the natural environment but are beneficial in the hatchery environment. In other words, in the hatchery, there is the absence of predators, as well as the provisioning of food, medical treatment, and mate acquisition, such that hatchery selection can be considered relaxed or maladaptive relative to the natural environment. This hatchery selection may be termed domestication selection — any change in the selection regime of a cultured population relative to that experienced by the natural population (Waples 1999). Although the genetic effects of domestication selection may be offset by natural selection eliminating the hatchery salmon not suited to survive in the natural environment, the traits exposed to natural selection during later life stages after release (i.e. smolt to adult) may not be the same as the early life stages (i.e. egg to juvenile), such that the mortality experienced after release may not fully offset any genetic change to populations. Nonetheless, there are methods to limit the genetic changes imposed by the hatchery environment. That is, minimize (1) selection intensity, (2) genetic diversity, (3) effective population size, and (4) number of generations (Frankham 2008), as well as (5) equalize family sizes (Fisch *et al*. 2015). There may be logistical constraints to applying some or all these methods in a hatchery. However, the method that may have the most benefits is minimizing domestication selection intensity via the use of semi-natural rearing techniques, resembling more the natural environment, because of the increases in post-release juvenile survival (Maynard *et al*. 2004, Näslund and Johnsson 2016). Ultimately, there is a general promotion in the literature for semi-natural rearing techniques to improve hatchery salmon survival and potentially limit genetic consequences on populations (Lynch and O'Hely 2001, Tave and Hutson 2019).

1.3 > GENETICS TO GENOMICS

The difference between genetics and genomics can be loosely defined by the number of markers: genetics (10s to 100s of DNA markers) and genomics (> 1000s of DNA markers). Recently (past 10 years) genomics technologies have rapidly replaced traditional genetic markers, e.g. microsatellites (Bernatchez *et al*. 2017). Albeit, genomics can be tied to genetics, e.g. a genomic scan can identify thousands of markers that is then simplified to a smaller number of the most powerful markers for addressing a given question. Genomics generally use next generation sequencing (NGS) — a rapid and high-volume DNA sequencing technology (McMahon *et al*. 2014, Garner *et al*. 2016, Waples *et al*. 2020). NGS can identify single-nucleotide polymorphisms (SNPs) — a single DNA base pair that is variable. Most SNP markers represent only two variant alleles for a gene (Waples *et al*. 2020). SNP markers can be examined on a SNP array — an analysis platform used to assess SNP genotypes in a high-throughput manner, even millions of markers. Also, SNP markers can be examined using genome-resequencing technology. Restriction site-associated DNA sequencing (RAD-seq) typically covers a small proportion of the genome (< 10%) such that it is economical for 2,000–6,000 SNP genes. Pool-seq can cover a higher proportion of the genome because a pool of DNA from individuals is fully sequenced together to identify allele frequencies. As the cost of sequencing decreases, soon whole genomes at the individual level (instead of the pool level) may be common. Future projections are that genomic technologies will be integrated into resource management (Bernatchez *et al*. 2017). For example, genomics technologies will be used to monitor important individual and population level processes (Waples *et al*. 2020).

Photo by: Mitch Miller

1.4 > RISE OF EPIGENETICS

Despite using local and wild integrated brood stock, hatchery salmon generally have lower reproductive success or other fitness-related traits than wild salmon in the natural environment (Fraser 2008, Christie *et al*. 2014). For a captive breeding hatchery, one generation in the hatchery found no decrease in reproductive success of hatchery salmon in the wild, but there can be decreases for later closed generations, i.e. ~40% lower reproductive success per captive-reared generation (Araki *et al*. 2007). Overall, for augmentation or supplementation hatcheries using local and predominantly wild-origin broodstock, Christie *et al*. (2014) found that: (1) hatchery salmon had about half the reproductive success of wild salmon, (2) the effect was more severe for hatchery males than females, and (3) all species were affected by the hatchery environment. The mechanism for decreased hatchery salmon performance does not appear to be underlying DNA differences because genomic scans have found few gene differences between hatchery and wild salmon (e.g. Mäkinen *et al*. 2015, Christie *et al*. 2016, Le Luyer *et al*. 2017, Gavery *et al*. 2018, 2019). For example, using thousands of SNP markers (1751–4733), there were few outlier gene differences between hatchery and wild Atlantic salmon even with 10 generations of hatchery rearing (Mäkinen *et al*. 2015). It appears as if detecting genetic signals of the hatchery environment maybe challenging unless the selection intensity is strong, and the traits show simple inheritance patterns.

One branch of genomics with particular relevance to hatcheries that has risen rapidly is epigenetics (i.e. chemical tags on DNA). Epigenetic programming differences may be a mechanistic explanation for the decreased hatchery salmon performance despite a similar DNA background as wild salmon (Fraser 2008, Tave and Hutson 2019). Epigenetic programming is non-coding changes to DNA that can have profound trait or phenotypic effects and may span multiple generations (Best *et al*. 2018). Specifically, epigenetic programming causes changes in gene expression without changing the underlying DNA sequence. Extrinsic (e.g. stressor and nutrition) and intrinsic (e.g. nervous, endocrine, immune, and metabolite system) factors can stimulate epigenetic effects (Best *et al*. 2018). The molecular basis of epigenetic programming is gene expression activation or silencing (Jonsson and Jonsson 2014). Gene expression can be activated by decreased DNA methylation or increased histone modification and silenced using the opposite. Another mechanism regulating the abundance and translation of gene expression is microRNAs (Best *et al*. 2018). The nature and extent of epigenetic effects has only recently been explored in salmon at the molecular level (Waples *et al*. 2020).

The environmental differences between hatchery and natural rearing is the likely explanation for epigenetic programming differences. Early life stages of salmon appear to be critical periods for development, as well as epigenetic programming (reviewed by Jonsson and Jonsson 2014). The epigenetic programming can be mediated by conditions experienced by the egg, alevin, fry, or juvenile (i.e. environmental effect) or conditions experienced by the parents (i.e. parental effect). Known traits influenced by environmental effects in salmon are social behaviours, age-at-maturity, and migration life history. For example, there may be a difference in the juvenile competitive abilities between hatchery and wild salmon (Weber and Fausch 2003, Huntingford 2004). Later in the natural environment during the spawning period, hatchery males have been found to incur higher wounding and mortality, as well as lower reproductive success than wild males (Fleming *et al*. 1997). A lower intensity of female than male competition was probably responsible for the lack of differences in performance between hatchery and wild females. Furthermore, higher growth rate is associated with the hatchery than natural environment. Younger (parr) age-at-maturity is associated with fast initial growth, whereas as later (adult) age-at-maturity is associated with slow growth (e.g. Vøllestad *et al*. 2004). Although there is a precedence for epigenetic programming differences between hatchery and wild salmon, the few studies at the molecular level confirming such differences are the beginning of a new research initiative that may provide guidance on how to best improve hatchery performance (Waples *et al*. 2020).

1.5 > FOUR 'OMICS' TECHNOLOGIES

Beyond genomics technologies, there are another three 'omics' technologies that cover the genotype to the phenotype (Alfaro and Young 2018): transcriptomics, proteomics, and metabolomics (Figure 2).

Genomics aims to sequence DNA and utilize the information provided by a genome that is ideally annotated, i.e. the gene function or relationship to a phenotype or salmon trait.

Transcriptomics aims to sequence messenger RNA (mRNA) or transcripts (coded by DNA) to highlight which genes are changing in expression (i.e. upregulated or downregulated) for a given condition, such as a change in the environment.

Proteomics aims to study proteins. In particular, certain proteins are translated from mRNA (previously transcribed from DNA), and changes in protein expression can reflect responses to a given condition, such as a change in the environment.

Metabolomics is the study of small molecules (< 1 kDa or kilodalton) or metabolites, with enzymes (proteins) involved in the regulation and operation of metabolic pathways. In particular, changes in metabolite expression may provide the best mechanistic explanations for what is happening at the physiological level.

Currently, research using 'omics' technologies is common because of its high-throughput nature, e.g. many markers. In addition, the analyses enable comparisons to databases containing functional information for genes, transcripts, proteins, and metabolites. The four 'omics' technologies can also use non-lethal tissue samples such as fin clips, gill, blood, seminal fluid, and ovarian fluid. Altogether, using more than one 'omics' technology can provide a wider mechanistic vision on salmon physiology in response to a given condition (Forné *et al*. 2010).

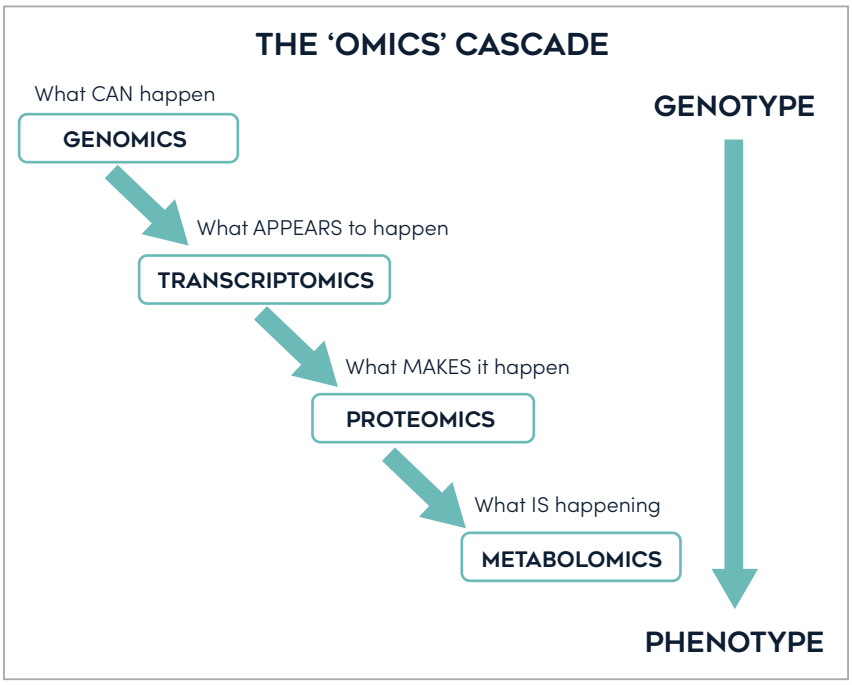

Figure 2.
Four 'omics' technologies along the gradient of genotype to phenotype (source Alfaro and Young 2018).

Out of the four 'omics' technologies, there are plenty of examples of applications using genomics and transcriptomics technologies; however, proteomics and metabolomics technologies are newer and there are fewer examples of applications (Alfaro and Young 2018). Furthermore, proteomics and metabolomics technologies have been mostly examined for aquaculture (Rodrigues *et al*. 2012, Alfaro and Young 2018), and have not yet transitioned to fisheries or hatcheries. However, proteomics and metabolomics have ties to salmon health (e.g. disease and immunology), reproduction (e.g. egg and sperm quality), and other fitness-related traits (Alfaro and Young 2018), which are also important to fisheries and hatcheries.

The current salmon research focus on genomics and transcriptomics is interesting considering their poorer ability to provide mechanistic explanations relative to proteomics and metabolomics. Proteomics is more mechanistic (closer to the phenotype) than genomics and transcriptomics because it incorporates post-translational modifications and protein degradation (Rodrigues *et al*. 2012). Although extrinsic factors (e.g. environmental stressors) can cause changes in the transcriptome (i.e. gene expression) and proteome (i.e. protein expression), such changes are amplified at the metabolome (i.e. metabolite expression), implying that it is the most sensitive of the four 'omics' technologies (Lankadurai *et al*. 2013, Alfaro and Young 2018). Furthermore, metabolite functions (e.g. biochemical networks) are conserved across species such that analytical tools are easily transferable, whereas genes, transcripts, and proteins may differ in structure and functions across species such that a reference (e.g. species genome) is needed for the analyses (Jones *et al*. 2013, Alfaro and Young 2018). Nonetheless, there are disadvantages for proteomics and metabolomics technologies relative to genomics and transcriptomics technologies. There is no amplification method for proteins and metabolites in contrast to DNA and RNA, such that examination of low numbers is problematic (Zhou *et al*. 2012). The metabolome is also highly sensitive to intrinsic factors, i.e. age, diet, diurnal cycle, and reproductive cycle (Jones *et al*. 2013, Alfaro and Young 2018), as well as handling stress or anaesthetics (Young and Alfaro 2018) relative to the genome, transcriptome, and proteome. These intrinsic factors may produce noise that make it more difficult to detect metabolite expression changes to a given condition. Despite certain disadvantages, research using novel proteomics and metabolomics technologies may become common given the tighter mechanistic link to salmon physiology.

In the following sections, I provide examples of hatchery deliverables that may improve hatchery performance using the molecular information derived from the four 'omics' technologies. How the technology works and analytical methods, such as reagents and statistical analyses, are not described in detail, as the focus is on highlighting the hatchery deliverables. Important considerations are emphasized in bold. Where appropriate, I specify the hatchery type, i.e. augmentation, supplementation, or captive breeding, for a deliverable. Given the abundance of salmon research using genomics and transcriptomics technologies, there are clear examples of hatchery deliverables. In contrast, the proteomics and metabolomics technologies are newer, and most salmon research is described for aquaculture. Yet, the aquaculture deliverables may be transferable to hatcheries, e.g. improvements in salmon health and condition. In a couple of cases, there are potential hatchery deliverables using environmental RNA, which are theoretical as they require further research to demonstrate feasibility.

2. BROAD-SENSE GENOMICS

Broad-sense genomics is described as an improvement of traditional genetics analyses (Waples *et al*. 2020). That is, such analyses have higher statistical power because of the higher number of markers, i.e. SNP markers. Examples of broad-sense genomics deliverables are population structure, evolutionary history, parentage analysis, hybridization and admixture analysis, inbreeding analysis, effective population size, and species identification. For fisheries, the most common broad-sense genomics deliverable is the delineation of the spatial extent and structure of populations (Bernatchez *et al*. 2017). For hatcheries, identified broad-sense genomics deliverables are (1) fishery and escapement contributions, (2) other hatchery metrics, (3) trait architecture, and (4) genetic connection.

2.1 > FISHERY AND ESCAPEMENT CONTRIBUTIONS

The original incentive for parentage-based tagging (PBT) was the increased demand from fisheries managers for precise information on mixed fisheries stock contributions (Steele *et al*. 2019). Across the Pacific Rim, salmon hatcheries cover many stocks, such that the hatchery source can be utilized for examining stock, as well as the direct contribution of the hatchery itself to the fishery. For PBT, a tissue sample (typically a fin clip) is collected from all individuals forming the hatchery brood stock. The tissue sample is preserved dry on absorptive paper, e.g. Whatman chromatography paper that can hold 50–100 samples. The tissue samples of the parents are genotyped using hundreds to thousands of SNP markers; the number of SNP markers needed to provide enough resolution increases with the number of parents, which can be determined using parentage analysis simulations. Given that there are genotypes for all the hatchery parents, the resulting offspring are all genetically tagged, even millions of juveniles upon hatchery release. The later captured offspring are assigned back to their parents using another tissue sample, e.g. fin clip, its genotype, and a parentage analysis computer programs that can handle hundreds to thousands of markers. The hatchery assignment accuracy is near perfect (> 99%) for steelhead trout using > 100 SNP markers (Steele *et al*. 2013), coho salmon using 304 SNP markers (Beacham *et al*. 2017, 2019a), and Chinook salmon using 321 SNP markers (Beacham *et al*. 2018). Given the large amount of data (e.g. genotypes for many SNP markers and many individuals), a large repository or database is needed for handling the information (Steele *et al*. 2019). Also given the logistics, e.g. hatcheries are coast-wide, there is a need for a committee or organization to oversee the PBT sampling program.

Ideally, hatchery and wild salmon should be externally identifiable within the mixed fishery or spawning run (escapement). For example, most juvenile coho salmon in southern BC, Washington, and Oregon receive an adipose fin clip before hatchery release (Beacham *et al*. 2017). Certain hatcheries coast-wide also provide an adipose fin clip to juvenile Chinook salmon receiving a coded-wire tag (CWT); however, the CWT tagging rate is approximately 10% of hatchery salmon (Beacham *et al*. 2018). Presently, the main reason for the adipose clipping is mass mark hatchery salmon for selective harvest to sustain fishing and decrease fishing pressure on wild (unclipped) salmon. However, the mass marking also enables hatchery salmon to be used for PBT analyses (Beacham *et al*. 2019b). Hatchery salmon with a low parentage assignment probability (e.g. < 85%) using PBT can still be assigned to a stock using a genetic stock identification (GSI) method, e.g. coho salmon (Beacham *et al*. 2017, 2019a) and Chinook salmon (Beacham *et al*. 2018, McKinney *et al*. 2019). Given that PBT can genetically tag all hatchery offspring and has a high assignment accuracy, the hatchery and stock information offers coverage that is greater than with the current CWT programs (Steele *et al*. 2013, Beacham *et al*. 2017, 2018, 2019a). However, there is currently no coast-wide sampling protocol to utilize the higher coverage offered by PBT than by CWT.

Overall, **PBT can provide an estimation of hatchery contribution and families within hatcheries to fisheries and escapement.** An example of a hatchery performance deliverable identified using PBT is that Capilano Hatchery coho salmon contributed to 30% (1,904 out of 6,409) of hatchery-origin (clipped) coho salmon during the 2017 Strait of Georgia recreational fishery (Beacham *et al*. 2019a). An example of a deliverable is that the coho salmon parents of the earlier run produced more returning offspring (escapement) in 2016 than the later run, i.e. 3.06 vs. 1.87 offspring/spawner for Capilano brood stock in 2014 (Beacham *et al*. 2019b). The offspring/spawner ratio is a measure of reproductive success. These deliverables have different implications for different hatchery types. The fishery deliverable may have higher relevance to augmentation hatcheries with an objective of maximizing fishing opportunities (Beacham *et al*. 2019b), whereas the escapement deliverable may have higher relevance to supplementation and captive breeding hatcheries with an objective of reproduction in the natural environment (Steele *et al*. 2013, Waples *et al*. 2020). Other uses of PBT within hatchery populations are described in the following sections, e.g. release group performance (section 2.2) and trait architecture (the genetic basis of salmon traits, section 2.3).

2.2 > OTHER HATCHERY METRICS

Parentage assignment using PBT information, including other SNP genotype information, can also be linked to hatchery records for developing other hatchery metrics. Beyond hatchery source, hatchery records on parents and their offspring typically include information on species, sex, brood year, and release group (Steele *et al*. 2013, 2019).

PBT Information

Hatchery source data can be used for quantifying the amount of **straying**. Salmon straying is considered undesirable because interbreeding may decrease the fitness of locally adapted populations (Christie *et al*. 2014, Bernatchez *et al*. 2017). An example of a deliverable is that PBT information identified that early maturing (i.e. jack) hatchery coho salmon were straying to non-local rivers for Vancouver Island and southern BC (Beacham *et al*. 2018). Another example deliverable from PBT analysis is that 1% (out of 4,447 individuals) of the coho salmon brood stock in 2017 for 19 hatcheries was strays, with the highest value (40%) between geographically close (< 1 km) Inch Creek and Norrish Creek (Beacham *et al*. 2019b).

Brood year records can be used for offspring **age determination**. Specifically, the capture time of the offspring is compared to the parental spawning time. May be valuable for revealing salmon age without relying on scale or otolith readings. An example of a deliverable is that PBT information confirmed the age of putative 'jimmy' and 'jack' male Chinook salmon, i.e. age-1 and age-2, respectively (Beacham *et al*. 2018).

Release records can be used for evaluating **group performance**. With careful hatchery planning to not divide a family across different groups, the relative performance of each group can be compared using the PBT family assignments (Figure 3). The performance measures can be short-term (e.g. juvenile early marine survival) or long-term (e.g. adult contribution to the fishery or escapement).

Pedigrees covering several generations can be constructed using parental records and offspring assignments. Pedigrees can be used to visualize family sizes (i.e. offspring counts). Equalizing the family sizes, can limit the loss of genetic diversity (Fraser 2008) and genetic adaptation to captivity (Fisch *et al*. 2015). Pedigree relationships can also provide expected inbreeding coefficients, i.e. probability that the two alleles are identical by descent if offspring are produced, and pairwise relatedness values, i.e. fraction of alleles shared (Wang 2016). Realized or accurate versions are described below. For example, the expected inbreeding coefficient (F) for the offspring produced using brother/sister mating is 25%, half-brother/half-sister mating 12.5%, and first cousin mating is 6.25%, which is related to the predicted decrease in offspring fitness because of inbreeding depression. The expected pairwise relatedness (r) value between full siblings is 1/2, between half siblings is 1/4, and between first cousins is 1/8. Pedigrees, inbreeding coefficients, and pairwise relatedness may have applications to supplemental and captive breeding hatcheries to avoid highly unequal family sizes and inbreeding depression (Fraser 2008).

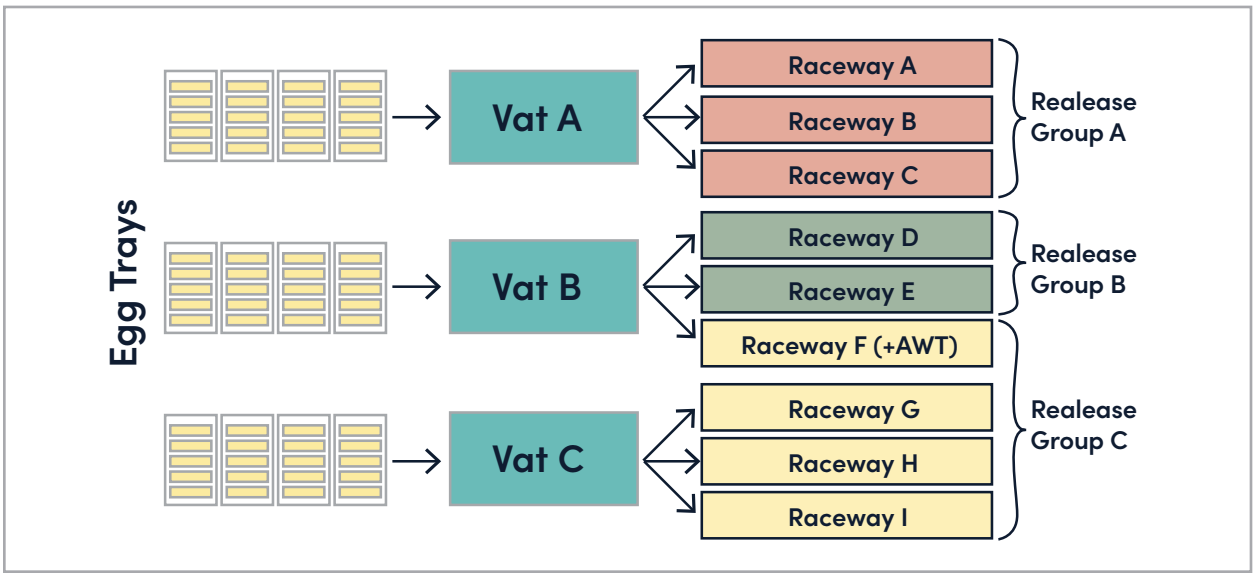

Figure 3.
Family loading plan for release groups identified by parentage-based tagging (source Steele *et al*. 2019). Generally, specific families are associated with specific release groups. Given that families (Vat B) are mixed with a different release group (Release Group C), another bulk tagging method is necessary (agency wire tag, AWT). Note that AWT is equivalent to CWT as every coded-wire tag is specific to an agency.

SNP Genotype Information

Sex records can be confirmed for brood stock using genetic **sex determination**. Certain SNP markers are associated with sex-linked genetic regions of the genotype, such that sex information can be revealed. These SNP markers are often examined for PBT (Steele *et al*. 2019). Other applications included general sex determination for unclear life stages, e.g. juveniles.

The species records can be used for determining the amount of **non-target species** or **inter-species hybridization**. Inter-species breeding is considered undesirable because of the potential decreased fitness of hybrid salmon, e.g. intrinsic outbreeding depression from genetic incompatibilities, such as chromosome rearrangements (Allendorf *et al*. 2001). Furthermore, hybrid salmon are of little conservation value relative to pure species, e.g. Westlope cutthroat trout (*O. clarkii lewisi*) × rainbow trout (*O. mykiss*) hybrid (Allendorf *et al*. 2001), which may have relevance to supplementation and captive breeding hatcheries dealing with threatened or endangered salmon populations. There are SNP genotypes to discern coho salmon, Chinook salmon, sockeye salmon, chum salmon, pink salmon, masu salmon (*O. masou*), rainbow (steelhead) trout, cutthroat trout, and Atlantic salmon (Beacham and Wallace 2019). Examples of deliverables are the percentage of non-target species in the brood stock and the percentage of inter-species hybridization for offspring in the hatchery (Beacham and Wallace 2019).

SNP markers across the genome can be used to provide *realized* **inbreeding coefficients** and **relatedness values**. Both metrics generally utilize the proportion of homozygous alleles and are considered more accurate than the *expected* versions above (Wang 2016). For example, the expected relatedness (r) value of 1/2 between full siblings is considered an average. That is, there is random sampling of alleles during fertilization, such that certain sibling pairs are genetically more similar or more dissimilar. Genomics data can be used to directly measure these genetic relationships for pairs of individuals. Given that SNP markers are already used for PBT to provide parentage assignments, which may be used to construct pedigrees, the existing SNP genotype information may also be used to provide realized inbreeding coefficients and relatedness values between all possible pairs of individuals. As described above, supplemental and captive breeding hatcheries may use the genomic data to avoid inbreeding depression (Fraser 2008).

Overall, PBT information coupled with hatchery records can provide performance indicator metrics such as straying, age determination, or group contributions to the fishery and escapement. Genotypes at specific SNP markers can also provide sex determination (same with PBT), as well as hatchery performance metrics for non-target species and inter-species hybridization contributions. The parentage assignments and SNP genotypes across the genome can be used to construct pedigrees, as well as provide inbreeding coefficients and pairwise relatedness values. More broadly, PBT now provides the ability to study (1) the **performance of hatchery salmon**, including individuals and families and (2) **the environmental effect** of different hatchery culture practices using groups.

Photo by: Mitch Miller

2.3 > TRAIT ARCHITECTURE

Similarities between parents and offspring as well as among siblings imply that certain salmon traits have a genetic basis. For example, coho salmon PBT results in British Columbia have highlighted a similar return timing between parents and offspring for Chilliwack River, i.e. 62% returning in the same month (Beacham *et al*. 2019b). PBT results also indicated family level (sibling) associations, for traits such as fishery contribution and escapement contribution (Beacham *et al*. 2019b). Determining whether a given salmon trait has a heritable component is important to assessing evolutionary potential, as the response to selection (e.g. fishing or climate change) has a direct relationship to the amount of additive genetic variance or narrow-sense heritability (Falconer and Mackay 2009). Trait architecture, i.e. genetic and environmental contributions, provide a greater scope on the factors explaining a given salmon trait. More broadly, salmon traits can also be influenced by genotype × environment interactions and genetic correlations.

Salmon trait architecture, including heritability, is best measured using a full factorial mating design. That is, females and males are mated in all possible combinations with the offspring reared in a common environment (Lynch and Walsh 1998). PBT can then be used for identifying the family relationships. The minimum full factorial design is a block of two females by two males (2 X 2); albeit, to achieve sufficient statistical power, more parents, e.g. 5 X 5, and/or more blocks, e.g. several 2 X 2 blocks, may be needed, which can be estimated using power analyses (Houde and Pitcher 2016). Beyond an estimate of additive genetic variance (i.e. heritability), the offspring trait variance can be partitioned into estimates of non-additive genetic variance (i.e. dominance and epistasis[3]) and maternal variance (Table 2). The latter two can also influence evolutionary potential under certain circumstances. Non-additive genetic variance can be converted to additive genetic variance, such as during a population bottleneck (Carson 1990). Maternal effects (genetic and environmental) can modify the rate and direction of a change in response to selection based on the correlation between maternal and offspring traits and the phenotypic plasticity[4] of female traits (Mousseau and Fox 1998, Räsänen and Kruuk 2007). Furthermore, the relative amount of additive and non-additive genetic variance can reveal mate choice for 'good genes' or 'compatible genes', respectively (Neff and Pitcher 2005). An example of a deliverable from this sort of analysis is that the trait of emergent fry body size was explained by an average 5% additive genetic variance, 14% non-additive genetic variance, and 24% maternal variance across two populations of Atlantic salmon populations using a block of 5 X 5 per population (Houde *et al*. 2013, 2015). The small additive genetic effects suggest that this trait may not evolve rapidly in response to selection pressures. The higher non-additive than additive genetic effects imply mate choice for 'compatible genes', such that this trait is influenced more by the suitability of female-male pairing rather than male quality *per se*.

Table 2. Summary of results for Chinook salmon alevin survival using a full factorial breeding design (source Houde and Pitcher 2016).

EFFECT	VARIANCE	PHENOTYPE
Dam	0.788	17.8%
Sire	0.166	3.8%
Dam X Sire	0.167	3.8%
Residual	3.294	
Total	4.415	

ARCHITECTURE	VARIANCE	PHENOTYPE
Maternal	0.622	14.1%
Additive	0.664	15.0%
Non-additive	0.668	15.1%

Note: A random effects model, similar to two-way analysis of variance (ANOVA), contained random effects for dam, sire, and dam × sire interaction. Maternal variance is the dam variance minus the sire variance. Additive genetic variance is four times the sire variance. Non-additive genetic variance is four times the Dam X Sire variance. Phenotype is the percentage out of trait total variance. Data source is Pitcher and Neff (2007). Calculations source is Lynch and Walsh (1998).

3. Dominance is an allele that increases offspring fitness when in a specific genotype at the same gene. Epistasis is an allele that increases offspring fitness dependent on the genotype at another gene.
4. Ability of genotypes to produce different phenotypes when exposed to different environmental conditions

Few studies have examined genotype × environment (G X E) interactions, i.e. the fitness of a gene or genotype in more than one environment. A familiar example of a G X E interaction in salmon is local adaption, i.e. local populations have higher fitness in their local than non-local environment (Figure 4). An estimate of the amount of a G X E effects can be provided with independent families (e.g. single pair matings) with the offspring split across the environments (Lynch and Walsh 1998). For example, for hatchery release, the number of individuals for a family is divided equally between environments. Independent families can also be represented by the diagonal families of the full factorial breeding design (Wellband *et al.* 2018), such that each dam and each sire is represented once.

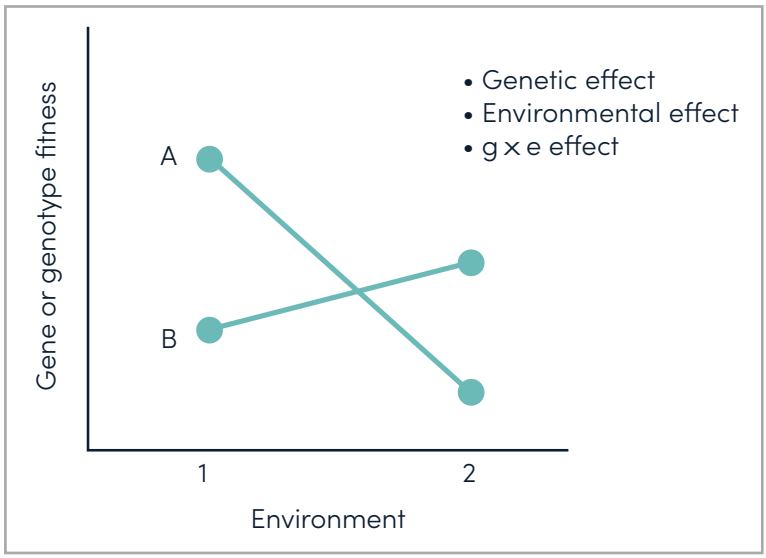

Figure 4.
Local adaptation genotype × environment (G X E) interaction. Population A has higher fitness in its local environment 1. Similarly, population B has higher fitness in its local environment 2 (source Neff and Pitcher 2005).

The offspring trait variance can be partitioned into estimates of genetic variance, environmental variance, and G X E interaction variance (Table 3). PBT can be used for family assignments within and across environments, opening doors for further studies on the extent of G X E interactions. The studies can include different natural environments (vs. laboratory environments), providing the most relevant information of the quality of a gene or genotype for hatchery salmon (Neff and Pitcher 2005).

Table 3. Summary of results for barley varieties yields examining the genotype × environment interaction (source Lynch and Walsh 1998).

EFFECT	VARIANCE	PHENOTYPE
Genotypes	0.0235	8.5%
Environment	0.1510	54.6%
G X E	0.0160	5.8%
Residual	0.0365	
Total	0.2770	

Note: A random effects model, similar to two-way analysis of variance (ANOVA), contained random effects for genetic, environment, and genotype X environment (G X E) interaction. A mixed-effects model can also be used with a fixed effect for environment and random effects for genetic and G X E (Lynch and Walsh 1998); the variance estimate of fixed effects can now be extracted from mixed-effects models (e.g. Nakagawa and Schielzeth 2013)

Beyond heritability, often traits can be genetically correlated and may influence breeding selection. Genetic correlations between traits may occur because a single gene influences several traits (i.e. pleiotrophy) or a group of alleles that tend to be inherited together (i.e. linkage disequilibrium) (Lynch and Walsh 1998). A simple method to estimate a genetic correlation is a regression of average parent values for trait 1 and average offspring values for trait 2. PBT can be used for assigning offspring to parents, and trait measures can be collected for the parents and offspring, e.g. age and body size at maturity. The correlations can be positive, facilitating breeding selection or negative, impeding breeding selection. That is, breeding selection for trait 1 could lead to a favourable or unfavorable response for trait 2 based on the direction of the genetic correlation. Genetic correlations have also been considered for fishing selection impacts on salmon. For example, simulated Chinook salmon mean age response generally depended on the strength of the genetic correlation to length and the fishery selection imposed, e.g. size-based vs. terminal (Eldridge *et al*. 2010, Bromaghin *et al*. 2011). Accurate estimates of genetic correlations between salmon traits important to hatcheries is key for predictions of trait responses to breeding or fishing selection.

Overall, examining salmon trait architecture requires multiple matings of females and males, i.e. single pair matings do not provide the necessary information, with PBT used for identifying the family relationships. Beyond the heritability of a given trait, the architecture estimates can provide information on evolutionary potential and mate choice type. PBT can also be used for examining genotype × environment interactions using the natural environment and genetic correlations for predictions of trait responses to selection. More broadly, PBT now provides the ability to study salmon traits (phenotypes), including the genetic basis using hatchery-produced families (genotypes). Ultimately, **the information may be useful for hatchery breeding protocols to increase offspring genetic quality**.

2.4 > GENETIC CONNECTION

Although there may be little genetic (DNA) difference between hatchery and wild salmon within local populations (e.g. Le Luyer *et al*. 2017), genetic differences arising from using non-local hatchery sources can be utilized for measuring hatchery genetic introgression into populations. Low genetic introgression may be considered valuable to hatcheries because it indicates lower non-local (or stray) salmon genetic influences on populations. Using a 57,501 SNP array for steelhead trout, 360 outlier SNP markers were identified that differentiated wild and hatchery salmon (Larson *et al*. 2018). The non-local hatchery salmon were mainly form one area of South Puget Sound (i.e. Chamber Creek), whereas the populations sampled were from other areas of Puget Sound. Hatchery genetic introgression for chum salmon of Prince William Sound, Alaska was measured using 135 SNP markers, archived scale samples (1964–1982), contemporary samples (2008–2010), and a sink-source model that could detect even weak genetic introgression (Jasper *et al*. 2013). A temporal shift towards a higher non-local hatchery allele frequency in the populations indicated stronger introgression. Interestingly, stronger introgression was associated with a similar spawning date between the hatchery salmon and the populations rather than the proximity to the hatchery or the intensity of straying. An example deliverable is a stronger hatchery chum salmon introgression rate (m) for Wells River, i.e. 0.257 (95% CI: 0.209–0.328) than Constantine Creek, i.e. 0.011 (0.004–0.017) (Jasper *et al*. 2013). Such information may have applications to augmentation hatcheries, avoiding hatchery salmon reproduction in the natural environment, as introgression is a proxy of hatchery salmon genetic influences on wild salmon.

Hatcheries can minimize the genetic divergence between local hatchery and wild salmon. Two brood stock management options were compared using 48,528 SNP markers and a calculation of genetic divergence between wild and hatchery Chinook salmon (Waters *et al*. 2015, 2016). The 'integrated' option (i.e. wild returning salmon were collected for hatchery brood stock every year) displayed little genetic divergence over four generations, whereas the 'segregated' option (i.e. hatchery-origin returning salmon were collected for brood stock every year) displayed such divergence because of genetic drift[5]. The amount of genetic divergence increased with every generation for the segregated option. An example deliverable is lower genetic differentiation (FST) for the integrated option (i.e. 0.0033) than the segregated option (i.e. 0.0125) during the fourth generation (Waters *et al*. 2016). Although using all wild salmon for the brood stock may not be feasible for hatcheries, different proportions of wild salmon representation in the hatchery brood stock should be evaluated for their potential to limit genetic divergence. Such information may have applications to supplemental and captive breeding hatcheries, with a goal of hatchery salmon breeding in the natural environment, as to limit genetic divergence because of domestication selection or genetic drift.

Overall, SNP markers, also available with PBT, can provide very **accurate estimates of hatchery genetic interactions with natural populations**.

5. Genetic drift is the change in the frequency of an allele because of random sampling every generation.

3. NARROW-SENSE GENOMICS

Narrow-sense genomics is defined as novel genetic analyses that were not possible before with traditional genetic analyses (Waples *et al*. 2020). Examples of narrow-sense genomics deliverables are linking genotypes and phenotypes (e.g. gene biomarkers), genomic signals of adaptation analysis, genomic evolution analysis, admixture/outbreeding depression analysis, epigenetic programming analysis, and historical effective population size estimate. Such genomic analyses have recently become common because it has taken time to develop the necessary tools to fully understand the complexities of the salmon genome. In particular, there is more detailed information now available on the structure and function of genes. Furthermore, reference genomes for Pacific salmon, i.e. Chinook salmon (Otsh_v1.0, January 2018), sockeye salmon (Oner_v1.0, June 2019), and coho salmon (Okis_v2.0, November 2019), have only been recently produced because of decreasing DNA sequencing costs. In this section, environmental DNA (i.e. animal release of DNA into their environment, eDNA) is considered as a narrow-sense genomics technology given its novelty and growing acceptance in resource management (Cristescu 2019). In particular, as of May 2019, there is a new journal *Environmental* DNA dedicated to the study and use of environmental DNA for basic and applied science. For hatcheries, identified narrow-sense genomics deliverables are (1) gene biomarkers, (2) trait epigenetic programming, (3) epigenetic programming differences, and (4) species distribution and abundance.

3.1 > GENE BIOMARKERS

Certain salmon traits of value to hatcheries are under the control of specific genes that may serve as biomarkers. For example, older (adult) rather than younger (parr) age-at-maturity is a valuable salmon trait for hatcheries, likely because of their smaller body size from a fisheries context (e.g. Middleton *et al*. 2019). Genomics technologies can be used to identify gene biomarkers for salmon traits (Piccolo 2016). That is, statistical models (genome-wide association methods) on genomics data can identify a small number of SNP markers (linked to genes) that explain the most variation for a given salmon trait. Such an approach was used for examining age-at-maturity with 208,704 SNP markers across 57 populations of Atlantic salmon (Barson *et al*. 2015). An example of a deliverable is that a gene biomarker (vestigial-like family member 3, VGLL3) is associated with age-at-maturity. Interestingly, VGLL3 explained a relatively high amount (39%) of the age-at-maturity variation and displayed sex-dependent dominance, i.e. causing younger maturation in males but older maturation in females. Another example of a deliverable is a gene biomarker (oestrogen in breast cancer-like gene, GREB1L) is associated with spawning run timing for steelhead trout (Hess *et al*. 2016). GREB1L explained 46% of the variation in run timing (summer vs. winter). Remarkably the same gene biomarker (GREB1L) also separated spring and fall spawning Chinook salmon (Prince *et al*. 2017). The above studies used RAD-seq, a coarse genomics resequencing method that covered 0.5–2% of the steelhead trout genome. Recent studies using Pool-seq, a whole genome resequencing method, identified additional gene biomarkers of smaller effect that also contributed to run timing for steelhead trout (Micheletti *et al*. 2018) and Chinook salmon (Narum *et al*. 2018).

Captive breeding hatcheries may want to conserve certain locally adapted traits. For example, Lahontan cutthroat trout (*O. clarkii henshawi*) reside in desert areas such that they display high thermal tolerance relative to other populations (Amish *et al*. 2019). Hatchery brood stock was compared to two control lakes using 4,644 SNP markers. A deliverable was that nine SNP biomarkers explained the variation in thermal tolerance (Amish *et al*. 2019). Beyond thermal tolerance, the identified biomarkers were also associated with other important biological functions including immune response, growth, and anaerobic resiliency. Overall, **identifying the gene biomarkers underlying valued salmon traits may serve as targets for hatchery breeding**.

3.2 > TRAIT EPIGENETIC PROGRAMMING

Although specific genes encode particular salmon traits, the expression of these genes can be modified by environmental features that can also influence these traits. For example, younger age-at-maturity is also linked to fast initial growth (e.g. Vøllestad *et al*. 2004), which is typical of hatcheries. Genomics technology can be used for examining differences in DNA methylation patterns (linked to gene expression), as an explanation for the trait differences despite similar genetic background (Jonsson and Jonsson 2014). Such trait epigenetic programming studies have only been possible recently because of the advancements in methods to examine DNA methylation patterns (Gavery *et al*. 2018). Earlier studies used the methylation-sensitive amplified polymorphism (MSAP) method, whereas current studies use the reduced-representation bisulfite sequencing (RRBS) technique, because it provides much higher sensitivity and

resolution than MSAP for examining DNA methylation patterns. Both methods examine the cytosine-phosphate-guanine (CpG) context, the responsive DNA methylation context in vertebrates. An example of a deliverable is that there are 74 differentially methylated regions (DMR) for the testes, 70 DMR for the brain, and 12 DMR for the liver between younger (parr) and older (adult) maturing male Atlantic salmon (Morán and Pérez-Figueroa 2011). The epigenetic effects, i.e. gene expression modifications, were directed at the gonads and brain, but not the liver, which is also supported by another study that identified gene expression patterns of the testes and brain for predicting younger age-at-maturity in male steelhead trout (Middleton *et al*. 2019).

Trait epigenetic programming studies are just beginning for salmon research and there are currently few examples. The two examples above examined salmon of different age-at-maturity within a common environment (Morán and Pérez-Figueroa 2011, Middleton *et al*. 2019). Comparisons of trait epigenetic programming within and between environments (e.g. high and low food availability) would provide more information on the environmental influences on traits at the molecular level. For example, a study examined epigenetic programming of juvenile steelhead trout reared in hatchery tanks with commercial pellets to satiation or a simulated stream with natural invertebrate drift for eight months (Gavery *et al*. 2019). The deliverable was that there were 413 DMR for liver tissue, with the majority (251 DMR) hypomethylated (or activated gene expression) for the hatchery-reared group. Indeed, the differences in early rearing environment stimulated epigenetic programming differences in salmon of similar genetic background. However, this study did not examine the association between the programming and a particular salmon trait. Altogether, **additional research is needed for identifying rearing environment features that contribute to salmon traits, and whether these effects can be managed in the hatchery setting**.

3.3 > EPIGENETIC PROGRAMMING DIFFERENCES

Epigenetic programming differences between hatchery and wild salmon have been suggested as an explanation for the decreased performance (Fraser 2008, Tave and Hutson 2019). An example of a deliverable is that there are 100 DMR for white muscle between hatchery and wild juvenile coho salmon (Le Luyer *et al*. 2017). A functional annotation analysis on the results can reveal differences in key biological processes. For instance, hatchery salmon displayed hypermethylation (or silenced gene expression) relative to wild salmon for biological processes that may decrease juvenile performance. For example, hypermethylation of neuromuscular communication (e.g. Ca2+ levels) which may decrease swimming performance. Another example of a deliverable is that there are 85 DMR for red blood cells and 108 DMR for sperm cells between hatchery and wild returning steelhead trout (Gavery *et al*. 2018). The DMR were almost equally hypermethylated and hypomethylated for red blood cells, whereas the DMR were mostly hyper-methylated (66 DMR) for sperm cells. The epigenetic effects on the sperm cells could potentially be passed to future generations as a transgenerational paternal effect. There are currently few studies examining epigenetic programming differences between hatchery and wild salmon, especially across environments and generations. **Additional research is needed for identifying rearing environmental features that contribute to epigenetic differences, and whether these effects can be managed in the hatchery setting**.

3.4 > SPECIES DISTRIBUTION AND ABUNDANCE

Hatcheries may want to examine a salmon species distribution across a landscape. For example, a distribution expansion over time may indicate colonization of available habitats by threatened or endangered populations. Genomics technologies using environmental DNA (eDNA) can be used for rapid determination of a rare species distribution across large landscapes. At this time eDNA sensitivity is limited to species; however, there may be opportunities for strains if there are there are differences in key genotypes, e.g. mitochondrial DNA haplotypes (Adams *et al*. 2019). Currently, DFO researchers are in the initial stages of using eDNA to monitor farmed and wild Atlantic salmon based on differences in haplotypes (Baillie *et al*. 2019).

Water samples are collected across the landscape, filtered, the eDNA extracted from the filter, then the eDNA is quantified for the DNA using a species-specific quantitative polymerase chain reaction (qPCR) assay. The eDNA methods can be quite sensitive such that there is a need for field controls and other negative controls to test for contamination (Laramie *et al*. 2015). The eDNA methods are also evolving, e.g. glass fiber filters were better than mixed cellulose ester filters at capturing eDNA (Lacoursière-Roussel *et al*. 2016). An example of a deliverable is that Chinook salmon was detected at 93% of the known sites (out of 48 water samples) in August, as well as outside of their known range, across the sub-basins of the Upper Columbia River (Laramie *et al*. 2015). Another study used eDNA to evaluate the performance of a landscape habitat quality model for juvenile Chinook salmon. The deliverable was that Chinook salmon

was confirmed in 80% and 51% of the sites modelled as high and low habitat quality, respectively (Matter *et al*. 2018). The distribution of a salmon species across a landscape may be important for supplementation and captive breeding hatcheries. **Locations can be prioritized based on salmon presence for assigning limited hatchery resources**, e.g. release sites (Matter *et al*. 2018).

Hatcheries can also examine salmon species abundance non-invasively. Environmental DNA methods using water samples can provide rough species abundance estimates without handling individuals. That is, there are positive relationships between water eDNA content and salmon count or biomass, e.g. juvenile brook trout (*Salvelinus fontinalis*) (Lacoursière-Roussel *et al*. 2016), juvenile Chinook salmon (Shelton *et al*. 2019), and spawning sockeye salmon (Tillotson *et al*. 2018). However, the strength of the abundance relationship is dependent on the environment. For example, the abundance relationship was stronger with warmer (14°C) than cooler (7°C) water, possibly because of increased salmon eDNA release, e.g. shedding of skin, mucus, feces, and urine, in warmer water (Lacoursière-Roussel *et al*. 2016). Thus, there is a need to for research to examine the abundance relationship for a specific environmental condition. Furthermore, there may be spatial and temporal limits to salmon detectability. For instance, tens of metres from the salmon for eDNA detection in streams (Tillotson *et al*. 2018) and estuaries (Shelton *et al*. 2019). Also, eDNA may degrade within 24 hours in stream environments, such that the eDNA content may be reflective of the salmon abundance for the day of water sample collection only (Tillotson *et al*. 2018).

The eDNA examples above used qPCR assays to quantify the DNA of a targeted salmon species in a water sample. An alternative approach, i.e. eDNA metabarcoding, can identify 'all' species (dependent on the library) in a water sample (Deiner *et al*. 2017). The eDNA is amplified using PCR, sequenced on a high-throughput platform, and then DNA sequences are matched to a species library, which may be publicly available. The resulting information can then be used for species richness and abundance estimates. An example of a deliverable is an estimate of salmon abundance or biomass at a site based on the correlation with water eDNA content. Similar to distribution, a **rough estimate of population size over time in natural environments** may be important for supplementation and captive breeding hatcheries. As a cautionary note, a recent DFO review concluded that abundance quantification is more relative than exact using current eDNA methods (Baillie *et al*. 2019).

Photo by: Mitch Miller

4. TRANSCRIPTOMICS

The previous section on genomics examined the DNA molecule. This section on transcriptomics examines the RNA molecule; specifically, the messenger RNA (mRNA) molecule that is transcribed from genes or DNA, and its role in gene expression. The previous section also indirectly discussed transcription within epigenetic programming, as DNA methylation patterns can influence gene expression. Earlier transcriptomics technologies examined gene expression using microarrays — a high-throughput platform involving a glass slide spotted with thousands of target sequences (Connon *et al*. 2018). More recent transcriptomics technologies have examined gene expression using RNA-sequencing (RNA-seq) — high-throughput sequencing of RNA covering most of the transcriptome. Expression of specific genes can be examined with quantitative polymerase chain reaction (qPCR) — the continuous collection of a fluorescent signal by PCR to quantify mRNA or transcripts. A benefit of transcriptomics is that it provides information on the molecular and biochemical mechanisms for animal responses to changing environments or stressors (Connon *et al*. 2018). Although not currently in the definition of transcriptomics (in this section) environmental RNA (i.e. animal release of RNA into their environment, eRNA) is considered as a transcriptomics technology because of the examination of the RNA molecule (also see Cristescu 2019). For hatcheries, identified transcriptomics deliverables are (1) gene expression biomarkers, (2) gene expression differences, (3) living and dead components, and (4) environmental gene expression.

4.1 > GENE EXPRESSION BIOMARKERS

Gene expression may change in response to intrinsic (e.g. developmental) or extrinsic (e.g. environmental) factors. For example, specific genes can change expression in response to a certain stressor, such that they are biomarkers (Connon *et al*. 2018). Recently, there are high-throughput platforms for examining many gene expression biomarkers across many individuals. In particular, the Fluidigm BioMark™ platform can examine 96 gene expression biomarkers by 96 samples at once; the types of biomarkers and samples examined are customizable. Currently, there is a focus on examining gene expression in salmon using gill tissue because it can be collected non-lethally. The ultimate vision using this platform is the 'Salmon Fit-Chip', containing customizable gene expression biomarkers to assess the physiological health and condition of many individuals (Houde *et al*. 2019a).

Gill gene expression biomarkers for smoltification have been developed using several groups of Pacific salmon (Houde *et al*. 2019b). Specifically, gene expression panels (top 10 biomarkers) for smoltification were identified for each of coho salmon, sockeye salmon, ocean-type Chinook salmon, and stream-type Chinook salmon, as well as across these four groups, as a generalized salmon. The gene expression patterns for ocean-type Chinook salmon smoltification were compared to a companion study using seawater challenges on pre-smolt, smolt, and de-smolt groups to produce a seawater tolerance model. The deliverable was the classification of individuals as seawater intolerant (i.e. pre-smolt or de-smolt) or seawater tolerant (i.e. smolt) using gill samples collected over time. Although smoltification can be measured using changes in gill Na+/K+-ATPase activity or plasma hormones, the gill gene expression biomarkers appeared to be a more sensitive indicator of smoltification and seawater preparedness than Na+/K+-ATPase activity (Houde *et al*. 2019a). **Hatcheries may use the smoltification biomarkers for refining juvenile release strategies**, e.g. releasing the juveniles when they are majority smolt classification. Furthermore, given the general lower seawater tolerance of hatchery than wild salmon (e.g. Shrimpton *et al*. 1994, Chittenden *et al*. 2008), with additional research, the smoltification biomarkers may help identify features in the hatchery environment that alter the physiological smoltification window.

Other gill gene expression biomarkers have been developed for responses to specific environmental stressors. Specifically, a gene expression panel (eight biomarkers) for thermal stress was identified across sockeye salmon and pink salmon (Akbarzadeh *et al*. 2018). This study examined the candidate gene expression biomarkers under a single stressor condition (i.e. temperature), whereas a follow up study examined multi-stressor conditions (i.e. combinations of temperature, salinity, and dissolved oxygen). Gene expression panels were identified using Chinook salmon specific to thermal stress (10 biomarkers) and salinity stress (11 biomarkers), but not hypoxic stress (Houde *et al*. 2019a). Beyond environmental stressor specificity, the temperature and salinity biomarkers were also robust across mortality status (i.e. live vs. moribund or dead) and smolt status (pre-smolt, smolt, and de-smolt). The deliverable was high classification (> 98% correct) of individuals for responses to freshwater vs. salinity (brackish water or seawater) and three temperatures (10, 14, and 18°C). Recently, gene expression biomarkers for hypoxic stress have been developed using additional research with RNA-seq (Akbarzadeh *et al*., in prep). Although environmental monitoring may indicate the potential effects of stressors, examining the individual salmon can assess the direct physiological impact. Supplementation and captive breeding hatcheries, dealing with threatened or endangered salmon populations exposed to environmental stressors, may benefit from physiological impact evaluations at the individual or salmon level. Indeed, thermal and hypoxic stress gene expression biomarkers are being examined in the context of endangered Cultus Lake sockeye salmon, for which there is egg incubation and juvenile rearing support at Inch Creek Hatchery.

More broadly, there are gill gene expression biomarkers associated with imminent mortality or a general stress response. Ten gene expression biomarkers were linked to the imminent mortality of juvenile sockeye salmon during smolt migration (Jeffries *et al*. 2014). Although the juvenile salmon that eventually died generally also had infectious haematopoietic necrosis virus (IHNV) loads, the gene expression (10 biomarkers) pattern was a better predictor of imminent mortality than IHNV load. There were several gene expression biomarkers associated with differences between live and moribund adult sockeye salmon during the spawning run (Jeffries *et al*. 2012). Similarly, a gene expression panel (seven biomarkers) was associated with differences between live and moribund or dead juvenile Chinook salmon (Houde *et al*. 2019a). The deliverable was the classification of individuals as live or distressed (moribund or dead). However, these biomarkers had links to the secondary stress response, i.e. heat shock protein, metabolite, and immune functions, after the primary release of hormones, and there was variability among the live individuals along the primary biomarker axis. Potentially a weaker response for certain biomarkers may indicate a general stress response in live salmon, whereas a stronger response may indicate imminent mortality. **Hatcheries may use these gene expression biomarkers for optimizing practices to minimize stress for salmon**.

Gene expression biomarkers can also be examined with microbe loads simultaneously on the same Fluidigm BioMark™ platform. Load assays for 46 microbes (bacteria, viruses, and microparasites), suspected or known to cause disease in salmon, were evaluated using salmon tissue, i.e. DNA and RNA molecules and pools of several tissues (Miller *et al*. 2016). The load assay limit of detection was under 40 microbe copies, but most load assays had a limit of detection of under 10 microbe copies. The load assays also had high specificity (> 98%) to the target microbe and high repeatability (average 96% between two technicians). An example of a deliverable is that Chinook salmon smolts in Cowichan Hatchery had lower microbe prevalence, load, and diversity than wild salmon in Cowichan River (Thakur *et al*. 2018). Later in seawater, the microbe profiles were similar between hatchery and wild salmon in seawater, implying that both groups were susceptible to the microbes in a common environment. Regardless, the presence of a microbe does not necessarily mean that the salmon is in a diseased state, i.e. the salmon could be a carrier. Recently, gene expression biomarkers were developed to differentiate salmon in an active disease state from a latent (carrier) state, as well as between viral and bacterial disease states, i.e. viral disease development (VDD) biomarkers (Miller *et al*. 2017). Specifically, a VDD gene expression panel (11 biomarkers) was identified across RNA-viral species, several salmon species, and several tissues. The deliverable was the classification of individuals as an active disease state or latent state. Paired with the microbe load assays, individuals in an active disease state but with low microbe loads may warrant further investigation. **Hatcheries may use these microbe load assays and VDD biomarkers to identify and mitigate exposure to infectious agents**.

In summary, gene expression biomarkers have been developed for smoltification, thermal stress, salinity stress, hypoxic stress, general stress, imminent mortality, and viral disease development. Microbe load assays can also be examined on the same platform. There is a focus on gill tissue because it can be collected non-lethally. There are also plans for developing additional gene expression biomarkers for toxicant, starvation/fasting, and domestication selection. The ultimate objective is the production of the 'Salmon Fit-Chip', with customizable gene expression biomarkers to examine the physiological condition of hundreds of salmon at once. Hatcheries may use the gene expression biomarkers to determine environmental requirements and critical thresholds, e.g. stress responses (Connon *et al*. 2018). The biomarkers or microbe assays can also provide valuable predictive screening information, e.g. health and condition, of hatchery salmon prior to release (Connon *et al*. 2018), which may be used to explain changes in hatchery performance over time or differences among hatcheries. Furthermore, paired with biotelemetry studies (e.g. acoustic tags), hatcheries may use the biomarkers or microbe assays to provide insights on the physiological links to salmon survival, e.g. juvenile sockeye salmon during smolt migration (Jeffries *et al*. 2014, Stevenson *et al*. 2019).

4.2 > GENE EXPRESSION DIFFERENCES

Physiological differences between hatchery and wild salmon may be revealed using gene expression. Examinations of early life stages have identified gene expression differences using a transcriptome scan, with potential links to differences in the rearing environment. An example of a deliverable is that there are 723 differentially expressed genes for steelhead trout using RNA-seq on the whole swim-up fry (Christie *et al*. 2016). Key biological processes that differed were wound healing (possibly because the hatchery environment may have a higher incidence of abraded fins than the natural environment) and immunity (possibly because of high immune stimulation from crowding in the hatchery environment). Another example of a deliverable is that there are 808 differentially expressed genes between hatchery (flat) and natural (gravel) rearing substrate for the head tissue of emergent fry Atlantic salmon using a microarray containing ~44,000 probes (Evans *et al*. 2015). Biological processes that differed were broadly muscle, heart, and limb development. In particular, there was the upregulation of a growth-related gene, i.e. tissue inhibitor of matrix metalloproteinase 2, as a potential explanation for the heavier body mass on natural relative to hatchery rearing substrate.

Given that transcriptomics provides greater mechanistic information than genomics (Alfaro and Young 2018), gene expression differences between hatchery and wild salmon may help better explain the decreased hatchery salmon performance than gene or epigenetic programming differences. However, **additional research using transcriptomics technologies is needed for providing more mechanistic explanations for hatchery-wild differences and for different groups within the hatchery setting.**

4.3 > LIVING AND DEAD COMPONENTS

Hatcheries may want to quantify the living and dead components of a salmon species across a landscape. Particularly, during the spawning run where there may both live and dead individuals. This may be achieved using transcriptomics technologies for environmental RNA (eRNA). RNA represents the living component, whereas DNA represents both the living and dead components of the community (Zaiko *et al*. 2018). An example of a deliverable is a proxy of the metabolic activity (live component) of communities, i.e. RNA:DNA ratio (Xu *et al*. 2017). It was recently also highlighted that the detection of both species-specific RNA and DNA may help resolve some of the eDNA methods false positives (i.e. detection of a species not present, e.g. contamination, natural transport to site) and false negatives (i.e. species undetected when present, e.g. dilution, natural transport away from site) (Cristescu 2019). Similar to the eDNA methods, water samples are collected across the landscape, filtered, then the eRNA extracted from the filter. In contrast to the eDNA methods, as far as I am aware, salmon species-specific qPCR assays to quantify the RNA do not currently exist, such that they would need to be developed. Yet, similar to the eDNA metabarcoding technique, RNA may be identified to species and quantified by matching the sequences to a species library, which may be publicly available.

Tillotson *et al*. (2018) examined eDNA as an estimate of the abundance of live an dead sockeye salmon seperately during the spawning period. Interestingly, dead individuals appeared to release more eDNA than live individuals, such that there was a weaker relationship between water eDNA content and the abundance of live individuals (Tillotson *et al*. 2018). Conceivably, the relationship may be stronger between eRNA content and the abundance of live individuals. Albeit, similar to eDNA, there is a need for research to examine whether such an eRNA relationship exists and under different environmental conditions. Given that RNA and DNA sequences can be species-specific, supplementation and captive breeding hatcheries dealing with threatened or endangered salmon populations may use eRNA paired with eDNA methods (e.g. live and dead individuals may be counted throughout the spawning period to assess the spawning time using the live/dead ratio. eDNA and eRNA method may provide rough estimates of the abundance of dead and live individuals, as well as estimate the live/dead ratio for spawning time).

4.4 > ENVIRONMENTAL GENE EXPRESSION

Hatcheries may want to examine gene expression without handling the salmon. Given that eRNA may contain messenger RNA (mRNA), there is the prospect of examining gene expression using water samples (Cristescu 2019). Generally, eRNA methods have been less developed because of the perception that RNA is much less persistent in the environment than DNA (Cristescu 2019). However, the RNA released by organisms into their environment can remain for sufficient time in some cases, e.g. protected within extracellular vesicles (RNA carriers) that leave with bodily fluids. If tissue gene expression is proportional to the representation of the underlying mRNA in the water (i.e. biological representation), then eRNA methods would be a non-invasive approach to the examination of biomarkers for a group of salmon. As far as I am aware, there are no examples of gene expression studies using eRNA methods. Certainly, it would be advantageous to apply the existing gene expression biomarkers (above) on the eRNA content of water samples to limit the stress impacts on salmon from handling, as well as assess health and condition without capture or interruption of behaviour. A potential deliverable is estimating the seawater tolerance of juvenile salmon (as a group) within a tank using water samples collected over time and the smoltification gene expression biomarkers. If feasible, the use of eRNA methods to examine gene expression would be an entirely new research field (Cristescu 2019). Additional research is needed on technological feasibility and biological representation.

5. PROTEOMICS

The previous sections on genomics and transcriptomics examined DNA and RNA molecules, respectively. This section on proteomics examines proteins, specifically the entire protein pool of a tissue (Tomanek 2011). Proteins are generally translated from mRNA or transcripts. Briefly for proteomics, the proteins of a tissue sample are digested to peptides using proteases and then the peptides may be separated using a gel (Tomanek 2011, Zhou *et al*. 2012). Peptides are then identified and quantified using mass spectrometry and aligned to a database. Recently, mass spectrometry with ion mobility of peptides has increased the coverage of proteins by 40% relative to without ion mobility (Gombar *et al*. 2017). Similar to genomics and transcriptomics, proteomics requires reference sequence information for the database alignments, i.e. arranging the sequences of DNA, RNA, or protein to identify regions of similarity. (Rodrigues *et al*. 2012). A benefit of proteomics is that it incorporates post-translational modifications and protein degradation to provide information closer to the phenotype than transcriptomics (Rodrigues *et al*. 2012). Albeit, there is variability in the strength of the relationship between proteomics and transcriptomics. That is, the correlation can indeed be poor ($r < 0.5$) such that proteomics may provide better information than transcriptomics (Feder and Walser 2005, Maier *et al*. 2009); however, the correlation can also be moderate ($0.5 < r < 0.7$) such that transcriptomics or proteomics generally provide similar information (Schwanhäusser *et al*. 2011, Kanerva *et al*. 2014). For hatcheries, identified proteomics deliverables are (1) protein expression biomarkers and (2) protein expression differences.

5.1 > PROTEIN EXPRESSION BIOMARKERS

Proteomics research using salmon has generally focussed on improving their physiological health and condition, especially for aquaculture (Forné *et al*. 2010). For example, proteomics examinations of infected salmon have provided a better understanding of the molecular mechanisms during disease progression (Tomanek 2011, Rodrigues *et al*. 2012). Such information simultaneously can provide information on the diagnosis of a disease and the development of new vaccines or immune stimulating treatments. Similar to gene expression, specific protein expression may change in response to certain stressors, e.g. disease, such that they are biomarkers. For instance, the expression of certain proteins may increase during cellular stress: (1) chaperones that stabilize other denaturing proteins, (2) enzymes that regulate protein turnover, (3) enzymes that sense and repair DNA and RNA damage, (4) enzymes for fatty acid and energy metabolism, and (5) enzymes for redox regulation (Tomanek 2011).

Protein expression biomarkers have been identified for certain diseases, as well as microbe antigens that stimulate the immune system. Proteomics technologies are used to compare healthy and infected tissues. An example of a deliverable is at least four protein expression biomarkers were associated with infectious hematopoietic necrosis (IHN) or bacterial kidney disease (BKD) using Atlantic salmon liver and kidney tissue (Booy *et al*. 2005). Although some of the expressed proteins were expected because of their known relationship to acute infection and inflammatory responses, there were other expressed proteins associated with unexpected cellular pathways, e.g. protein synthesis related to host resistance, that were strong biomarkers for diagnostics and vaccine development. Antigen membrane proteins of the bacterium associated with cold-water disease (CWD) were evaluated for an immune system response using rainbow trout and proteomics technology (Dumetz *et al*. 2008). Another example of a deliverable is out of the 36 antigen proteins, 25 proteins elicited an immune response, which may serve as targets for salmon vaccines. **Hatcheries may use existing candidate protein expression biomarkers for salmon disease diagnostics and vaccine development. Diagnostic biomarkers and vaccine targets can also be developed for other hatchery disease concerns.**

Protein expression biomarkers have also been identified for reproductive quality. Proteomics technologies were applied to ovarian or seminal fluid in association with reproductive traits or variation among individuals. An example of a deliverable is at least eight protein expression biomarkers were associated with egg quality (i.e. post-ovulatory ageing) using rainbow trout ovarian fluid (Rime *et al*. 2004). Interestingly, it was suggested that the leakage of egg proteins into the ovarian fluid decreased egg fertility and developmental competence. Similarly, there is 174 proteins that differed in expression among 25 female Chinook salmon using ovarian fluid (Johnson *et al*. 2014). Given that there is a strong female by male interaction for sperm performance (i.e. cryptic female choice), it is possible that the protein expression variation among females mediates such an interaction. Another example of a deliverable is that there are several protein expression biomarkers suspected for sperm performance (e.g. spermatogenesis and motility) using Chinook salmon seminal fluid (Gombar *et al*. 2017). For the last two studies, there was no direct association between protein expression and reproductive traits. If there are clear associations between protein expression and a reproduction trait, then a deliverable is protein expression biomarkers.

Overall, in contrast to the gene expression biomarkers, protein expression biomarkers have not yet been studied across a range of conditions to determine specificity to a certain stressor (e.g. single study vs. meta-analysis) or validated using new samples, e.g. a different species. These discovery and validation approaches are important for robust biomarker development (e.g. Miller *et al*. 2017). At this time, there are candidate protein expression biomarkers for (1) IHN and BKD using liver and kidney tissue and (2) egg quality using ovarian fluid that may be validated and used in hatcheries. **Additional research is needed to develop robust protein expression biomarkers for salmon traits important to hatcheries**.

5.2 > PROTEIN EXPRESSION DIFFERENCES

Better mechanistic explanations for the lower hatchery salmon performance than wild salmon may be explored using proteomics over genomics or transcriptomics. Albeit, given the novelty of proteomics technologies, there are currently no examples of protein expression differences between hatchery and wild salmon. However, there is one example of a comparison between groups displaying a difference in reproductive performance. A proteome scan of the seminal fluid of Chinook salmon identified 30 proteins that differed in expression between jack and adult males (Gombar *et al*. 2017). Jacks generally have higher sperm quality than adult males, i.e. relatively larger testes and faster swimming sperm in water. Key biological processes that differed were hormone transport, energy metabolism, and proteolysis, as potential mechanistic explanations for the sperm quality differences. With further studies comparing groups using proteomics, a deliverable is the number of differentially expressed proteins and the key biological processes involved. As discussed in previous sections, **additional research using proteomics technologies is needed for providing more mechanistic explanations for hatchery-wild salmon differences and for different groups within the hatchery setting.**

Photo by: Sam James

6. METABOLOMICS

The previous sections on genomics, transcriptomics, and proteomics examined DNA, RNA, and proteins, respectively. This section on metabolomics examines metabolites (e.g. sugars, organic acids, amino acids, vitamins, lipids, and nucleotides), specifically the entire metabolite pool of a tissue (Jones *et al*. 2013, Alfaro and Young 2018). Certain metabolic pathways are regulated and operated by enzymes (proteins). Metabolites are the small molecule (< 1 kDa) that are the products of metabolism, which may be strongly amplified in response to stressors (Lankadurai *et al*. 2013, Alfaro and Young 2018). Because metabolic processes continue in tissues even after dissection, tissues should be preserved quickly (e.g. liquid nitrogen) and stored at or below -80°C to inactivate these processes (Young and Alfaro 2018). Non-targeted (all) metabolites or targeted metabolites can be examined (Lankadurai *et al*. 2013). Briefly, nuclear magnetic resonance (NMR) spectroscopy is typically used for non-targeted metabolomics because of its low selectivity, but only metabolites in significant quantities are detected (Lankadurai *et al*. 2013, Young and Alfaro 2018). In contrast, mass spectrometry (MS) with separation by chromatography or electrophoresis is typically used for targeted metabolomics because of its high selectivity and ability to detect metabolites in low quantities. However, both metabolomics technologies can be used for biomarker development, i.e. NMR to screen all metabolites and discover candidate biomarkers, followed by validation or monitoring using MS. A benefit of metabolomics is that it may provide the most mechanistic information relative to genomics, transcriptomics, and proteomics (Alfaro and Young 2018). For hatcheries, similar to proteomics, identified deliverables are (1) metabolite expression biomarkers and (2) metabolite expression differences.

6.1 > METABOLITE EXPRESSION BIOMARKERS

Similar to proteomics, metabolomics research using salmon has generally focussed on improving their physiological health and condition, especially for aquaculture (Alfaro and Young 2018). In particular, metabolomics examinations of infected salmon have provided a better understanding of the molecular mechanisms during disease progression, including early diagnosis (Jones *et al*. 2013). Metabolomics research may provide mechanistic and functional biochemical information that may help solve issues with hatchery performance (Young and Alfaro 2018). Similar to gene expression and protein expression, metabolite expression may change in response to certain environmental stressors or culture conditions (Alfaro and Young 2018). For example, toxin exposure increases the metabolites of redox homeostasis, e.g. glutathione and NADPH (Young and Alfaro 2018).

Metabolite expression biomarkers have been identified for certain salmon diseases. Metabolomics technologies are used to compare healthy and infected tissues using NMR. An example of a deliverable is that there are 18 metabolite biomarkers for bacteria *Aeromonas salmonicida* infection using Atlantic salmon blood plasma (Solanky *et al*. 2005). The bacterial infection caused major modifications in lipid metabolism, with minor modification in carbohydrate and amino acid metabolism. Another example of a deliverable is 18 metabolite expression biomarkers were associated with A. *salmonicida* infection using Atlantic salmon kidney tissue (Liu *et al*. 2016). The infection caused absorption inhibition of amino acids and disturbed protein and cell metabolism. **Beyond existing diagnosis biomarkers for *A. salmonicida*, metabolite expression biomarkers may be developed for other hatchery disease concerns.**

Metabolite expression biomarkers have also been identified for salmon health or condition. An example of a deliverable is at least nine metabolite expression biomarkers associated with long-term stress were identified using NMR on juvenile Atlantic salmon blood plasma (Karakach *et al*. 2009). Generally, carbohydrates increased with stress, whereas lipids decreased. Another example of a deliverable is out of the 137 metabolites monitored using MS on blood plasma, 27 metabolite expression biomarkers were associated with starvation or fasting of female Atlantic salmon (Cipriano *et al*. 2015). Generally, fatty acid concentrations decreased with fasting, whereas sugar alcohols concentrations increased. Albeit, the study used inappetent (lacking appetite) individuals during the spawning run, when the salmon may be naturally tolerant of prolonged fasting or starvation than earlier life stages. Another example of a deliverable is out of the 95 metabolites monitored using MS on liver tissue, 23 and 9 metabolite expression biomarkers for were associated with possible toxin (e.g. copper) exposure for Skeena River female and male sockeye salmon, respectively (Benskin *et al*. 2014). It was suspected that toxin exposure caused endocrine disruption during spawning migration into Babine Lake, such that certain females displayed more masculine profiles and certain males displayed more feminine profiles. However, the results could also be explained by natural variability of individuals rather than toxin exposure *per se*.

Overall, similar to the discussion on protein expression biomarkers, metabolite expression biomarkers have not yet been studied across conditions for specificity to a certain stressor or validated using new samples. A possible explanation is the novelty of both proteomics and metabolomics technologies such that there are a small number of studies for discovery and validation approaches. However, these approaches are important for robust biomarker development, as demonstrated for gene expression biomarkers (e.g. Miller *et al*. 2017), as well as highlighted in the metabolomics strategies for aquaculture (i.e. Young and Alfaro 2018). At this time, there are candidate metabolite expression biomarkers for (1) A. *salmonicida* infection using blood plasma or kidney tissue and (2) long-term stress using blood plasma that may be validated and used in hatcheries. **Additional research is warranted to develop robust metabolite expression biomarkers for salmon traits important to hatcheries.**

6.2 > METABOLITE EXPRESSION DIFFERENCES

Metabolomics, relative to genomics, transcriptomics, and proteomics, may provide the best mechanistic explanations for the lower hatchery salmon performance relative to wild salmon. However, at this time, there are no examples of metabolite expression differences between hatchery and wild salmon, likely because of the novelty of metabolomics. However, there is one example comparing farmed and wild Atlantic salmon. The deliverable was that there is 187 differentially expressed metabolites using NMR on lipid extracts of muscle tissue (Aursand *et al*. 2009). Although there was no specific analysis to determine the differences in key biological processes to reveal mechanistic explanations, farmed salmon displayed higher concentrations of n-6 fatty acids than wild salmon. This is not surprising given that n-6 fatty acids are abundant in the vegetable oils used for farmed salmon diets (i.e. commercial feeds). That is, the different diets between farmed and wild salmon was a major confounding factor because of their strong influence on the metabolome. Nonetheless, using the top 12 metabolites, another deliverable was high classification (96% correct) of individuals as farmed or wild salmon. Similar to previous sections, with further studies comparing hatchery and wild salmon, a deliverable is a number of differentially expressed metabolites and the key biological processes involved. **Additional research using metabolomics technologies, specifically NMR, is needed for providing more mechanistic explanations for hatchery-wild differences and for different groups within the hatchery setting.**

Photo by: Sam James

7. KNOWLEDGE GAPS

Beyond describing deliverables using 'omics' technologies, the present review also examined the scientific literature on hatchery issues to identify knowledge gaps. Filling in the knowledge gaps with studies using 'omics' technologies can increase the understanding of how to best improve hatchery performance. Here, I highlight key experimental considerations and supporting background information for filling in four knowledge gaps: (1) group performance, (2) offspring quality, (3) health and condition biomarkers, and (4) hatchery-wild differences. Key 'omics' technologies or main text sections are in bold.

7.1 > GROUP PERFORMANCE

There are few examples of relative group performance evaluations using 'omics' technologies. Importantly, a clear strength of **parentage-based tagging (PBT)** is the ability to genetically tag all hatchery salmon to the family level with the families assigned to specific groups (see Figure 3). Based on increases in hatchery salmon survival, additional insight may be achieved examining by group attributes such as (1) ecological smolt window, (2) physiological smolt window, and (3) semi-natural rearing.

7.1.1 ECOLOGICAL SMOLT WINDOW

Juvenile release group timing can be examined with zooplankton bloom timing during the 'ecological smolt window' (McCormick *et al*. 1998, Björnsson *et al*. 2011). Certain salmon populations have poor survival during the early marine period, such that an increase in juvenile survival during this period may translate to increased salmon production (Beamish *et al*. 2009, 2010). For example, there were lower juvenile captures in the estuary near Discovery Passage (juveniles moved farther away from shore) and higher adult male returns (1.5 to 3 times) for hatchery coho salmon that were released within a week of the zooplankton bloom (> 1,000 zooplankton/m3) than salmon released earlier or later (Chittenden *et al*. 2010). Hatcheries may release multiple groups of juvenile salmon progressively during the spring, with each group tagged for monitoring purposes according to their **PBT** family. The movement and performance of the release groups can be examined for associations with features of the local zooplankton bloom, e.g. timing.

7.1.2 PHYSIOLOGICAL SMOLT WINDOW

Another important consideration for juvenile release timing strategies is smolt status during the 'physiological smolt window' (McCormick *et al*. 1998, Björnsson *et al*. 2011). That is, juvenile survival in seawater is dependent on smolt status. For instance, juvenile mortality is higher for pre-smolt (16%, 21 of 128 fish) or de-smolt (13%, 25 of 192 fish) than smolt (1%, 1 of 96 fish) Chinook salmon after six days of 10°C or 14°C seawater (28 or 29 PSU) exposure using a controlled set-up (Houde *et al*. 2019a). Interestingly, juvenile mortality was low (< 1%, 4 of 416 fish) for the same duration and temperature in brackish water (20 PSU) across smolt status. This salinity was likely similar to that internally of the fish, such that the energy expenditure in brackish water may be less than seawater for maintaining homeostasis. Similar to low food periods, pre-smolt and de-smolt salmon may be exposed to increased predation, as they remain in freshwater or brackish water estuary habitats longer. In other words, there is a positive relationship between the time to seawater entry (or estuary exit) and smolt status (Stich *et al*. 2015, 2016). Furthermore, hatchery salmon also generally have lower seawater tolerance than wild salmon (e.g. Shrimpton *et al*. 1994, Chittenden *et al*. 2008). Thus, to limit juvenile mortality after release, hatcheries can classify the smolt status using weekly gill samples and the **smoltification gene expression biomarkers**, such that juveniles are released as majority smolt. However, the optimum for juvenile survival consists of overlapping ecological and physiological smolt windows (McCormick *et al*. 1998, Björnsson *et al*. 2011). Using the multiple groups tagged by separate **PBT** families (above) for the zooplankton bloom, hatcheries may align juvenile smoltification with the targeted release date. For example, smoltification can be advanced by increasing water temperature in the hatchery environment (McCormick *et al*. 1998, Björnsson *et al*. 2011).

7.1.3 SEMI-NATURAL REARING

There has been a general promotion of the idea that the more the hatchery environment resembles the natural environment, the better for juvenile salmon survival (Lynch and O'Hely 2001, Tave and Hutson 2019). For example, semi-natural rearing techniques associated with increased post-release juvenile survival are: (1) cobble substrates (e.g. embedded in concrete pavers), (2) in-water structure (e.g. plastic crates), (3) overhead cover (e.g. camouflage net), and (4) anti-predator training (e.g. visual, acoustic, or chemical cues) (Maynard *et al*. 2004, Näslund and Johnsson 2016). Other juvenile performance improvements are associated with additional techniques: (5) supplementing diets with live food (i.e. increasing in prey hunting ability) and (6) exercise current velocities (i.e. increased swimming performance). Although the semi-natural rearing techniques are associated with increased performance at the juvenile level, they have not been evaluated at the adult level (Maynard *et al*. 2004), possibly because of logistical constraints that may now be alleviated with use of PBT for tracking groups. Hatcheries may release different groups of juvenile salmon reared using traditional hatchery and combinations of semi-natural rearing techniques, with each group tagged by separate **PBT** families. Another group that can be compared is natural or wild juvenile salmon as a benchmark for performance or as another control for year effects. The parents of the natural group may not have been examined with PBT, so another bulk tagging technique for these salmon would be necessary.

Overall, using groups tagged by separate **PBT** families as well as considering smolt status, the ultimate post-release performance measures for hatchery salmon are the contribution to the fishery for augmentation hatcheries, and the contribution to escapement or reproductive success for supplementation and captive breeding hatcheries. Groups that demonstrate higher relative performance may be pursued further by the hatcheries to achieve their respective goals.

7.2 > OFFSPRING QUALITY

The hatchery offspring quality may increase by allowing mate choice or the by targeting the specific genes underlying important salmon traits. In particular, across a range of taxa, allowing mate choice may increase offspring fitness by a median of 36% (Neff *et al*. 2011). However, there are few examples of the quantitative benefits of mate choice in salmon, as well as the trait architecture beyond early life stages, i.e. egg to fry. Examinations of the mate choice benefits and trait architecture at later life stages require the breeding of individuals with more than one partner using a full factorial breeding design, such that **PBT** is a useful tool for tagging all offspring and providing assignments back to specific families. The salmon **trait architecture** may also provide the guidance for follow-up studies at the molecular level using 'omics' technologies to identify specific **biomarkers** linked to improvements in offspring quality.

7.2.1 MATE CHOICE BENEFITS

Mate choice benefits can be measured using spawning channels and a full factorial breeding design. Spawning channels containing female and male salmon can facilitate sexual selection via male-male competition and female mate choice (Neff *et al*. 2011). The female's chosen male can be recorded using behavioural observation criteria (Puurtinen *et al*. 2009). Then, using a full factorial breeding design, the female can be mated with the recorded male along with at least one other male. In other words, the concept is to allow mate choice in spawning channels, then remove the female and chosen male to artificially produce offspring with the mate choice pair and a control pair. For example, using two females by two males (2 X 2), the egg batch of each female is divided in half, with the first half fertilized by the chosen male and the remaining half fertilized by the other male. Thus, the total of four families is represented by two mate choice families and two non-mate choice families, with each family tagged separately by **PBT**. The resulting offspring survival and fitness-related traits (e.g. body size) can be measured for the two mate choice groups using the family assignments. Furthermore, with the full factorial breeding design, the amount of additive genetic and non-additive genetic effects explaining the offspring traits can be estimated, which can speak to the importance of 'good genes' and 'compatible genes' for salmon mate choice systems (Neff and Pitcher 2005). Several parents (e.g. 5 X 5) and/or blocks (e.g. several 2 X 2 blocks) may be needed for sufficient statistical power, which can be determined using a power analysis (Houde and Pitcher 2016). If mate choice benefits are sufficiently large (e.g. 33% higher offspring survival for mate choice than non-mate choice), then hatcheries may want to consider allowing mate choice to increase offspring quality.

7.2.2 TRAIT ARCHITECTURE

The full factorial breeding design can also be used for examining **trait architecture** for more than early life stages. In particular, maternal and non-additive genetic effects explain a large amount of variation in salmon traits during early life stages, i.e. egg to fry (Houde *et al*. 2013, 2015, 2016), but it is not clear if the importance of these effects continue into later life stages, i.e. smolt to adult. Maternal effects should decrease as maternal investments (e.g. egg resources) are utilized, while genetic effects should increase in importance for later life stages traits (Wilson and Réale 2006). Conceivably, non-additive genetic effects may remain important for later life stages traits with strong ties to fitness, e.g. survival, whereas for traits with weaker ties to fitness, e.g. body size, additive genetic effects may be more important (Crnokrak and Roff 1995, Roff and Emerson 2006). If non-additive genetic effects do remain important at later life stages, the benefits of mate choice for 'compatible genes' may be key across the entire life cycle. Given that hatchery performance goals are generally focussed on the adult life stage, e.g. contribution to the fishery or escapement and reproductive success, the full factorial breeding design with family tagged separately by **PBT** may provide information on the underlying architecture and mate choice system explaining such important traits.

7.2.3 BROOD STOCK BIOMARKERS

Genetic or maternal effects explaining a large amount of a given salmon trait can direct complementary brood stock biomarker approaches. Traits explained by large genetic (additive and/or non-additive) effects may warrant further investigation using **gene biomarkers** to identify the specific genes contributing to higher offspring quality. Traits explained by large maternal variance may warrant further investigation using **protein expression biomarkers**, **metabolite expression biomarkers**, and/or chemical associations to identify the features, i.e. specific proteins, small molecules, and/or chemicals, in the egg or ovarian fluid that are contributing to higher offspring quality. With information on the specific genes or maternal feature explaining salmon traits, hatcheries may screen brood stock with the biomarkers to assess the value of a female ('beneficial genes' and maternal), a male ('beneficial genes'), and/or a female-male pair ('compatible genes').

Overall, a full factorial breeding design with families tagged by **PBT** can help fill the knowledge gaps on the quantitative benefits of mate choice and trait architecture of later life stages. Using **trait architecture** as a guide, the corresponding **biomarker** approach may identify specific genes or maternal features that may be beneficial to keep or promote in the brood stock. Hatchery offspring quality may be increased by allowing mate choice (if the quantitative benefits are large) or by screening potential brood stock with the resulting biomarkers to determine specific matings.

7.3 > HEALTH AND CONDITION BIOMARKERS

Beyond brood stock biomarkers for increasing offspring quality, **expression biomarkers** can also be produced for other aspects of the health and condition of hatchery salmon. Although there are existing gene expression biomarkers (i.e. 'Salmon Fit-Chip') for smoltification, salinity stress, thermal stress, hypoxic stress, general stress, imminent mortality, and viral disease development using primarily gill tissue, there are knowledge gaps for biomarkers of other aspects of hatchery salmon health and condition. For instance, reproductive development (egg and sperm) and fungal disease development have not been studied for suitable biomarkers. Current expression biomarkers use tissue samples, which may be lethally (e.g. liver) or non-lethally (e.g. gill) collected, but both require the handling of salmon. There are also knowledge gaps about what tissues and 'omics' technologies may be the most appropriate for biomarker development. To develop robust biomarkers, discovery and validation are needed (Young and Alfaro 2018). Biomarkers that would be applicable across various species and environmental conditions could then be developed with information from several studies and considering a range of samples to create even more robust and useful biomarkers (Connon *et al.* 2018). Furthermore, the handling of salmon may not be required if relevant tissue molecules (i.e. mRNAs, proteins, and metabolites) are excreted into the environment, such that water samples can be analyzed for expression biomarkers. However, additional knowledge gaps remain to be overcome, such as how long these molecules remain in the water before degrading (i.e. quality) and whether current 'omics' technologies have adequate sensitivity to detect such molecules in low quantity.

7.3.1 CHOICE OF TISSUE AND 'OMICS' TECHNOLOGY

The identification of expression biomarkers may use different tissues and 'omics' technologies, depending on the specific salmon health and condition attribute of interest. Ideally, target tissues would be non-lethally collected, e.g. blood, gill, ovarian fluid, and seminal fluid. The choice of tissue and corresponding 'omics' technology may be determined using background information. For example, the blood circulates leukocytes (i.e. white blood cells) involved in immune responses, with gills important to pathogen resistance and blood circulation (Bond 1996), such that either tissue could be examined for aspects of disease progression or resistance. The initial response to external (e.g. pathogen) and intrinsic (e.g. hormone) factors is a change in transcripts or mRNA (Young and Alfaro 2018), such that **gene expression biomarkers** may be appropriate for disease response and salmon development. Furthermore, ovarian fluid and seminal fluid can have tight links to reproductive success (e.g. egg and fertilization quality), with fluid protein and metabolite composition explaining a portion of the success (e.g. Lahnsteiner *et al*. 1995, 1998), such that **protein expression biomarkers** and/or **metabolite expression biomarkers** may be appropriate. Otherwise, more than one tissue or 'omics' technology may be examined using an exploratory approach. For instance, male reproductive development is associated with testes gene expression (Middleton *et al*. 2019); however, it is unknown if the relevant mRNA is also detectable in seminal fluid, or possibly ovarian fluid for female reproductive development. Using more than one 'omics' technology, can also provide a higher coverage of mechanistic information for salmon physiology (Forné *et al*. 2010), as well as provide higher coverage for identifying the biomarkers with the tightest links to a given aspect of salmon health and condition.

7.3.2 ROBUST BIOMARKERS

Discovery and validation approaches are needed for the development of robust biomarkers. Comparison groups may be defined by quantified differences in salmon health and condition traits. If these group details are not known at the time, tissue samples can be collected and stored (e.g. -80°C freezer) until such details become available. Alternatively, salmon health and condition traits that are influenced by time (e.g. reproductive quality) may be represented by the collection of tissue samples at regular time intervals. For comparison groups, a base study can discover candidate biomarkers using an 'omics' screen with a statistical model to identify differences in molecule expression, e.g. identify the genes differing in expression across the transcriptome. For time intervals, a base study can discover candidate biomarkers using an 'omics' screen to correlate or otherwise associate molecules to traits; a multivariate simplification technique (e.g. principal component analysis) may be useful in this case. Previous research on certain candidate biomarkers were not subsequently validated using new samples, e.g. **protein expression biomarkers** and **metabolite expression biomarkers**. This biomarker validation process is important for decreasing type II errors (i.e. false negative results) and increasing reproducibility, ultimately translating to higher sensitivity and specificity (Young and Alfaro 2018).

Extensions of the discovery and validation approaches can result in the development of even more robust expression biomarkers. However, the data requirements are higher because a larger number of studies and samples are needed. Existing studies are mined or are subjected to a meta-analysis to identify the candidate expression biomarkers, i.e. transcripts, proteins, or metabolites, that are shared in response to a given condition (Connon *et al*. 2018). The resulting candidate biomarkers are then validated using new samples covering a range of intrinsic (e.g. species, age, sexes, or life stages) and extrinsic (e.g. environmental conditions or multiple stressors) factors (Young and Alfaro 2018). Such an extended approach was used for developing the robust **gene expression biomarkers**, e.g. viral disease development (Miller *et al*. 2017) and smoltification (Houde *et al*. 2019b).

7.3.3 NON-INVASIVE BIOMARKERS

Considering that hatchery salmon are typically reared in closed or semi-closed tanks, as well as at higher densities than natural conditions, tissue molecules should be present in sufficient quality and quantity in tank water for environmental analyses. It is known that environmental DNA (eDNA) can remain in water for a sufficient amount of time (i.e. adequate quality, slow degradation) to allow measurements of species distribution and abundance. Environmental RNA (eRNA), including mRNA, conceivably may also remain in water for a sufficient amount of time to allow measurements of environmental gene expression (Cristescu 2019). Similar cases can be made for environmental proteins and metabolites considering future research beyond eDNA (Seymour 2019). There are also examples of 'omics' technologies using environmental molecules (i.e. RNA, proteins, and metabolites) for examining microbial communities of surface water biofilms or soil samples (e.g. Gotelli *et al*. 2012, Kido Soule *et al*. 2015, Zaiko *et al*. 2018). However, using the current 'omics' technologies, environmental gene expression, protein expression, and metabolite expression *per se* have not been demonstrated. Regardless of environmental molecule quality

or current technological feasibility, environmental molecules are generally in low quantities because of low excretion rate or water dilution. Although molecule amplification technology exists for DNA and RNA, it currently does not for proteins and metabolites (Zhou *et al*. 2012). Given that water samples are passed through filters to collect the eDNA, an option for other environmental molecules is to increase the volume of water that is sampled. Future technological advancements for proteomics and metabolomics are also expected to have increased sensitivity (Zhou *et al*. 2012, Alfaro and Young 2018), enabling measurements of proteins and metabolites in low quantity. Certainly, an initial area of investigation for environmental molecules is determining the volume of the water sample to be filtered, resulting in the necessary molecule quantity for current 'omics' technologies. Given the concerns of RNA, protein, and metabolite degradation, another research need is to evaluate molecule quality limits for current 'omics' technologies.

Assuming there are few quantity or quality limitations to using current 'omics' technologies on environmental molecules, expression biomarkers may be examined for hatchery salmon at the tank level using water samples. For example, tank water samples are used for non-invasive stress evaluations of fish using cortisol measurements (Ellis *et al*. 2004) and most recently using ß-D-glucose measurements (Makaras *et al*. 2020). Yet, biological representation of the environmental molecules has not been tested, i.e. does the environmental molecule (i.e. mRNA, proteins, and metabolites) representation in a water sample reflect the tissue expression? For eDNA research, relationships were detected between water eDNA content and salmon abundance or biomass using correlation tests. A similar approach may be used for testing possible relationships between environmental molecules and the relevant tissue expression used in the development of the biomarkers. For example, is there a significant positive relationship between tank water mRNA content and the average gill gene expression using the smoltification biomarkers? If biological representation is supported for environmental molecules, **environmental gene expression**, **environmental protein expression**, and **environmental metabolite expression** can decrease the handling of salmon for tissue collection used for expression biomarkers; thus, limiting stress responses (Wagner *et al*. 2002). This water sampling approach could be used for assessing reproductive development by decreasing the regular handling of salmon (i.e. testing for ripeness) during the spawning period. Relevant expression biomarkers for reproductive development would first need to be developed for tissues using the methods described in the two earlier sections (7.3.1 and 7.3.2), followed by testing water sample suitability using the methods described in this section.

In summary, there are existing **gene expression biomarkers** that have hatchery deliverables, e.g. smoltification biomarkers for smolt status. Other relevant tissue expression biomarkers can be produced for aspects of salmon health and condition that are important to hatcheries, e.g. reproductive development. To produce robust **expression biomarkers**, the choice of tissue and 'omics' technology, as well as discovery and validation approaches are described. Although environmental DNA methods have been demonstrated for technological feasibility and biological representation using water samples, environmental molecule (i.e. mRNA, protein, and metabolite) expression using biomarkers have not to date. If plausible, environmental molecules in water samples would be a non-invasive approach to simultaneously decrease the handling of salmon and obtain measures of health and condition using relevant expression biomarkers.


Photo by: Mitch Miller

7.4 > HATCHERY-WILD DIFFERENCES

The performance of hatchery salmon is generally lower than wild salmon in the natural environment (Fraser 2008, Christie *et al*. 2014). Mechanistic explanations have traditionally been explored using broad-sense genomics, i.e. gene differences, which generally revealed that the underlying DNA background between hatchery and wild salmon is similar (e.g. Mäkinen *et al*. 2015, Christie *et al*. 2016, Le Luyer *et al*. 2017, Gavery *et al*. 2018, 2019). Recently, mechanistic explanations have been explored using narrow-sense genomics, i.e. **epigenetic programming differences**, which highlighted that DNA methylation patterns differences may produce gene expression differences (e.g. Le Luyer *et al*. 2017, Gavery *et al*. 2018). Transcriptomics has supported such gene expression differences between hatchery and wild salmon (Christie *et al*. 2016) or between rearing environments, i.e. hatchery vs. semi-natural rearing (Evans *et al*. 2015). Nonetheless, there are knowledge gaps for mechanistic explanations using proteomics and metabolomics, which have closer links to salmon traits or phenotypes than genomics and transcriptomics (Rodrigues *et al*. 2012, Alfaro and Young 2018). Furthermore, the evidence for epigenetic programming differences is currently limited to early life stages prior to hatchery release. There are knowledge gaps on the amount of epigenetic effects explaining traits, as well as the epigenetic timeline after hatchery release (Gavery *et al*. 2018, 2019). Specifically, are juvenile hatchery salmon epigenetic effects persistent (i.e. influence adult traits), reversible (i.e. same as wild salmon if within a common environment), or transgenerational? Indeed, epigenetic effects studies for early environmental effects on later performance in fishes is currently a poorly developed field (Jonsson and Jonsson 2014). There are also knowledge gaps on the epigenetic effects of semi-natural rearing techniques. Given that semi-natural rearing techniques may produce more wild-type salmon traits (Maynard *et al*. 2004, Näslund and Johnsson 2016), conceivably semi-natural rearing techniques may decrease the epigenetic differences between hatchery and wild salmon (Tave and Hutson 2019).

7.4.1 MECHANISTIC EXPLANATIONS

Hatchery and wild salmon differences may be examined using proteomics and metabolomics to explore mechanistic explanations for lower hatchery salmon performance. Across the whole proteome and metabolome, **protein expression differences** and **metabolite expression differences** can be identified. These differences can then be subjected to a functional annotation analysis to reveal the key biological processes involved. In particular, the key biological processes using metabolomics may provide the best mechanistic explanations for the hatchery and wild salmon differences, followed by the results using proteomics (see Young and Alfaro 2018). As far as I am aware, there is no study that has examined brain tissue directly for an analysis of hatchery and wild salmon differences. Domestication selection typically influences behaviour first (Mayr 1963), there are differences between hatchery and wild salmon in behaviour, e.g. aggression, feeding, and anti-predator (Weber and Fausch 2003, Huntingford 2004), and some of these behavioural differences are linked to post-release juvenile survival (Maynard *et al*. 2004, Näslund and Johnsson 2016). Examining metabolite expression differences and protein expression differences for brain tissue may be an important first step. Additional tissues associated with observed differences, e.g. liver tissue for metabolic differences, could also be examined to reveal mechanistic explanations for hatchery and wild salmon differences.

7.4.2 TRAIT EPIGENETIC PROGRAMMING

Trait differences between hatchery and wild salmon may have an underlying epigenetic programming basis stimulated by rearing environment differences. Threshold trait (e.g. run timing) or meristic trait (i.e. age-at-maturity, e.g. jimmy, jack, and adult ages) differences can be examined at the molecular level using **trait epigenetic programming**, highlighting any trait associations to DNA methylation patterns. A functional annotation analysis of the associations can reveal the key biological processes involved to provide more information on the mechanistic underpinnings for the trait differences. Of note, trait epigenetic effects cannot be quantified using an extension of trait architecture methods, i.e. splitting the families across the rearing environments and incorporating environmental variance and genotype X environment variance (Lynch and Walsh 1998). That is, epigenetic effects can impact all components explaining the trait, i.e. genetic variance, environmental variance, genotype X environment variance, genotype-environment covariance, and residual variance (Banta and Richards 2018), such that examinations at the molecular level are necessary to explain traits. However, it is possible to estimate the amount of trait transgenerational or epigenetic variance using a complicated pedigree and analytical model incorporating the covariances between (1) parents and offspring, (2) siblings, and (3) uncles and nephews (Tal *et al*. 2010). This type of quantitative genetic analysis may be pursued if it is revealed that transgenerational epigenetic effects are problematic for hatchery salmon performance in the natural environment (see below).

7.4.3 EPIGENETIC TIMELINE

Differences in hatchery and wild salmon epigenetic effects may be examined just before release and post-release, as well as across generations. Just before release, juvenile **epigenetic programming differences** can serve as a benchmark for rearing environment influences on early life.

Persistence of epigenetic effects into the adult life stage may be determined using adults compared to the juvenile benchmark. Given the similarities between juvenile and adult Atlantic salmon competitive behaviour (Fleming *et al*. 1997), certain behavioural traits may have underlying persistent epigenetic effects for the brain. Ideally, the comparison between life stages would be at the individual level, but the lethal collection of certain tissues (e.g. brain) may necessitate comparisons at the life stage group level.

Reversibility of epigenetic effects may be determined using post-release hatchery juveniles compared to wild juveniles after a few months in a common environment. Given the similarities between hatchery and wild post-smolt gene expression in the early marine environment using brain, gill, white muscle, and liver tissues (Houde *et al*. 2019c), certain traits may have reversible epigenetic effects. Albeit, the similarities between hatchery and wild salmon may be explained by higher hatchery than wild salmon mortality, such that the surviving hatchery salmon more closely resemble wild salmon (see Waples 1999). This caveat may be addressed by minimizing mortality, e.g. by using a semi-natural common environment, and by a comparison to the benchmark.

Transgenerational potential of epigenetic effects can be determined using a comparison of hatchery to wild reproductive cells (i.e. egg and sperm), as well as the offspring. Given the epigenetic programming differences between hatchery and wild male steelhead trout sperm cells (Gavery *et al*. 2018), certain traits may have trans-generational potential. However, the transgenerational potential should be confirmed early in the next generation, i.e. fertilized eggs, before the environment can further influence epigenetic programming. Similarly, offspring produced using reciprocal hatchery X wild hybrid crosses (i.e. female X male and male X female) may provide insights on the epigenetic effects that may integrate into populations. A comparison between hatchery genera-tional benchmarks at the same juvenile life stage may also highlight if there has been an increase in the number of epigenetic programming differences. In particular, transgenerational epigenetic effects have been suggested as a mechanism for the continued decrease of hatchery salmon performance in the natural environment with every generation (Christie *et al*. 2014, 2016), such that an increase in the number of epigenetic programming differences for hatchery salmon from one generation to the next may support this hypothesis.

7.4.4 SEMI-NATURAL REARING

Semi-natural rearing techniques may decrease the epigenetic programming differences between hatchery and wild salmon, but this has not been tested empirically. The benchmark for the amount of epigenetic effects can be represented by the number of **epigenetic programming differences** for hatchery juveniles just before release. A similar analysis can be done using another rearing environment, i.e. semi-natural vs. wild salmon, at the same life stage. If semi-natural rearing techniques decrease the amount of epigenetic effects, then the number of differences semi-natural vs. natural should be less than the hatchery vs. wild benchmark. However, early life stage epigenetic effects may be reversible or persistent considering later life stages (Jonsson and Jonsson 2014). Reversible epigenetic effects are the ideal for hatcheries, i.e. there are few epigenetic programming differences after hatchery and wild salmon spend a period of time in the early marine environment. Yet, persistent epigenetic effects that influence adult traits in the natural environment, e.g. poor competitive behaviour (Fleming *et al*. 1997), may be undesirable. Furthermore, given that hatchery salmon may breed in the natural environment and/or interbreed with wild salmon, transgenerational or epigenetic effects for traits maladaptive in the natural environ-ment may decrease population fitness (Fraser 2008, Tave and Hutson 2019). Albeit, traits with epigenetic effects may not be transgenerational, i.e. not represented in reproductive cells or transferred to offspring (Jonsson and Jonsson 2014). Persistence, reversibility, and transgenerational potential of a given trait at later life stages should be examined (described above). Semi-natural rearing techniques producing epigenetic programming closer to wild salmon may be targeted as an objective by hatcheries to decrease impacts at the adult life stage (i.e. persistent epigenetic effects) and on populations (i.e. transgenerational epigenetic effects).

Overall, mechanistic explanations for differences in performance of hatchery and wild salmon should be explored using metabolomics and proteomics because of their closer ties to traits, especially for brain tissue given the behavioural observations. Differences in traits between hatchery and wild salmon can also be explored using DNA methylation patterns to identify any epigenetic programming associations, which may be due to the differences in the rearing environment. Pre-release juvenile epigenetic effects may be examined for reversibility, persistence, and transgenerational potential using specific comparisons across life stages (i.e. reversibility: hatchery-wild differences as post-release juvenile, and persistence: juvenile vs. adult) and across generations (i.e. transgenerational potential: hatchery-wild differences as adult reproductive cells and offspring as fertilized eggs). Conceivably, semi-natural rearing techniques may decrease the amount of epigenetic effects between hatchery and wild salmon. Examining the number of epigenetic programming differences with comparisons of hatchery vs. wild salmon and semi-natural vs. natural salmon can help fill this knowledge gap. Hatcheries may benefit from the additional mechanistic insight explaining the lower performance of hatchery than wild salmon. Recently, juvenile **epigenetic programming differences** because of differences in the rearing environment have been suggested as a potential mechanism; however, there are questions on the epigenetic timeline and whether semi-natural rearing techniques can decrease the amount of epigenetic effects. In light of minimizing impacts on populations, hatcheries may benefit from further insight on transgenerational epigenetic effects, as maladapted traits may be passed to the next generation from hatchery salmon breeding in the natural environment or interbreeding with wild salmon.

Photo by: Sam James

8. SUMMARY AND CONCLUSION

Salmon research has progressed from using genetics to genomics technologies, such that there is an increase in statistical power of traditional analyses (i.e. broad-sense genomics) based on the higher number of markers, as well as opportunities for novel analyses (i.e. narrow-sense genomics) based on improved technology (e.g. high-throughput platforms), references (e.g. quality genomes), and databases (e.g. functional annotation). Beyond genomics, salmon research also uses transcriptomics technologies, examining the changes in gene expression. However, aquaculture salmon research has progressed further than hatchery salmon research by using newer proteomics and metabolomics technologies, with the benefit of closer ties to salmon traits or phenotypes. Furthermore, environmental DNA from water samples can be used for assessing species distribution and abundance. There may be opportunities for examining other environmental molecules, i.e. mRNA, proteins, and metabolites, in water samples as a non-invasive approach for monitoring expression biomarkers to assess the health and condition of hatchery salmon within a tank. Although there are reviews of 'omics' technology for fisheries and aquaculture, there has been no specific review for hatcheries. In this review, I have also highlighted the significant knowledge gaps in our understanding of how to apply these new technologies to best improve hatchery salmon performance.

8.1 > DELIVERABLES AND KNOWLEDGE GAPS

Given the current research focus on salmon genomics and transcriptomics technologies, there are several examples of hatchery deliverables. A broad-sense genomics tool that shows up repeatedly is parentage-based tagging (PBT), because of the amount of new information that can be provided for assessing hatchery performance and addressing knowledge gaps. In contrast, the newer proteomics and metabolomics technologies have fewer examples of hatchery deliverables so far. There are also potential hatchery deliverables that could be derived using environmental RNA, but these are theoretical and require further research to demonstrate feasibility. Knowledge gaps were identified, and experiments were described using 'omics' technology to gather more information, with the objective of increasing our understanding of how to apply these new technologies to best improve hatchery salmon performance. Organized by main themes, below are lists of the deliverables and knowledge gaps, with short summaries of the applicable 'omics' technologies and key considerations. At the end, there is are two tables summarizing the deliverables (1) applicable to the different hatchery types and (2) applicable to all hatcheries.

8.1.1 HATCHERY PERFORMANCE

Deliverables:

- Fishery and escapement contributions (section 2.1), e.g. % contribution and offspring/spawner ratio for a fishery or escapement. Uses PBT.
- Other hatchery metrics (section 2.2), e.g. % contribution and offspring/spawner ratio for straying, age determination, run timing, and release group strategy. Uses PBT with hatchery records. Sex determination, non-target species, and inter-species hybridization are available using SNP genotypes. Pedigree construction, inbreeding coefficients, and pairwise relatedness are available using PBT and SNP genotypes across the genome.
- Genetic connection (section 2.4), e.g. hatchery genetic introgression rate and hatchery-wild genetic differentiation. Uses SNP markers across the genome.
- Species distribution and abundance (section 3.4), e.g. % species detection at sites and rough estimate of salmon count or biomass. Uses eDNA collected from water samples across the landscape.
- Living and dead components (section 4.3), e.g. theoretical % live species detection at sites and rough estimate of live salmon count or biomass. Uses eRNA and eDNA collected from water samples across the landscape. Paired with eDNA, rough estimate of live/dead ratio for spawning time.

Knowledge gaps on group performance for hatchery release timing and rearing techniques:

- Ecological smolt window (section 7.1.1), e.g. survival of different release timing groups around zooplankton blooms. Uses PBT with hatchery planning for the families representing the groups.
- Physiological smolt window (section 7.1.2), e.g. pair smolt status with targeted release date. Uses PBT with hatchery planning for the families representing the groups, as well as smoltification gene expression biomarkers for smolt status. Examined in concert with ecological smolt window.
- Semi-natural rearing (section 7.1.3), e.g. survival of traditional hatchery vs. semi-natural rearing groups. Uses PBT with hatchery planning for the families representing the groups.

8.1.2 OFFSPRING QUALITY

Deliverables:

- Trait architecture (section 2.3), e.g. % genetic and maternal, heritability (% additive genetic), as well as mate choice type (i.e. beneficial genes vs. compatible genes). Uses PBT with a full factorial breeding design and trait measures. Genotype × environment interactions and genetic correlations are available using PBT and trait measures.
- Gene biomarkers (section 3.1), e.g. age-at-maturity and run timing prediction, as well as thermal tolerance basis. Uses SNP markers across the genome to compare between groups.
- Trait epigenetic programming (section 3.2), e.g. precocious parr vs. adult. Uses DNA methylation patterns across the genome, such as RRBS technique, to compare between trait groups.

Knowledge gaps for improving offspring quality:

- Mate choice benefits (section 7.2.1), e.g. % higher offspring survival for mate choice than non-mate choice. Uses spawning channels with behavioural observation criteria, as well as PBT with a full factorial mating design and trait measures.
- Trait architecture (section 7.2.2), e.g. contribution to fishery or escapement, smolt to adult traits. Uses trait architecture method.
- Brood stock biomarkers (section 7.2.3), e.g. screen brood stock for female value, male value, and female-male pair value. Follow-up to trait architecture results to identify the specific genes, proteins, and metabolites. Uses gene biomarker method for large genetic effects. Uses protein expression biomarkers and metabolite expression biomarkers methods for large maternal effects.

8.1.3 HEALTH AND CONDITION BIOMARKERS

Deliverables:

- Gene expression biomarkers (section 4.1), e.g. smoltification, stress response, and viral disease development. Discovery process uses mRNA markers across the transcriptome to compare between groups and identify candidate biomarkers. Validation or monitoring process uses qPCR for biomarkers, e.g. 'Salmon Fit-Chip'. Same technology can be used for examining microbe loads.
- Protein expression biomarkers (section 5.1), e.g. disease diagnosis, vaccine targets, egg quality, and sperm quality. Uses protein markers across the proteome to compare between groups and identify candidate biomarkers.
- Metabolite expression biomarkers (section 6.1), e.g. disease diagnosis, long-term stress, starvation or fasting, and toxin exposure. Uses metabolite markers across the metabolome to compare between groups and identify candidate biomarkers.
- Environmental gene expression (section 4.4), e.g. theoretical non-invasive gene expression. Uses the mRNA component of eRNA in water samples collected from tank with gene expression biomarkers developed for tissues.

Knowledge gaps on biomarker development for aspects of hatchery salmon health and condition:

- Choice of tissue and 'omics' technology (section 7.3.1), e.g. reproductive development biomarkers. Uses background guidance or exploratory approach for the selection of tissues and expression biomarker methods.
- Robust biomarkers (section 7.3.2), e.g. across environmental conditions. Discovery process uses 'omics' wide screen to identify candidate biomarkers; validation process uses new samples. Extension for discovery uses meta-analysis to identify candidate biomarkers; extension for validation process uses a range of samples, e.g. several species, life stages, and environmental conditions.
- Non-invasive biomarkers (section 7.3.3), e.g. expression biomarkers. Uses the environmental mRNA, proteins, and metabolites in water samples collected from a tank with expression biomarkers designed for tissues.

8.1.4 HATCHERY-WILD DIFFERENCES

Deliverables:

- Epigenetic programming differences (section 3.3), e.g. number of DNA methylation differences at genes and the key biological processes involved (i.e. mechanistic explanations). Uses DNA methylation patterns across the genome to compare between salmon (i.e. hatchery vs. wild) or rearing (i.e. hatchery vs. natural) groups.
- Gene expression differences (section 4.2), e.g. number of mRNA differences at genes and the key biological processes involved. Uses mRNA markers across the transcriptome to compare groups.
- Protein expression differences (section 5.2), e.g. number of protein differences and the key biological processes involved. Uses protein markers across the proteome to compare groups.
- Metabolite expression differences (section 6.2), e.g. number of metabolite differences and the key biological processes involved. Uses metabolite markers across the metabolome to compare groups.

Knowledge gaps on mechanistic explanations, especially epigenetic effects:

- Mechanistic explanations (section 7.4.1), e.g. proteins and metabolites have closer ties to traits or phenotypes. Uses protein expression differences and metabolite expression methods.
- Trait epigenetic programming (section 7.4.2), e.g. jimmy or jack vs. adult. Uses trait epigenetic programming method.
- Epigenetic timeline (section 7.4.3), e.g. persistence, reversibility, and/or transgenerational potential of epigenetic effects. Uses specific comparisons across life stages (i.e. reversibility: hatchery-wild differences as post-release juvenile, and persistence: juvenile vs. adult) and across generations (i.e. transgenerational potential: hatchery-wild differences as adult reproductive cells and offspring as fertilized eggs). Uses DNA methylation patterns across the genome between groups.
- Semi-natural rearing (section 7.4.4), e.g. % decrease of epigenetic effects for semi-natural than traditional hatchery. Uses DNA methylation patterns across the genome between hatchery vs. wild salmon and semi-natural vs. wild salmon.

Table 4. Summary of deliverables applicable to the different hatchery types across 'omics' technologies.

DELIVERABLE	METHOD	UTILITY
Augmentation hatchery		
Fishery contributions	PBT	Hatchery salmon counts to fishery
Genetic introgression	SNP markers	Hatchery salmon genetic influences on wild salmon
Supplementation and captive breeding hatcheries		
Escapement contributions	PBT	Hatchery salmon counts to escapement
Pedigrees	PBT with several generations	Avoid highly unequal family sizes and inbreeding depression. Expected inbreeding coefficients and relatedness values
Non-target species contributions	SNP genotypes	Limit inter-species breeding, brood stock level
Inter-species hybridization contributions	SNP genotypes	Limit inter-species breeding, offspring level
Realized inbreeding coefficients	SNP genotypes	Avoid inbreeding depression
Realized relatedness values	SNP genotypes	Avoid inbreeding depression
Genetic differentiation	SNP markers	Limit hatchery-wild genetic divergence because of domestication selection or genetic drift
Species distribution	eDNA across landscape	Locations can be prioritized based on salmon presence for assigning limited hatchery resources

Species abundance	eDNA with surveys	Rough estimate of population size over time in natural environments
Thermal stress biomarkers	qPCR	Physiological impact evaluations at the individual or salmon level
Hypoxic stress biomarkers	qPCR	Physiological impact evaluations at the individual or salmon level
Salinity stress biomarkers	qPCR	Physiological impact evaluations at the individual or salmon level
Living/dead species distribution (theoretical)	eRNA and eDNA across landscape	Locations of live and dead salmon during spawning run
Living/dead species abundance (theoretical)	eRNA and eDNA with surveys	Rough estimate of live and dead salmon during spawning run, live/dead ratio for spawning time

Table 5. Summary of deliverables applicable to all hatchery types by 'omics' technology.

DELIVERABLE	METHOD	UTILITY
Broad-sense genomics		
Straying	PBT	Hatchery salmon in non-target locations
Age determination	PBT	Age confirmation, etc.
Group performance	PBT with loading plan	Release group counts to fishery, etc.
Sex determination	SNP genotypes (PBT)	Sex confirmation, etc.
Additive genetic variance (or heritability)	Full factorial breeding with PBT	Breeding protocols/offspring quality, evolutionary potential, 'good genes' mating system
Non-additive genetic variance	Full factorial breeding with PBT	Breeding protocols/offspring quality, 'compatible genes' mating system
Maternal variance	Full factorial breeding with PBT	Breeding protocols/offspring quality, importance of maternal effects
Genotype × environment interactions	Single pair breeding with PBT	Breeding protocol/offspring quality, gene or genotype quality in natural environments
Genetic correlations	Parent-offspring with PBT	Breeding protocol/offspring quality, predict trait responses to selection, e.g. breeding or fishing
Narrow-sense genomics		
Gene biomarkers	SNP markers across genome	Identify the gene biomarkers underlying valued salmon traits that may serve as targets for hatchery breeding
Trait epigenetic programming	RRBS	DNA methylation patterns explaining salmon traits. Can these be managed in the hatchery setting?
Epigenetic programming differences	RRBS	DNA methylation differences between hatchery and wild salmon or rearing environments. Can these be managed in the hatchery setting?

Transcriptomics		
Smoltification biomarkers	qPCR	Refining juvenile release strategies, e.g. releasing the juveniles when they are majority smolt classification
General stress or imminent mortality biomarkers	qPCR	Optimize practices to minimize stress for salmon
Viral disease development (VDD) biomarkers and microbe load assays	qPCR	Identify and mitigate exposure to infectious agents
Gene expression biomarkers (discovery)	RNA-seq	Identify the biomarkers underlying important salmon traits, e.g. reproductive development
Gene expression differences	RNA-seq	Differences between hatchery and wild salmon or groups within hatchery setting
Environmental gene expression (theoretical)	eRNA, qPCR	Biomarkers without handling the salmon, e.g. limit stress response

Proteomics		
IHN or BKD biomarkers (candidate)	Mass spectrometry	Disease diagnostics
Antigen proteins	Mass spectrometry	Vaccine development
Egg quality biomarkers (candidate)	Mass spectrometry	Reproductive quality of females
Protein expression biomarkers (discovery)	Mass spectrometry	Identify the biomarkers underlying important salmon traits, e.g. other hatchery disease concerns
Protein expression differences	Mass spectrometry	Differences between hatchery and wild salmon or groups within hatchery setting

Metabolomics		
A. *salmonicida* infection biomarkers (candidate)	MS	Disease diagnostics
Long-term stress (candidate)	MS	Optimize practices to minimize stress for salmon
Metabolite expression biomarkers (discovery)	NMR	Identify the biomarkers underlying important salmon traits, e.g. other hatchery disease concerns
Metabolite expression differences	NMR	Differences between hatchery and wild salmon or groups within hatchery setting

8.2 > CONCLUSION

The different production objectives for hatcheries (augmentation, supplementation, and captive breeding) each have different goals, but each can benefit from the use of 'omics' technologies. Augmentation hatcheries would benefit from an increased understanding of how to improve survival and production but may be limited by the impacts to local natural populations. Supplementation hatcheries may have similar goals for survival, less for production, but increased interests for genetic diversity within the hatchery and natural populations. Captive breeding for restoration of a natural population must increase the numbers of adults and protect the genetic traits and diversity within the enhanced production. The use of **parentage-based tagging** (PBT) families and the goals to increase smolt readiness and survival overlaps between the three hatchery production objectives.

An increased understanding of the higher early marine survival of salmon and minimizing negative impacts on natural populations are applicable across all hatchery types. In this review, I have promoted that trait architecture, using a full factorial breeding design and separate families tagged by PBT, can provide estimates of the underlying genetic (i.e. additive and non-additive) and maternal effects for important salmon traits, e.g. contribution to fishery or escapement. Large effects may be pursued at the molecular level using 'omics' technology for identifying specific targets, e.g. genes. PBT can be used for several hatchery performance metrics, e.g. % contribution and offspring/spawner ratio, as well as performance by other grouping, e.g. release timing, for fisheries, escapement, reproductive success, and early marine survival. The four 'omics' technologies can be used for gathering relevant information at the molecular level, i.e. genes, mRNA, proteins, and metabolites, for salmon trait associations (e.g. biomarkers) and salmon trait differences between groups (e.g. hatchery vs. wild).

Biomarkers may be used as indicators for brood stock values and physiological measures for improving hatchery performance via offspring quality and salmon health and condition. The biomarkers can also provide valuable predictive screening information of hatchery salmon prior to release, which may be used to explain patterns in hatchery performance over time or differences among hatcheries. Comparisons of hatchery and wild salmon can reveal the number of differences and the key biological processes involved (i.e. mechanistic explanations). Conceivably, hatcheries may minimize the impacts on natural populations by decreasing the amount of hatchery -wild epigenetic effects using semi-natural rearing techniques. This can be tested using the number of DNA methylation pattern differences between hatchery vs. wild salmon compared to semi-natural vs. natural salmon at the life stage just prior to release.

Filling in the four knowledge gaps identified in this review, namely (1) group performance, (2) offspring quality, (3) health and condition biomarkers, and (4) hatchery-wild differences, using 'omics' technologies will increase the understanding of how to improve hatchery performance and ultimately increase the hatchery salmon contributions to fisheries, escapement, and reproductive success. Furthermore, there can be an increase in the understanding of how to minimize negative impacts on natural populations, e.g. limit transgenerational epigenetic effects for hatchery salmon breeding in the natural environment, as well as for hatchery salmon interbreeding with wild salmon.

9. LITERATURE CITED

Adams, C.I.M., Knapp, M., Gemmell, N.J., Jeunen, G.-J., Bunce, M., Lamare, M.D., and Taylor, H.R. 2019. Beyond biodiversity: Can environmental DNA (eDNA) cut it as a population genetics tool? Genes 10(3):192. DOI: 10.3390/genes10030192

Akbarzadeh, A., Günther, O.P., Houde, A.L., Li, S., Ming, T.J., Jeffries, K.M., Hinch, S.G., and Miller, K.M. 2018. Developing specific molecular biomarkers for thermal stress in salmonids. BMC Genomics 19(1):749. DOI: 10.1186/s12864-018-5108-9

Alfaro, A.C. and Young, T. 2018. Showcasing metabolomic applications in aquaculture: A review. Reviews in Aquaculture 10(1):135–152. DOI: 10.1111/raq.12152

Allendorf, F.W., Leary, R.F., Spruell, P., and Wenburg, J.K. 2001. The problems with hybrids: Setting conservation guidelines. Trends in Ecology & Evolution 16(11):613–622. DOI: 10.1016/S0169-5347(01)02290-X

Amish, S.J., Ali, O., Peacock, M., Miller, M., Robinson, M., Smith, S., Luikart, G., and Neville, H. 2019. Assessing thermal adaptation using family-based association and F_{ST} outlier tests in a threatened trout species. Molecular Ecology 28(10):2573–2593. DOI: 10.1111/mec.15100

Anderson, J.H., Warheit, K.I., Craig, B.E., Seamons, T.R., and Haukenes, A.H. 2020. A review of hatchery reform science in Washington State. Washington Fish and Wildlife Commission.

Araki, H., Berejikian, B.A., Ford, M.J., and Blouin, M.S. 2008. Fitness of hatchery-reared salmonids in the wild: Fitness of hatchery fish. Evolutionary Applications 1(2):342–355. DOI: 10.1111/j.1752-4571.2008.00026.x

Araki, H., Cooper, B., and Blouin, M.S. 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318(5847):100–103. DOI: 10.1126/science.1145621

Araki, H. and Schmid, C. 2010. Is hatchery stocking a help or harm? Aquaculture 308:S2–S11. DOI: 10.1016/j.aquaculture.2010.05.036

Aursand, M., Standal, I.B., Præl, A., McEvoy, L., Irvine, J., and Axelson, D.E. 2009. ^{13}C NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed, and geographical origin. Journal of Agricultural and Food Chemistry 57(9):3444–3451. DOI: 10.1021/jf8039268

Baillie, S.M., McGowan, C., May-McNally, S., Leggatt, R., Sutherland, B.J.G., and Robinson, S. 2019. Environmental DNA and its applications to Fisheries and Oceans Canada: National needs and priorities. Canadian Technical Report of Fisheries and Aquatic Sciences 3329.

Banta, J.A. and Richards, C.L. 2018. Quantitative epigenetics and evolution. Heredity 121(3):210–224. DOI: 10.1038/s41437-018-0114-x

Barson, N.J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G.H., Fiske, P., Jacq, C., Jensen, A.J., Johnston, S.E., Karlsson, S., Kent, M., Moen, T., Niemelä, E., Nome, T., Næsje, T.F., Orell, P., Romakkaniemi, A., Sægrov, H., Urdal, K., Erkinaro, J., Lien, S., and Primmer, C.R. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528(7582):405–408. DOI: 10.1038/nature16062

Beacham, T.D., Wallace, C., Jonsen, K., McIntosh, B., Candy, J.R., Willis, D., Lynch, C., Moore, J.-S., Bernatchez, L., and Withler, R.E. 2019a. Comparison of coded-wire tagging with parentage-based tagging and genetic stock identification in a large-scale coho salmon fisheries application in British Columbia, Canada. Evolutionary Applications 12(2):230–254. DOI: 10.1111/eva.12711

Beacham, T.D., Wallace, C., Jonsen, K., McIntosh, B., Candy, J.R., Willis, D., Lynch, C., and Withler, R.E. 2019b. Variation in migration pattern, broodstock origin, and family productivity of coho salmon hatchery populations in British Columbia, Canada, derived from parentage-based tagging. Ecology and Evolution 9(17):9891–9906. DOI: 10.1002/ece3.5530

Beacham, T.D., Wallace, C., MacConnachie, C., Jonsen, K., McIntosh, B., Candy, J.R., Devlin, R.H., and Withler, R.E. 2017. Population and individual identification of coho salmon in British Columbia through parentage-based tagging and genetic stock identification: An alternative to coded-wire tags. Canadian Journal of Fisheries and Aquatic Sciences 74(9):1391–1410. DOI: 10.1139/cjfas-2016-0452

Beacham, T.D., Wallace, C., MacConnachie, C., Jonsen, K., McIntosh, B., Candy, J.R., and Withler, R.E. 2018. Population and individual identification of Chinook salmon in British Columbia through parentage-based tagging and genetic stock identification with single nucleotide polymorphisms. Canadian Journal of Fisheries and Aquatic Sciences 75(7):1096–1105. DOI: 10.1139/cjfas-2017-0168

Beacham, T.D. and Wallace, C.G. 2019. Salmon species identification via direct DNA sequencing of single amplicons. Conservation Genetics Resources . DOI: 10.1007/s12686-019-01102-1. (http://link.springer.com/10.1007/s12686-019-01102-1)

Beamish, R.J., Sweeting, R.M., Lange, K.L., Noakes, D.J., Preikshot, D., and Neville, C.M. 2010. Early marine survival of coho salmon in the Strait of Georgia declines to very low Levels. Marine and Coastal Fisheries 2(1):424–439. DOI: 10.1577/C09-040.1

Beamish, R.J., Sweeting, R.M., and Neville, C. 2009. Plannning the management of Pacific Salmon in a changing climate. American Fisheries Society Symposium 69:155–173.

Benskin, J.P., Ikonomou, M.G., Liu, J., Veldhoen, N., Dubetz, C., Helbing, C.C., and Cosgrove, J.R. 2014. Distinctive metabolite profiles in in-migrating sockeye salmon suggest sex-linked endocrine perturbation. Environmental Science & Technology 48(19):11670–11678. DOI: 10.1021/es503266x

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D.T., Barth, J.M.I., Beacham, T.D., Maes, G.E., Martinsohn, J.T., Miller, K.M., Naish, K.A., Ovenden, J.R., Primmer, C.R., Young Suk, H., Therkildsen, N.O., and Withler, R.E. 2017. Harnessing the power of genomics to secure the future of seafood. Trends in Ecology & Evolution 32(9):665–680. DOI: 10.1016/j.tree.2017.06.010

Best, C., Ikert, H., Kostyniuk, D.J., Craig, P.M., Navarro-Martin, L., Marandel, L., and Mennigen, J.A. 2018. Epigenetics in teleost fish: From molecular mechanisms to physiological phenotypes. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 224:210–244. DOI: 10.1016/j.cbpb.2018.01.006

Björnsson, B.T., Stefansson, S.O., and McCormick, S.D. 2011. Environmental endocrinology of salmon smoltification. General and Comparative Endocrinology 170(2):290–298. DOI: 10.1016/j.ygcen.2010.07.003

Bond, C.E. 1996. Biology of Fishes, 2nd ed Edition. Thomson Learning, Toronto, Ontario.

Booy, A.T., Haddow, J.D., Ohlund, L.B., Hardie, D.B., and Olafson, R.W. 2005. Application of isotope coded affinity tag (ICAT) analysis for the identification of differentially expressed proteins following Infection of Atlantic salmon (*Salmo salar*) with infectious hematopoietic necrosis virus (IHNV) or *Renibacterium salmoninarum* (BKD). Journal of Proteome Research 4(2):325–334. DOI: 10.1021/pr049840t

Bromaghin, J.F., Nielson, R.M., and Hard, J.J. 2011. A model of Chinook salmon population dynamics incorporating size-selective exploitation and inheritance of polygenic correlated traits. Natural Resource Modeling 24(1):1–47. DOI: 10.1111/j.1939-7445.2010.00077.x

Carson, H.L. 1990. Increased genetic variance after a population bottleneck. Trends in Ecology & Evolution 5(7):228–230. DOI: 10.1016/0169-5347(90)90137-3

Chittenden, C.M., Jensen, J.L.A., Ewart, D., Anderson, S., Balfry, S., Downey, E., Eaves, A., Saksida, S., Smith, B., Vincent, S., Welch, D., and McKinley, R.S. 2010. Recent salmon declines: A result of lost feeding opportunities due to bad timing? PLoS ONE 5(8):e12423. DOI: 10.1371/journal.pone.0012423

Chittenden, C.M., Sura, S., Butterworth, K.G., Cubitt, K.F., Plantalech Manel-la, N., Balfry, S., Økland, F., and McKinley, R.S. 2008. Riverine, estuarine and marine migratory behaviour and physiology of wild and hatchery-reared coho salmon *Oncorhynchus kisutch* (Walbaum) smolts descending the Campbell River, BC, Canada. Journal of Fish Biology 72(3):614–628. DOI: 10.1111/j.1095-8649.2007.01729.x

Christie, M.R., Ford, M.J., and Blouin, M.S. 2014. On the reproductive success of early-generation hatchery fish in the wild. Evolutionary Applications 7(8):883–896. DOI: 10.1111/eva.12183

Christie, M.R., Marine, M.L., Fox, S.E., French, R.A., and Blouin, M.S. 2016. A single generation of domestication heritably alters the expression of hundreds of genes. Nature Communications 7(1):10676. DOI: 10.1038/ncomms10676

Cipriano, R.C., Smith, M.L., Vermeersch, K.A., Dove, A.D.M., and Styczynski, M.P. 2015. Differential metabolite levels in response to spawning-induced inappetence in Atlantic salmon *Salmo salar*. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 13:52–59. DOI: 10.1016/j.cbd.2015.01.001

Connon, R.E., Jeffries, K.M., Komoroske, L.M., Todgham, A.E., and Fangue, N.A. 2018. The utility of transcriptomics in fish conservation. The Journal of Experimental Biology 221(2):jeb148833. DOI: 10.1242/jeb.148833

Cristescu, M.E. 2019. Can environmental RNA revolutionize biodiversity science? Trends in Ecology & Evolution 34(8):694–697. DOI: 10.1016/j.tree.2019.05.003

Crnokrak, P. and Roff, D.A. 1995. Dominance variance: Associations with selection and fitness. Heredity 75(5):530–540. DOI: 10.1038/hdy.1995.169

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., Vere, N., Pfrender, M.E., and Bernatchez, L. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology 26(21):5872–5895. DOI: 10.1111/mec.14350

DFO. 2005. Canada's Policy for Conservation of Wild Pacific Salmon. Fisheries and Oceans Canada, Vancouver. 49 pp.

DFO. 2013. A Biological Risk Management Framework for Enhancing Salmon in the Pacific Region. Salmon Enhancement Program, Pacific Region.

Dumetz, F., Duchaud, E., Claverol, S., Orieux, N., Papillon, S., Lapaillerie, D., and Le Henaff, M. 2008. Analysis of the *Flavobacterium psychrophilum* outer-membrane subproteome and identification of new antigenic targets for vaccine by immunomics. Microbiology 154(6):1793–1801. DOI: 10.1099/mic.0.2008/016600-0

Eldridge, W.H., Hard, J.J., and Naish, K.A. 2010. Simulating fishery-induced evolution in chinook salmon: the role of gear, location, and genetic correlation among traits. Ecological Applications 20(7):1936–1948. DOI: 10.1890/09-1186.1

Ellis, T., James, J.D., Stewart, C., and Scott, A.P. 2004. A non-invasive stress assay based upon measurement of free cortisol released into the water by rainbow trout. Journal of Fish Biology 65(5):1233–1252. DOI: 10.1111/j.0022-1112.2004.00499.x

Evans, M.L., Hori, T.S., Rise, M.L., and Fleming, I.A. 2015. Transcriptomic responses of Atlantic salmon (*Salmo salar*) to environmental enrichment during juvenile rearing. PLOS ONE 10(3):e0118378. DOI: 10.1371/journal.pone.0118378

Falconer, D.S. and Mackay, T.F.C. 2009. Introduction to quantitative genetics, 4. ed. Edition. Pearson, Prentice Hall, Harlow.

Feder, M.E. and Walser, J.-C. 2005. The biological limitations of transcriptomics in elucidating stress and stress responses. Journal of Evolutionary Biology 18(4):901–910. DOI: 10.1111/j.1420-9101.2005.00921.x

Fisch, K.M., Kozfkay, C.C., Ivy, J.A., Ryder, O.A., and Waples, R.S. 2015. Fish hatchery genetic management techniques: Integrating theory with implementation. North American Journal of Aquaculture 77(3):343–357. DOI: 10.1080/15222055.2014.999846

Fleming, I.A., Lamberg, A., and Jonsson, B. 1997. Effects of early experience on the reproductive performance of Atlantic salmon. Behavioral Ecology 8(5):470–480. DOI: 10.1093/beheco/8.5.470

Forné, I., Abián, J., and Cerdà, J. 2010. Fish proteome analysis: Model organisms and non-sequenced species. Proteomics 10(4):858–872. DOI: 10.1002/pmic.200900609

Frankham, R. 2005. Genetics and extinction. Biological Conservation 126(2):131–140. DOI: 10.1016/j.biocon.2005.05.002

Frankham, R. 2008. Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17(1):325–333. DOI: 10.1111/j.1365-294X.2007.03399.x

Fraser, D.J. 2008. How well can captive breeding programs conserve biodiversity? A review of salmonids. Evolutionary Applications 1(4):535–586. DOI: 10.1111/j.1752-4571.2008.00036.x

Garner, B.A., Hand, B.K., Amish, S.J., Bernatchez, L., Foster, J.T., Miller, K.M., Morin, P.A., Narum, S.R., O'Brien, S.J., Roffler, G., Templin, W.D., Sunnucks, P., Strait, J., Warheit, K.I., Seamons, T.R., Wenburg, J., Olsen, J., and Luikart, G. 2016. Genomics in conservation: Case studies and bridging the gap between data and application. Trends in Ecology & Evolution 31(2):81–83. DOI: 10.1016/j.tree.2015.10.009

Gavery, M.R., Nichols, K.M., Berejikian, B.A., Tatara, C.P., Goetz, G.W., Dickey, J.T., Van Doornik, D.M., and Swanson, P. 2019. Temporal dynamics of DNA methylation patterns in response to rearing juvenile steelhead (*Oncorhynchus mykiss*) in a hatchery versus simulated stream environment. Genes 10(5):356. DOI: 10.3390/genes10050356

Gavery, M.R., Nichols, K.M., Goetz, G.W., Middleton, M.A., and Swanson, P. 2018. Characterization of genetic and epigenetic variation in sperm and red blood cells from adult hatchery and natural-origin steelhead, *Oncorhynchus mykiss*. G3 Genes|Genomes|Genetics 8(11):3723–3736. DOI: 10.1534/g3.118.200458

Gombar, R., Pitcher, T.E., Lewis, J.A., Auld, J., and Vacratsis, P.O. 2017. Proteomic characterization of seminal plasma from alternative reproductive tactics of Chinook salmon (*Oncorhynchus tswatchysha*). Journal of Proteomics 157:1–9. DOI: 10.1016/j.jprot.2017.01.019

Gotelli, N.J., Ellison, A.M., and Ballif, B.A. 2012. Environmental proteomics, biodiversity statistics and food-web structure. Trends in Ecology & Evolution 27(8):436–442. DOI: 10.1016/j.tree.2012.03.001

Hess, J.E., Zendt, J.S., Matala, A.R., and Narum, S.R. 2016. Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing. Proceedings of the Royal Society B: Biological Sciences 283(1830):20153064. DOI: 10.1098/rspb.2015.3064

Houde, A.L., Wilson, C.C., and Neff, B.D. 2013. Genetic architecture of survival and fitness-related traits in two populations of Atlantic salmon. Heredity 111(6):513–519. DOI: 10.1038/hdy.2013.74

Houde, A.L.S., Akbarzadeh, A., Günther, O.P., Li, S., Patterson, D.A., Farrell, A.P., Hinch, S.G., and Miller, K.M. 2019a. Salmonid gene expression biomarkers indicative of physiological responses to changes in salinity and temperature, but not dissolved oxygen. The Journal of Experimental Biology 222(13):jeb198036. DOI: 10.1242/jeb.198036

Houde, A.L.S., Black, C.A., Wilson, C.C., Pitcher, T.E., and Neff, B.D. 2015. Genetic and maternal effects on juvenile survival and fitness-related traits in three populations of Atlantic salmon. Canadian Journal of Fisheries and Aquatic Sciences 72(5):751–758. DOI: 10.1139/cjfas-2014-0472

Houde, A.L.S., Günther, O.P., Strohm, J., Ming, T.J., Li, S., Kaukinen, K.H., Patterson, D.A., Farrell, A.P., Hinch, S.G., and Miller, K.M. 2019b. Discovery and validation of candidate smoltification gene expression biomarkers across multiple species and ecotypes of Pacific salmonids. Conservation Physiology 7(1):coz051. DOI: 10.1093/conphys/coz051

Houde, A.L.S. and Pitcher, T.E. 2016. fullfact: An R package for the analysis of genetic and maternal variance components from full factorial mating designs. Ecology and Evolution 6(6):1656–1665. DOI: 10.1002/ece3.1943

Houde, A.L.S., Schulze, A.D., Kaukinen, K.H., Strohm, J., Patterson, D.A., Beacham, T.D., Farrell, A.P., Hinch, S.G., and Miller, K.M. 2019c. Transcriptional shifts during juvenile coho salmon (*Oncorhynchus kisutch*) life stage changes in freshwater and early marine environments. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 29:32–42. DOI: 10.1016/j.cbd.2018.10.002

Houde, A.L.S., Wilson, C.C., and Pitcher, T.E. 2016. Genetic architecture and maternal contributions of early-life survival in lake trout *Salvelinus namaycush*. Journal of Fish Biology 88(5):2088–2094. DOI: 10.1111/jfb.12965

Huntingford, F.A. 2004. Implications of domestication and rearing conditions for the behaviour of cultivated fishes. Journal of Fish Biology 65(Suppl. A):122–142.

Jasper, J.R., Habicht, C., Moffitt, S., Brenner, R., Marsh, J., Lewis, B., Creelman Fox, E., Grauvogel, Z., Rogers Olive, S.D., and Grant, W.S. 2013. Source-sink estimates of genetic introgression dhow influence of hatchery strays on wild chum salmon populations in Prince William Sound, Alaska. PLoS ONE 8(12):e81916. DOI: 10.1371/journal.pone.0081916

Jeffries, K.M., Hinch, S.G., Gale, M.K., Clark, T.D., Lotto, A.G., Casselman, M.T., Li, S., Rechisky, E.L., Porter, A.D., Welch, D.W., and Miller, K.M. 2014. Immune response genes and pathogen presence predict migration survival in wild salmon smolts. Molecular Ecology 23(23):5803–5815. DOI: 10.1111/mec.12980

Jeffries, K.M., Hinch, S.G., Sierocinski, T., Clark, T.D., Eliason, E.J., Donaldson, M.R., Li, S., Pavlidis, P., and Miller, K.M. 2012. Consequences of high temperatures and premature mortality on the transcriptome and blood physiology of wild adult sockeye salmon (*Oncorhynchus nerka*). Ecology and Evolution 2(7):1747–1764. DOI: 10.1002/ece3.274

Johnson, S.L., Villarroel, M., Rosengrave, P., Carne, A., Kleffmann, T., Lokman, P.M., and Gemmell, N.J. 2014. Proteomic analysis of Chinook salmon (*Oncorhynchus tshawytscha*) ovarian fluid. PLoS ONE 9(8):e104155. DOI: 10.1371/journal.pone.0104155

Jones, O.A.H., Maguire, M.L., Griffin, J.L., Dias, D.A., Spurgeon, D.J., and Svendsen, C. 2013. Metabolomics and its use in ecology. Austral Ecology 38(6):713–720. DOI: 10.1111/aec.12019

Jonsson, B. and Jonsson, N. 2014. Early environment influences later performance in fishes: Effects of early experiences. Journal of Fish Biology 85(2):151–188. DOI: 10.1111/jfb.12432

Kanerva, M., Vehmas, A., Nikinmaa, M., and Vuori, K.A. 2014. Spatial variation in transcript and protein abundance of Atlantic salmon during feeding migration in the Baltic Sea. Environmental Science & Technology 48(23):13969–13977. DOI: 10.1021/es502956g

Karakach, T.K., Huenupi, E.C., Soo, E.C., Walter, J.A., and Afonso, L.O.B. 2009. 1H-NMR and mass spectrometric characterization of the metabolic response of juvenile Atlantic salmon (*Salmo salar*) to long-term handling stress. Metabolomics 5(1):123–137. DOI: 10.1007/s11306-008-0144-0

Kido Soule, M.C., Longnecker, K., Johnson, W.M., and Kujawinski, E.B. 2015. Environmental metabolomics: Analytical strategies. Marine Chemistry 177:374–387. DOI: 10.1016/j.marchem.2015.06.029

Lacoursière-Roussel, A., Rosabal, M., and Bernatchez, L. 2016. Estimating fish abundance and biomass from eDNA concentrations: Variability among capture methods and environmental conditions. Molecular Ecology Resources 16(6):1401–1414. DOI: 10.1111/1755-0998.12522

Lahnsteiner, F., Berger, B., Weismann, T., and Patzner, R.A. 1998. Determination of semen quality of the rainbow trout, *Oncorhynchus mykiss*, by sperm motility, seminal plasma parameters, and spermatozoal metabolism. Aquaculture 163(1–2):163–181. DOI: 10.1016/S0044-8486(98)00243-9

Lahnsteiner, F., Weismann, T., and Patzner, R. 1995. Composition of the ovarian fluid in 4 salmonid species: *Oncorhynchus mykiss*, *Salmo trutta f lacustris*, *Salvelinus alpinus* and *Hucho hucho*. Reproduction Nutrition Development 35(5):465–474. DOI: 10.1051/rnd:19950501

Lankadurai, B.P., Nagato, E.G., and Simpson, M.J. 2013. Environmental metabolomics: An emerging approach to study organism responses to environmental stressors. Environmental Reviews 21(3):180–205. DOI: 10.1139/er-2013-0011

Laramie, M.B., Pilliod, D.S., and Goldberg, C.S. 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biological Conservation 183:29–37. DOI: 10.1016/j.biocon.2014.11.025

Larson, W.A., Palti, Y., Gao, G., Warheit, K.I., and Seeb, J.E. 2018. Rapid discovery of SNPs that differentiate hatchery steelhead trout from ESA-listed natural-origin steelhead trout using a 57K SNP array. Canadian Journal of Fisheries and Aquatic Sciences 75(7):1160–1168. DOI: 10.1139/cjfas-2017-0116

Le Luyer, J., Laporte, M., Beacham, T.D., Kaukinen, K.H., Withler, R.E., Leong, J.S., Rondeau, E.B., Koop, B.F., and Bernatchez, L. 2017. Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proceedings of the National Academy of Sciences 114(49):12964–12969. DOI: 10.1073/pnas.1711229114

Liu, P., Du, Y., Meng, L., Li, X., and Liu, Y. 2016. Metabolic profiling in kidneys of Atlantic salmon infected with *Aeromonas salmonicida* based on 1H NMR. Fish & Shellfish Immunology 58:292–301. DOI: 10.1016/j.fsi.2016.08.055

Lynch, M. and O'Hely, M. 2001. Captive breeding and the genetic fitness if natural populations. Conservation Genetics 2(4):363–378. DOI: 10.1023/A:1012550620717

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, Mass. 980 pp.

MacKinlay, D.D., Lehmann, S., Bateman, J., and Cook, R. 2004. Pacific salmon hatcheries in British Columbia. American Fisheries Society Symposium 44:57–75.

Maier, T., Güell, M., and Serrano, L. 2009. Correlation of mRNA and protein in complex biological samples. FEBS Letters 583(24):3966–3973. DOI: 10.1016/j.febslet.2009.10.036

Makaras, T., Razumienė, J., Gurevičienė, V., Šakinytė, I., Stankevičiūtė, M., and Kazlauskienė, N. 2020. A new approach of stress evaluation in fish using β-D-glucose measurement in fish holding-water. Ecological Indicators 109:105829. DOI: 10.1016/j.ecolind.2019.105829

Mäkinen, H., Vasemägi, A., McGinnity, P., Cross, T.F., and Primmer, C.R. 2015. Population genomic analyses of early-phase Atlantic salmon (*Salmo salar*) domestication/captive breeding. Evolutionary Applications 8(1):93–107. DOI: 10.1111/eva.12230

Matter, A.N., Falke, J.A., López, J.A., and Savereide, J.W. 2018. A rapid-assessment method to estimate the distribution of juvenile Chinook Salmon in tributary habitats using eDNA and occupancy estimation. North American Journal of Fisheries Management 38(1):223–236. DOI: 10.1002/nafm.10014

Maynard, D.J., Flagg, T.A., Iwamoto, R., and Mahnken, C.V.W. 2004. A review of recent studies investigation seminatural rearing strategies as a tool for Increasing Pacific salmon postrelease survival. In M.J. Nickum, P.M. Mazik, J.G. Nickum, Don.D. MacKinlay (eds.) Propagated fish in resource management. American Fisheries Society symposium 44. American Fisheries Society, Bethesda, Maryland.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Massachusetts.

McCormick, S.D., Hansen, L.P., Quinn, T.P., and Saunders, R.L. 1998. Movement, migration, and smolting of Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences 55(S1):77–92. DOI: 10.1139/d98-011

McKinney, G.J., Pascal, C.E., Templin, W.D., Gilk-Baumer, S.E., Dann, T.H., Seeb, L.W., and Seeb, J.E. 2019. Dense SNP panels resolve closely related Chinook salmon populations. Canadian Journal of Fisheries and Aquatic Sciences :cjfas-2019-0067. DOI: 10.1139/cjfas-2019-0067

McMahon, B.J., Teeling, E.C., and Höglund, J. 2014. How and why should we implement genomics into conservation? Evolutionary Applications 7(9):999–1007. DOI: 10.1111/eva.12193

Micheletti, S.J., Hess, J.E., Zendt, J.S., and Narum, S.R. 2018. Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous steelhead. BMC Evolutionary Biology 18(1):140. DOI: 10.1186/s12862-018-1255-5

Middleton, M.A., Larsen, D.A., Dickey, J.T., and Swanson, P. 2019. Evaluation of endocrine and transcriptomic markers of male maturation in winter-run steelhead trout (*Oncorhynchus mykiss*). General and Comparative Endocrinology 281:30–40. DOI: 10.1016/j.ygcen.2019.05.010

Miller, K.M., Gardner, I.A., Vanderstichel, R., Burnley, T., Schulze, A.D., Li, S., Tabata, A., Kaukinen, K.H., Ming, T.J., and Ginther, N. 2016. Report on the performance evaluation of the Fluidigm BioMark platform for high-throughput microbe monitoring in salmon. DFO Canadian Science Advisory Secretariat 2016/038.

Miller, K.M., Günther, O.P., Li, S., Kaukinen, K.H., and Ming, T.J. 2017. Molecular indices of viral disease development in wild migrating salmon. Conservation Physiology 5(1). DOI: 10.1093/conphys/cox036. (https://academic.oup.com/conphys/article-lookup/doi/10.1093/conphys/cox036)

Morán, P. and Pérez-Figueroa, A. 2011. Methylation changes associated with early maturation stages in the Atlantic salmon. BMC Genetics 12(1):86. DOI: 10.1186/1471-2156-12-86

Mousseau, T. and Fox, T.A. 1998. The adaptive significance of maternal effects. Trends in Ecology & Evolution 13(10):403–407. DOI: 10.1016/S0169-5347(98)01472-4

Nakagawa, S. and Schielzeth, H. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4(2):133–142. DOI: 10.1111/j.2041-210x.2012.00261.x

Narum, S.R., Di Genova, A., Micheletti, S.J., and Maass, A. 2018. Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon. Proceedings of the Royal Society B: Biological Sciences 285(1883):20180935. DOI: 10.1098/rspb.2018.0935

Näslund, J. and Johnsson, J.I. 2016. Environmental enrichment for fish in captive environments: Effects of physical structures and substrates. Fish and Fisheries 17(1):1–30. DOI: 10.1111/faf.12088

Neff, B.D., Garner, S.R., and Pitcher, T.E. 2011. Conservation and enhancement of wild fish populations: Preserving genetic quality versus genetic diversity. Canadian Journal of Fisheries and Aquatic Sciences 68(6):1139–1154. DOI: 10.1139/f2011-029

Neff, B.D. and Pitcher, T.E. 2005. Genetic quality and sexual selection: An integrated framework for good genes and compatible genes. Molecular Ecology 14(1):19–38. DOI: 10.1111/j.1365-294X.2004.02395.x

Piccolo, J.J. 2016. Conservation genomics: Coming to a salmonid near you. Journal of Fish Biology 89(6):2735–2740. DOI: 10.1111/jfb.13172

Pitcher, T.E. and Neff, B.D. 2007. Genetic quality and offspring performance in Chinook salmon: Implications for supportive breeding. Conservation Genetics 8(3):607–616. DOI: 10.1007/s10592-006-9204-z

Prince, D.J., O'Rourke, S.M., Thompson, T.Q., Ali, O.A., Lyman, H.S., Saglam, I.K., Hotaling, T.J., Spidle, A.P., and Miller, M.R. 2017. The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. Science Advances 3(8):e1603198. DOI: 10.1126/sciadv.1603198

Puurtinen, M., Ketola, T., and Kotiaho, J.S. 2009. The good genes and compatible genes benefits of mate choice. The American Naturalist 174(5):741–752. DOI: 10.1086/606024

Räsänen, K. and Kruuk, L.E.B. 2007. Maternal effects and evolution at ecological time-scales. Functional Ecology 21(3):408–421. DOI: 10.1111/j.1365-2435.2007.01246.x

Rime, H., Guitton, N., Pineau, C., Bonnet, E., Bobe, J., and Jalabert, B. 2004. Post-ovulatory ageing and egg quality: A proteomic analysis of rainbow trout coelomic fluid. Reproductive Biology and Endocrinology 2(1):26. DOI: 10.1186/1477-7827-2-26

Rodrigues, P.M., Silva, T.S., Dias, J., and Jessen, F. 2012. PROTEOMICS in aquaculture: Applications and trends. Journal of Proteomics 75(14):4325–4345. DOI: 10.1016/j.jprot.2012.03.042

Roff, D.A. and Emerson, K. 2006. Epistasis and dominance: Evidence for differental effects in life-history versus morphological traits. Evolution 60(10):1981–1990. DOI: 10.1554/06-165.1

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. 2011. Global quantification of mammalian gene expression control. Nature 473(7347):337–342. DOI: 10.1038/nature10098

Seymour, M. 2019. Rapid progression and future of environmental DNA research. Communications Biology 2(1):80. DOI: 10.1038/s42003-019-0330-9

Shelton, A.O., Kelly, R.P., O'Donnell, J.L., Park, L., Schwenke, P., Greene, C., Henderson, R.A., and Beamer, E.M. 2019. Environmental DNA provides quantitative estimates of a threatened salmon species. Biological Conservation 237:383–391. DOI: 10.1016/j.biocon.2019.07.003

Shrimpton, J.M., Bernier, N.J., Iwama, G.K., and Randall, D.J. 1994. Differences in measurements of smolt development between wild and hatchery-reared juvenile coho salmon (*Oncorhynchus kisutch*) before and after saltwater exposure. Canadian Journal of Fisheries and Aquatic Sciences 51(10):2170–2178. DOI: 10.1139/f94-218

Solanky, K., Burton, I., MacKinnon, S., Walter, J., and Dacanay, A. 2005. Metabolic changes in Atlantic salmon exposed to *Aeromonas salmonicida* detected by 1H-nuclear magnetic resonance spectroscopy of plasma. Diseases of Aquatic Organisms 65:107–114. DOI: 10.3354/dao065107

Steele, C.A., Anderson, E.C., Ackerman, M.W., Hess, M.A., Campbell, N.R., Narum, S.R., and Campbell, M.R. 2013. A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. Canadian Journal of Fisheries and Aquatic Sciences 70(7):1046–1054. DOI: 10.1139/cjfas-2012-0451

Steele, C.A., Hess, M., Narum, S., and Campbell, M. 2019. Parentage-based tagging: Reviewing the implementation of a new tool for an old problem. Fisheries 44(9):412–422. DOI: 10.1002/fsh.10260

Stevenson, C.F., Bass, A.L., Furey, N.B., Miller, K.M., Li, S., Rechisky, E.L., Porter, A.D., Welch, D.W., and Hinch, S.G. 2019. Infectious agents and gene expression differ between sockeye salmon (*Oncorhynchus nerka*) smolt age classes but do not predict migration survival. Canadian Journal of Fisheries and Aquatic Sciences :1–12. DOI: 10.1139/cjfas-2019-0113

Stich, D.S., Zydlewski, G.B., Kocik, J.F., and Zydlewski, J.D. 2015. Linking behavior, physiology, and survival of Atlantic salmon smolts during estuary migration. Marine and Coastal Fisheries 7(1):68–86. DOI: 10.1080/19425120.2015.1007185

Stich, D.S., Zydlewski, G.B., and Zydlewski, J.D. 2016. Physiological preparedness and performance of Atlantic salmon *Salmo salar* smolts in relation to behavioural salinity preferences and thresholds. Journal of Fish Biology 88(2):595–617. DOI: 10.1111/jfb.12853

Tal, O., Kisdi, E., and Jablonka, E. 2010. Epigenetic contribution to covariance between relatives. Genetics 184(4):1037–1050. DOI: 10.1534/genetics.109.112466

Tave, D. and Hutson, A.M. 2019. Is good fish culture management harming recovery efforts in aquaculture assisted fisheries? North American Journal of Aquaculture 81(4):333–339. DOI: 10.1002/naaq.10107

Thakur, K.K., Vanderstichel, R., Li, S., Laurin, E., Tucker, S., Neville, C., Tabata, A., and Miller, K.M. 2018. A comparison of infectious agents between hatchery-enhanced and wild out-migrating juvenile Chinook salmon (*Oncorhynchus tshawytscha*) from Cowichan River, British Columbia. FACETS 3(1):695–721. DOI: 10.1139/facets-2017-0113

Tillotson, M.D., Kelly, R.P., Duda, J.J., Hoy, M., Kralj, J., and Quinn, T.P. 2018. Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. Biological Conservation 220:1–11. DOI: 10.1016/j.biocon.2018.01.030

Tomanek, L. 2011. Environmental proteomics: Changes in the proteome of marine organisms in response to environmental stress, pollutants, infection, symbiosis, and development. Annual Review of Marine Science 3(1):373–399. DOI: 10.1146/annurev-marine-120709-142729

Vøllestad, L.A., Peterson, J., and Quinn, T.P. 2004. Effects of freshwater and marine growth rates on early maturity in male coho and Chinook salmon. Transactions of the American Fisheries Society 133(3):495–503. DOI: 10.1577/T03-033.1

Wagner, E., Arndt, R., and Hilton, B. 2002. Physiological stress responses, egg survival and sperm motility for rainbow trout broodstock anesthetized with clove oil, tricaine methanesulfonate or carbon dioxide. Aquaculture 211(1–4):353–366. DOI: 10.1016/S0044-8486(01)00878-X

Wang, J. 2016. Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? Theoretical Population Biology 107:4–13. DOI: 10.1016/j.tpb.2015.08.006

Waples, R.S. 1999. Dispelling some myths about hatcheries. Fisheries 24(2):12–21. DOI: 10.1577/1548-8446(1999)024<0012:DSMAH>2.0.CO;2

Waples, R.S., Naish, K.A., and Primmer, C.R. 2020. Conservation and management of salmon in the age of genomics. Annual Review of Animal Biosciences 8(1):annurev-animal-021419-083617. DOI: 10.1146/annurev-animal-021419-083617

Waters, C.D., Hard, J.J., Brieuc, M.S.O., Fast, D.E., Warheit, K.I., Waples, R.S., Knudsen, C.M., Bosch, W.J., and Naish, K.A. 2015. Effectiveness of managed gene flow in reducing genetic divergence associated with captive breeding. Evolutionary Applications 8(10):956–971. DOI: 10.1111/eva.12331

Waters, C.D., Hard, J.J., Brieuc, M.S.O., Fast, D.E., Warheit, K.I., Waples, R.S., Knudsen, C.M., Bosch, W.J., and Naish, K.A. 2016. What can genomics tell us about the success of enhancement programs in anadromous Chinook salmon? A comparative analysis across four generations. preprint. Evolutionary Biology. DOI: 10.1101/087973. (http://biorxiv.org/lookup/doi/10.1101/087973)

Weber, E.D. and Fausch, K.D. 2003. Interactions between hatchery and wild salmonids in streams: Differences in biology and evidence for competition. Canadian Journal of Fisheries and Aquatic Sciences 60(8):1018–1036. DOI: 10.1139/f03-087

Wellband, K.W., Heath, J.W., and Heath, D.D. 2018. Environmental and genetic determinants of transcriptional plasticity in Chinook salmon. Heredity 120(1):38–50. DOI: 10.1038/s41437-017-0009-2

Wilson, A.J. and Réale, D. 2006. Ontogeny of additive and maternal genetic effects: Lessons from domestic mammals. The American Naturalist 167(1):E23–E38. DOI: 10.1086/498138

Xu, D., Li, R., Hu, C., Sun, P., Jiao, N., and Warren, A. 2017. Microbial eukaryote diversity and activity in the water column of the South China Sea based on DNA and RNA high throughput sequencing. Frontiers in Microbiology 8:1121. DOI: 10.3389/fmicb.2017.01121

Young, T. and Alfaro, A.C. 2018. Metabolomic strategies for aquaculture research: A primer. Reviews in Aquaculture 10(1):26–56. DOI: 10.1111/raq.12146

Zaiko, A., Pochon, X., Garcia-Vazquez, E., Olenin, S., and Wood, S.A. 2018. Advantages and limitations of environmental DNA/RNA tools for marine biosecurity: Management and surveillance of non-indigenous species. Frontiers in Marine Science 5:322. DOI: 10.3389/fmars.2018.00322

Zhou, X., Ding, Y., and Wang, Y. 2012. Proteomics: Present and future in fish, shellfish and seafood. Reviews in Aquaculture 4(1):11–20. DOI: 10.1111/j.1753-5131.2012.01058.x

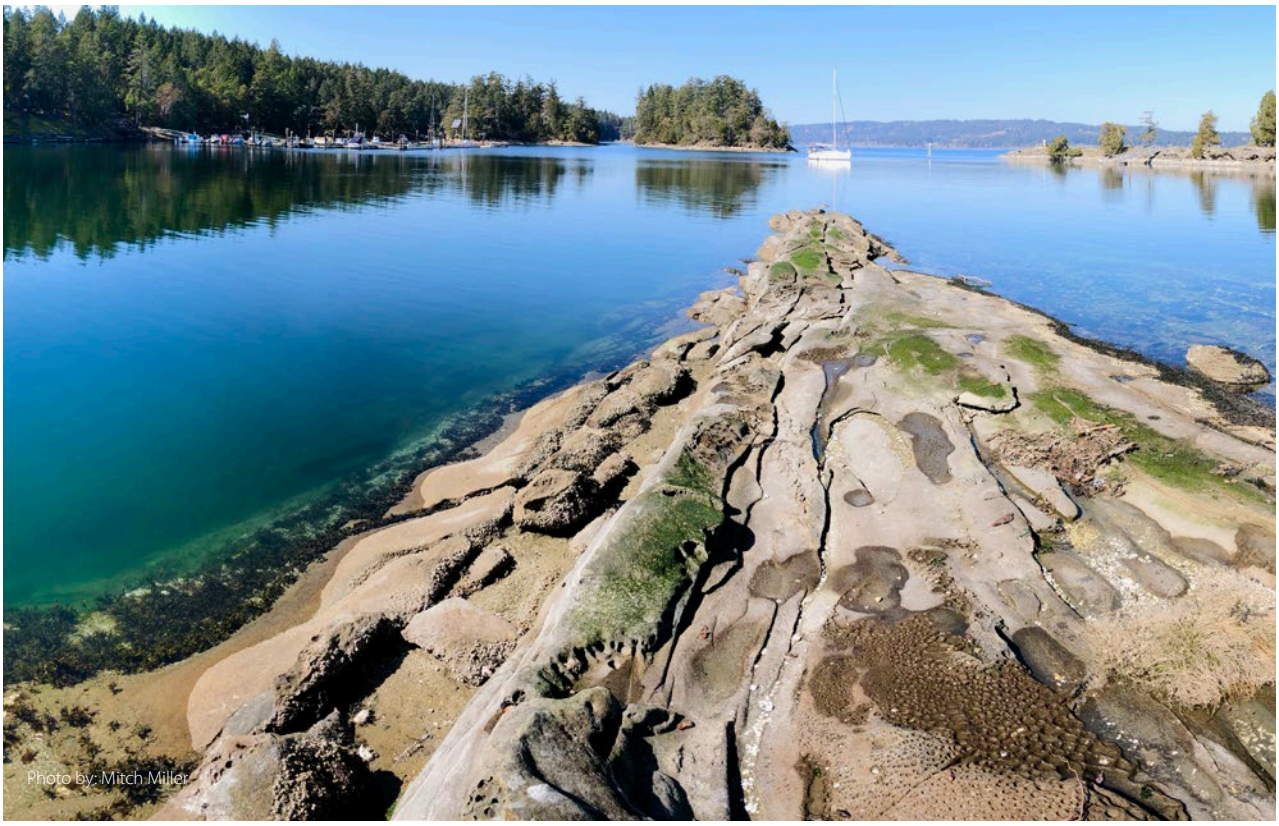
Photo by: Mitch Miller

10. GLOSSARY

Additive genetic variance. Also known as heritability. Genes that have an additive effect on a phenotypic trait. Has a positive relationship to evolutionary potential. If additive is greater than non-additive genetic variance, this may indicate 'good genes'. See non-additive genetic variance and trait architecture.

Alignment. Arranging DNA, RNA, or protein sequences to identify regions of similarity.

Alleles. Versions of the same gene with small differences in the underlying DNA sequence. These small differences contribute to the genotype.

Amino acid. Structural unit of proteins. They join to form short polymer chains called peptides (e.g. < 10 amino acids) or longer chains called either polypeptides (e.g. 10 amino acids) or proteins (e.g. 50 amino acids).

Antagonistic selection. Rare alleles that are detrimental one environment (e.g. natural environment), but beneficial in another environment (e.g. hatchery environment).

Augmentation hatchery. Increase salmon abundance for commercial and recreational fishing opportunities. Population status is stable. Reproduction in the natural environment is not a goal *per se*. See supplementation hatchery and captive breeding hatchery.

Biomarker [and robust biomarker]. A measurable substance (e.g. gene, gene expression, protein, or metabolite) that indicates a biological state. For example, stress response or disease state. Is designated as 'robust' if the biomarker works across species, life stages, environmental conditions, etc. See biomarker [discovery and validation].

Biomarker [discovery and validation]. Used to produce robust biomarkers. The discovery approach is a study to identify candidate biomarkers, e.g. RNA-seq. An extended discovery approach may include data mining several studies. The validation approach is a test of the candidate biomarkers on new samples, e.g. qPCR. An extended validation approach may include a wide range of new samples, e.g. species, life histories, sexes, etc.

Captive breeding hatchery. Salmon are unable to survive in the natural environment for at least a portion of their life cycle. Prevent imminent extinction; population status is endangered. Reproduction in the natural environment is a goal. See augmentation hatchery and supplementation hatchery.

'**Compatible gene**'. A gene with alleles that increases offspring fitness dependent on the genome, i.e. a specific genotype, representative of non-additive genetic variance. For example, males with compatible genes will produce offspring with either higher or lower average fitness dependent on the female pairing; thus, mate choice of each female should be for a different male, such that the compatible genes are in the direction of higher offspring fitness (i.e. female-male pairing). See 'good gene'.

Deoxyribonucleic acid (DNA). Hereditary material. A molecule containing a sequence (code) of four bases: adenine (A), guanine (G), cytosine (C) and thymine (T), which form base pairs as A-T and C-G. See gene.

Differentially expressed gene (DEG). Used in transcriptomics. Statically different gene expression pattern between two groups, e.g. hatchery vs. wild. Linking DEG to gene functions can reveal differences in key biological processes.

Differentially expressed metabolite (DEM). Used in metabolomics. Statically different metabolite expression pattern between two groups, e.g. hatchery vs. wild. Linking DEM to metabolite functions can reveal differences in key biological processes.

Differentially expressed protein (DEP). Used in proteomics. Statically different protein expression pattern between two groups, e.g. hatchery vs. wild. Linking DEP to protein functions can reveal differences in key biological processes.

Differentially methylated region (DMR). Used in genomics, specifically epigenetics and epigenetics programming. Statically different DNA methylation pattern between two groups, e.g. hatchery vs. wild. Linking the DMR to gene functions can reveal differences in key biological processes.

DNA methylation. Has a role in epigenetics and epigenetic programming. Methyl groups are added to the underlying DNA sequence of a gene to influence gene expression. Decreased DNA methylation (hypomethylation) activates gene expression. Increased DNA methylation (hypermethylation) silences gene expression.

Domestication selection. Any change in the selection regime of a cultured population relative to that experienced by the natural population.

Effective population size. The size of an ideal population, meeting genetic assumptions, typically smaller than the size of the census population. For example, the census population size of salmon is 100, but may include breeders and non-breeders, such that the effective populations size is the number of breeders.

Environmental DNA (eDNA). Animal release of DNA into their environment. Used for rare species detection and rough estimates of population size. Water samples are filtered, the eDNA is extracted from the filter, and subjected to a species-specific reaction.

Environmental variance. Non-genetic effect explaining a phenotypic trait. Effect is experienced by multiple individuals in a population. For example, low food vs. high food environment on growth rate. See trait architecture.

Environmental gene expression. The mRNA or transcript component of eRNA representing gene expression. Theoretical and requires further research to demonstrate feasibility. Conceivably, could be used for examining group of salmon gene expression using water samples.

Environmental RNA (eRNA). Animal release of RNA into their environment. RNA represents the living component, whereas DNA represents both the living and dead component of the community. There is the perception that RNA is much less persistent in the environment than DNA. Theoretical and requires further research to demonstrate feasibility. Conceivably, could be used for live/dead ratio, live rare species detection, rough estimates of live population size, and environmental gene expression.

Epigenetics. The study of changes in gene expression that do not involve changes in gene or the underlying DNA sequence. Involves chemical tags on genes (e.g. DNA methylation) that activate or silence gene expression.

Epigenetic programming. Epigenetic coding, i.e. chemical tags on genes or the underlying DNA, resulting from environmental factors, e.g. stress, nutrition, hormones, etc., typically during early life stages, e.g. juvenile. This programming may be permanent (i.e. influencing adult traits), reversible (i.e. disappears after some time in a common environment), or transgenerational (i.e. passed to the next generation).

Enzyme. A protein that accelerates chemical reactions in the body. Metabolic pathways depend on enzymes to conduct individual steps.

Evolutionary potential. The ability to evolve in response to selection, e.g. fishing or climate change. Has a positive relationship to the amount of additive genetic variance.

Fitness. Individual survival and reproductive success. Specifically, the ability to survive to reproductive age, find a mate, and produce offspring. The more offspring produced, the greater the fitness.

Fluidigm BioMarkTM platform. A transcriptomics technology using qPCR. A high-throughput platform to examine 96 gene expression biomarkers by 96 samples at once. See 'Salmon Fit-Chip'.

Full factorial breeding design. Females and males are mated in all possible combinations with the offspring reared in a common environment. This design is used for measures of trait architecture, e.g. additive genetic variance, non-additive genetic variance, and maternal environmental variance.

Gene. The basic functional unit of heredity. Carries the genetic instructions for development and function. A DNA sequence that typically codes for a protein using mRNA or a transcript as an intermediate step. Each gene typically has two alleles, one inherited from each parent.

Gene expression. Typically associated with the process of transcription, i.e. the copying of a gene or its underlying DNA sequence, to produce messenger RNA (mRNA) or a transcript. Examined in the terms of an increase or a decrease in amount of a specific mRNA(s) or transcript(s) because of a condition, e.g. environmental stressor.

Genetic correlation. Genetic association between two traits. May occur because a single gene influences several traits (i.e. pleiotrophy) or a group of alleles that tend to be inherited together (i.e. linkage disequilibrium). Positive correlations facilitate breeding selection, whereas negative correlations impede breeding selection. That is, breeding selection for trait 1 could lead to a favourable or unfavorable response for trait 2 based on the direction of the genetic correlation.

Genetic differentiation. A measure of the difference in allele frequencies between two populations. Related to genetic divergence between populations. May occur between separate populations (e.g. hatchery vs. wild) because of genetic drift or domestication selection. May be reduced with interbreeding between populations.

Genetic diversity. The total number of genetic characteristics in the genetic make-up of a population. The maintenance of genetic diversity is important, as its loss is associated with lower evolutionary potential. Large populations are more likely to maintain genetic diversity, whereas small populations are more likely to lose genetic diversity because of genetic drift.

Genetic drift. A change in the allele frequency over time because of random sampling every generation. For example, a rare allele may be lost in a small population because of chance.

Genetic introgression. Movement of a gene from one population (e.g. hatchery) into the gene pool of another population (e.g. wild) by interbreeding. Low genetic introgression may be considered valuable to hatcheries because it indicates lower non-local (or stray) salmon genetic influences on natural populations.

Genetics. A branch of biology concerned with the study of genes and heredity. The difference between genetics and genomics can be loosely defined by the number of markers: genetics (10s to 100s of markers) and genomics (> 1000s of markers).

Genome. An organism's complete set of DNA sequences, including genes.

Genomics. Aims to study the genome, involving the sequencing of genes or DNA. Tissue DNA is extracted using protease (breaks down proteins) then alcohol precipitation. Generally, uses next-generation sequencing (NGS) to identify single-nucleotide polymorphisms (SNPs). The difference between genetics and genomics can be loosely defined by the number of markers: genetics (10s to 100s of markers) and genomics (> 1000s of markers).

Genotype. The combination of two alleles for a gene. This combination influences the phenotype.

Genotype variance. Genetic effect explaining a trait. See trait architecture.

Genotype × Environment interaction (G × E) variance. Genetic by environment interaction explaining a trait. The fitness of a gene or genotype in more than one environment. A familiar example of a G × E interaction in salmon is local adaption, i.e. local populations have higher fitness in their home (local) than away (non-local) environment.

'**Good gene**'. A gene with an allele that increases offspring fitness independent of the genome, representative of additive genetic variance. For example, males with a good gene will produce offspring with higher fitness than males without the good gene; thus, the mate choice of all females should be for the male with the good gene (i.e. male quality). See 'compatible gene'.

Hatchery deliverables. Improvements in hatchery performance and hatchery salmon health and condition.

Heritability. See additive genetic variance.

Inbreeding coefficient (F). Probability that two alleles are identical by descent if offspring are produced between parents. For example, the expected inbreeding coefficient for the offspring produced using brother/sister mating is 25%, half-brother/half-sister mating 12.5%, and first cousin mating is 6.25%, which is related to the predicted decrease in offspring fitness be cause of inbreeding depression. See pedigree.

Inbreeding depression. Reduced fitness because of inbreeding or the breeding of related individuals. Offspring display two recessive alleles associated with a deleterious trait. The more genetically similar the parents, the more likely offspring may possess deleterious traits.

Key biological process. Processes that are vital for an organism to live, made up of many chemical reactions in the body. For example, osmoregulation and metabolism are key biological processes for salmon.

Mass spectrometry (MS). A proteomics and metabolomics technique. Measurements of mass-to-charge ratios. For proteomics, ion mobility, i.e. separation based on charge, may be used to increase the coverage of proteins. For metabolomics, used with chromatography (i.e. dissolved in liquid then mixture separated on paper) or electrophoresis (i.e. separation based on charge using an electric field). High selectivity for metabolites such that low quantities are detected.

Maternal environmental variance. Non-genetic maternal investments explaining a phenotypic trait. For example, egg resources such as fats and proteins. See trait architecture.

Messenger RNA (mRNA). Focus of transcriptomics. Also known as a transcript. Created during the process of transcription, to produce a mirror copy of a gene or DNA, as an intermediate step to producing a protein. See gene expression and RNA.

Metabolite. Focus of metabolomics. A product of metabolism, e.g. sugars, organic acids, amino acids, vitamins, lipids, and nucleotides. Metabolites are the small molecule (< 1 kDa or kilodalton). See enzyme.

Metabolite expression. Examined in the terms of an increase or a decrease in amount of specific metabolite(s) because of a condition, e.g. environmental stressor.

Metabolome. An organism, cell, or tissue's complete set of metabolites.

Metabolomics. Aims to study the metabolome, to highlight changes in metabolite expression. Although environmental stressors can cause changes in gene expression, and protein expression, such changes are amplified for metabolite expression, implying that metabolomics is the most sensitive of the four 'omics' technologies. Metabolites are preserved quickly, e.g. liquid nitrogen, and stored at or below -80°C. Tissue metabolites are extracted using solvents, e.g. methanol with water. Non-targeted metabolomics is all metabolites and typically uses nuclear magnetic resonance (NMR) spectroscopy. Targeted metabolomics is a sub-set of metabolites and typically uses mass spectrometry (MS).

Methylation-sensitive amplified polymorphism (MSAP). An epigenetics technique that identifies genomic regions with high DNA methylation. DNA sequences are cut using methylation-sensitive restriction enzymes. A sub-set of DNA fragments are then amplified using PCR. Amplicons are sequenced using a high-throughput technology, i.e. NGS. Used in earlier studies, more recent studies use reduced-representation bisulfate sequencing (RRBS).

Microarray. A transcriptomics technology. A high-throughput platform involving a glass slide spotted with thousands of target complimentary mRNA sequences. Used in earlier studies, more recent studies use RNA-sequencing (RNA-seq).

Non-additive genetic variance. Genes that have a non-additive effect on a trait, i.e. combined dominance and epistasis effects. Dominance is an allele that increases offspring fitness when in a specific genotype at the same gene. Epistasis is an allele that increases offspring fitness dependent on the genotype at another gene. If non-additive is greater than additive genetic variance, this may indicate 'compatible genes'. See additive genetic variance and trait architecture.

Next generation sequencing (NGS). A genomics technology. Rapid and high-volume DNA sequencing. Bases are identified using fluorescent signals (Illumina or Roche) or the release of hydrogen protons (Ion Torrent).

Nuclear magnetic resonance (NMR) spectroscopy. A metabolomics technology. Radio waves excite a sample in a magnetic field to reveal nuclear magnetic resonance that is detected by radio receivers. The resonance signature can identify the metabolite. Low selectivity for metabolites such that only significant quantities are detected.

'**omics**'. A field of study in biology ending in -omics, such as genomics, transcriptomics, proteomics, or metabolomic. Omics aims at the collective characterization and quantification of pools of biological molecules that translate into the structure, function, and dynamics of an organism or organisms.

Outbreeding depression. Reduced fitness because of the breeding of genetically distant individuals. Offspring may display (1) intermediate genotypes that are less fit than either parental form and (2) the breakdown of biochemical or physiological compatibility.

Pedigree. A visualization of parent-offspring and sibling relationships. Can also be used to visualize family sizes (i.e. offspring counts). Provides expected inbreeding coefficients and expected relatedness values.

Parentage-based tagging (PBT). A genomics technique using SNP markers to assign offspring back to parents. Tissue samples are collected from all individuals forming the hatchery broodstock. Tissue sample DNA is sequenced for hundreds to thousands of SNP markers, which provides a genetic tag for each parent. Tissue samples of offspring are sequenced for its genetic tag, which is assigned back to parents using a computer program.

Phenotype. The set of observable characteristics of an individual resulting from the genotype, environment, and genotype × environment interaction variance. The genotype may be further sub divided to additive genetic and non-additive genetic variance. The environment may be further sub divided to maternal environmental variance. See trait architecture.

Polymerase chain reaction (PCR). A widely used method to rapidly make millions to billions of copies of a specific DNA sequence. Small amounts of DNA sequences are exponentially amplified using cycles of temperature changes, primers (bounding the target DNA sequence), and DNA polymerase (enzyme for copying).

Pool-seq. A genomics technology, which re-sequences the genome. DNA is sequenced using a high-throughput technology, i.e. NGS. Covers a higher proportion (> 10%) of the genome than RAD-seq because a pool of DNA from individuals is fully sequenced together. As the cost of sequencing decreases, soon whole genomes at the individual level (instead of the pool level) may be common.

Protein. Focus of proteomics. One or more long chains of amino acids (> 20–30). Composed of peptide sub-units (< 20–30 amino acids). See enzyme.

Protein expression. Examined in the terms of an increase or a decrease in amount of specific protein(s) because of a condition, e.g. environmental stressor.

Proteome. An organism, cell, or tissue's complete set of proteins.

Proteomics. Aims to study the proteome, to highlight changes in protein expression. Proteomics is more mechanistic than genomics and transcriptomics because it incorporates protein modifications and degradation. Tissue proteins are digested with protease (i.e. broken down) to peptides (i.e. sub-unit of protein). Peptides may be separated using a gel prior to mass spectrometry (MS).

Quantitative polymerase chain reaction (qPCR). A transcriptomics and environmental DNA technology. The continuous collection of a fluorescent signal by PCR to quantify DNA. For transcriptomics, RNA is first converted to DNA, and is used for examining the expression of specific genes. For environmental DNA, used for examining the amount of DNA in a water sample using a species-specific reaction.

Reduced-representation bisulfate sequencing (RRBS). An epigenetics technique that identifies genomic regions with high DNA methylation. DNA sequences are cut using methylation-sensitive restriction enzymes. DNA fragments are then bisulfate converted, i.e. turns the unmethylated cytosine to uracil. Converted fragments are amplified using PCR. Amplicons are sequenced using a high-throughput technology, i.e. NGS. Used currently, as RRBS has much higher sensitivity and resolution than methylation-sensitive amplified polymorphism (MSAP).

Relatedness (r) value. Fraction of alleles shared between individuals. For example, the expected pairwise relatedness (r) value between full siblings is 1/2, between half siblings is 1/4, and between first cousins is 1/8. See pedigree.

Restriction site-associated DNA sequencing (RAD-seq). A genomics technology, which re-sequences the genome. DNA sequences are cut using at least one restriction enzymes. DNA fragments are then sequenced using a high-throughput technology, i.e. NGS. Typically covers a small proportion of the genome (< 10%) such that it is economical for 2,000–6,000 SNP genes. See pool-seq.

Ribonucleic acid (RNA). Material essential in various biological roles for coding, decoding, regulation, and gene expression. A molecule containing a sequence (code) of four bases: adenine (A), guanine (G), cytosine (C) and uracil (U), which forms base pairs as A-U and C-G. See messenger RNA (mRNA).

RNA-sequencing (RNA-seq). A transcriptomics technology. RNA is first converted to DNA. Rapid and high-volume DNA sequencing using NGS. Can cover most of the transcriptome.

'**Salmon Fit-Chip**'. A transcriptomics technique using the Fluidigm BioMarkTM platform technology. Contains customizable gene expression biomarkers to assess the physiological health hand condition of hundreds of salmon at once. Current biomarkers include smoltification, general stress response, salinity stress, thermal stress, hypoxic stress, and viral disease development (VDD). Microbe loads can also be examined on the same platform.

Sequence. The order of bases in DNA or RNA, as well as the order of amino acids in a protein. This order provides the instructions for biological development and function.

Sequencing. The process of determining the base (DNA or RNA) or amino acid (protein) sequence using technology. For example, high-throughput sequencing using NGS.

Single-nucleotide polymorphism (SNP). A single DNA base pair that is variable. Commonly used in genomics. For example, parentage-based tagging (PBT) uses many SNP marker for parentage assignments.

Species-specific reaction. An environmental DNA technique using qPCR. Contains primers binding to a species-specific region of DNA, typically in the mitochondria (energy production) or ribosome (protein synthesis) genes.

Supplementation hatchery. Improve the status of an existing population by intentional demographic integration of hatchery and wild salmon. Population status is threatened. Reproduction in the natural environment is a goal. See augmentation hatchery and captive breeding hatchery.

Trait architecture. Measure of variance components explaining a phenotypic trait, typically the components are presented as the percent explaining a trait. Within a single (common) environment, variance components may include additive genetic variance, non-additive genetic variance, and maternal environmental variance. For more than one environment, variance components may include genotype variance, environmental variance, and genotype X environment (G X E) interaction variance. See full factorial breeding design.

Transcription. Copying of a gene or its underlying DNA sequence, to produce messenger RNA (mRNA) or a transcript. The intermediate process between the gene and the protein. See gene expression.

Transcript. See messenger RNA (mRNA). Focus of transcriptomics.

Transcriptome. An organism, cell, or tissue's complete set of mRNA or transcripts.

Transcriptomics. Aims to sequence mRNA or transcripts, to highlight changes in gene expression. Provides greater mechanistic information than genomics. mRNA is preserved quickly, e.g. RNAlater solution, liquid nitrogen, or dry ice, and stored at or below -80°C. Tissue mRNA is extracted using an acid solution, e.g. TRIzol. mRNA is typically converted to DNA because it is more stable.

Translation. Also known as protein synthesis. The code from mRNA or a transcript is used to build a specific protein. See transcription.

Photo by: Eiko Jones

PACIFIC SALMON FOUNDATION

1682 West 7th Ave,
Vancouver, BC, V6J 4S6
Tel: 604-664-7664
Email: salmon@psf.ca

BRITISH COLUMBIA

Fisheries and Oceans Canada

Pêches et Océans Canada