Genetic Stock Identification of Fraser River Pink Salmon: Methodology and Management Application

Bruce White

March, 1996

Pacific Salmon Commission Technical Report No. 7

The Pacific Salmon Commission is charged with the implementation of the Pacific Salmon Treaty, which was signed by Canada and the United States in 1985. The focus of the agreement is salmon stocks originating from rivers in one country that are subject to interception by the other country. The objectives of the Treaty are to: (1) conserve the six species of Pacific salmon in order to achieve optimum production, and (2) to provide for each country to receive benefits equivalent to the production of salmon originating in its waters.

Technical Reports of the Pacific Salmon Commission present results of completed or ongoing investigations carried out by the Pacific Salmon Commission that are deemed of sufficient interest to be made available to the scientific community and the public.

Pacific Salmon Commission
600-1155 Robson Street
Vancouver, B.C.
V6E 1B5
(604) 684-8081

Pacific Salmon Commission
Technical Report No. 7

Genetic Stock Identification of Fraser River Pink Salmon: Methodology and Management Application

Bruce White

March, 1996

Correct citation for this publication:

White, B.A. 1996. Genetic stock identification of Fraser River pink salmon: methodology and management application. Pacific Salmon Comm. Tech. Rep. No. 7: 44 p.

TABLE OF CONTENTS

ABSTRACT

Genetic stock identification (GSI) is used by the Pacific Salmon Commission (PSC) to estimate contributions of Fraser River pink salmon (*Oncorhynchus gorbuscha*) to mixed-stock fishery catches. GSI provides the basic stock composition data needed by the Fraser River Panel to achieve its mandates of accounting for Fraser River pink salmon wherever they are caught and managing fisheries targeting on these stocks within the Fraser River Panel Area. The main components of the GSI program are discussed, including: (1) collection and electrophoretic analysis of tissue samples of pink salmon from spawning grounds of major, representative stocks; (2) creation of baselines after statistical analysis of the stock-specific genetic data; (3) planning and execution of the in-season GSI program; (4) application of the stock composition estimates for fisheries management; and (5) post-season analyses of stock contribution estimates made during the in-season period. Over the five cycle-years (1987, 1989, 1991, 1993, and 1995) that the GSI program has been conducted, it has aided the Panel in achieving catch and escapement goals for Fraser River pink salmon and has provided useful information for in-season run size estimation and determination of migration routes and timing of these fish.

INTRODUCTION

The Pacific Salmon Treaty, signed in 1985 by the Governments of Canada and the United States, established the Pacific Salmon Commission (PSC) and its Panels. The Fraser River Panel is responsible for in-season regulation of Fraser River sockeye *(Oncorhynchus nerka)* and pink salmon (*Oncorhynchus gorbuscha*) fisheries in the Fraser River Panel Area. Its two primary mandates involving Fraser River pink salmon are: (a) to manage fisheries focusing on Fraser River pink salmon within the Fraser River Panel Area (Figure 1); and (b) to account for the catch of Fraser River pink salmon in all fisheries in both Canada and the United States. The Panel is responsible for meeting gross escapement targets set by Canada and for achieving agreed international and domestic catch allocations.

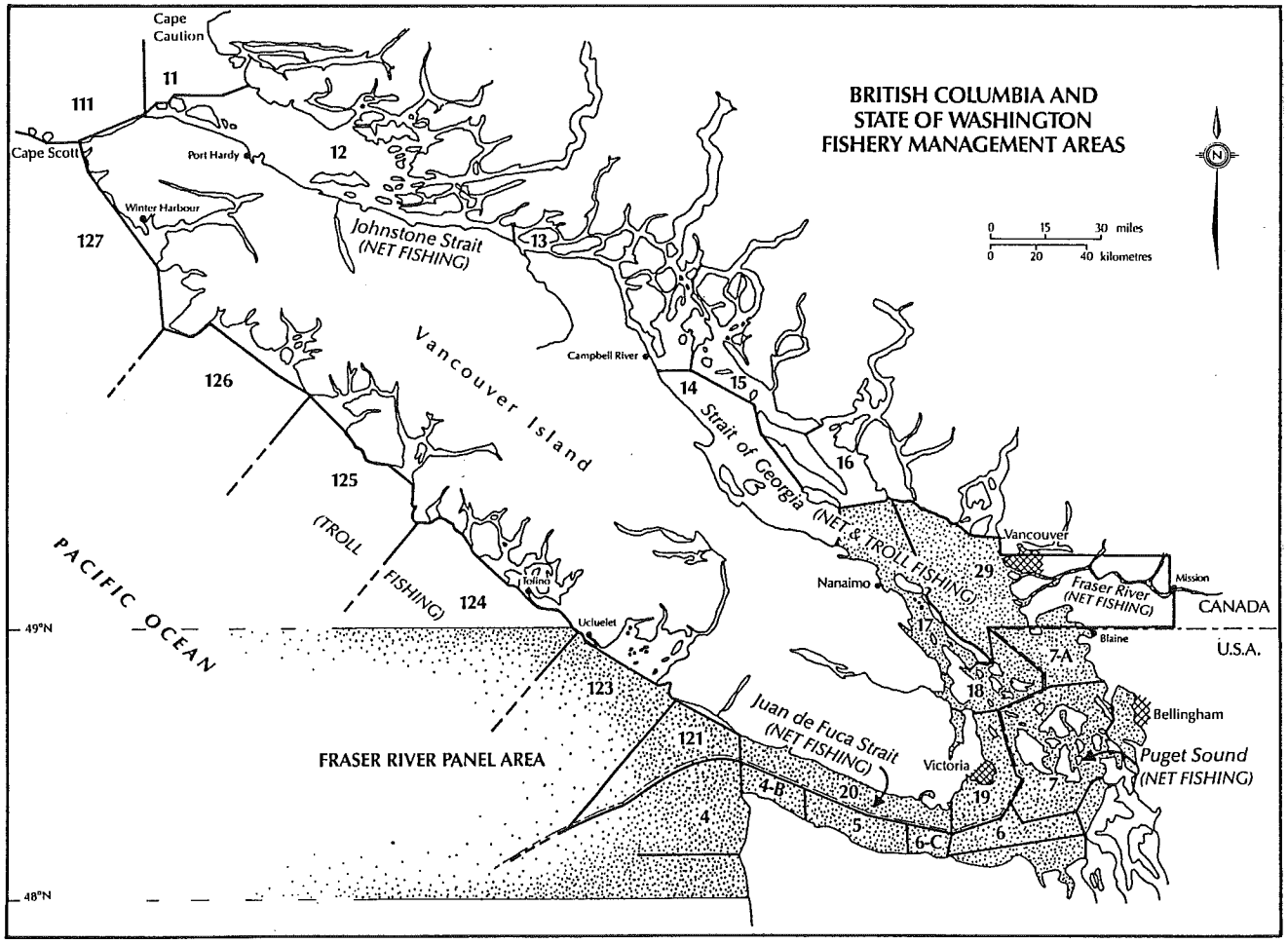


Figure 1. Fishery management areas (shaded) in the Fraser River Panel Area (PSC 1994).

The Fraser River and its tributaries produce the largest run of pink salmon in British Columbia with an average run of approximately 15 million fish over the last 10 cycles; 1977-1995 (PSC 1994 and unpublished data). Fraser River pink salmon have a rigid two-year life cycle and spawn every odd-numbered year. The major populations in the Fraser River watershed spawn in the mainstem of the Fraser River (between Chilliwack and Hope), Seton Creek, Thompson River, Harrison River, and Chilliwack-Vedder River (PSC 1990) (Figure 2). The enhanced component of the Fraser River pink salmon run is mainly comprised of four spawning channels at upper and lower Seton Creek, Weaver Creek and Jones Creek. They contribute about five percent of the overall watershed production of Fraser River pink salmon (DFO 1995).

During their marine phase, pink salmon from the Fraser River intermingle with pink salmon stocks from Alaska, Washington and other stocks from British Columbia and are harvested in Canadian and United States fisheries. Consequently, to fulfil the Fraser River Panel's responsibility for managing fisheries to achieve escapement and catch allocation objectives, the PSC developed an in-season method to identify and account for catches of Fraser River pink salmon rapidly in mixed-stock fisheries.

Prior to 1987, the Pacific Salmon Commission (and formerly the International Pacific Salmon Fisheries Commission, IPSFC) used run-reconstruction models for post-season accounting of Fraser River pink salmon catches. Some of the disadvantages of run-reconstruction are that it does not provide fisheries managers with in-season stock composition information for developing management strategies and does not account for interannual variation in migration routes. Other methods of stock identification such as scale pattern analysis, morphometrics, and meristics have thus far shown limited potential for identifying Fraser River pink salmon in mixed-stock fisheries.

In the early 1980's fisheries geneticists (e.g., Milner et al. 1985) began actively developing genetic stock identification (GSI) methods for identifying Pacific salmon (*Oncorhynchus* spp.) stocks in mixed-stock fisheries. This method is based on the principle that fish stocks existing as discrete reproductive units may become genetically distinct over time (Shaklee et al. 1990a). The genetic differences between stocks can be quantified using starch-gel electrophoresis, which involves separating electrically charged molecules (e.g., enzymes) in an electric field. The genetic traits of individual fish or stocks can be measured because of the relation between the genetic code (DNA) and the enzyme biochemical phenotypes that are expressed as banding patterns on starch gels (Utter et al. 1987).

GSI has several advantages over many other methods of stock identification: the traits (allele frequencies of loci) used to differentiate the stocks are relatively stable over time, which reduces the need to re-baseline stocks every cycle; the genetic marks are inherited naturally; and the gene markers can be analyzed rapidly and in most cases at a reasonable cost (Pella and Milner 1987; Shaklee et al. 1990a).

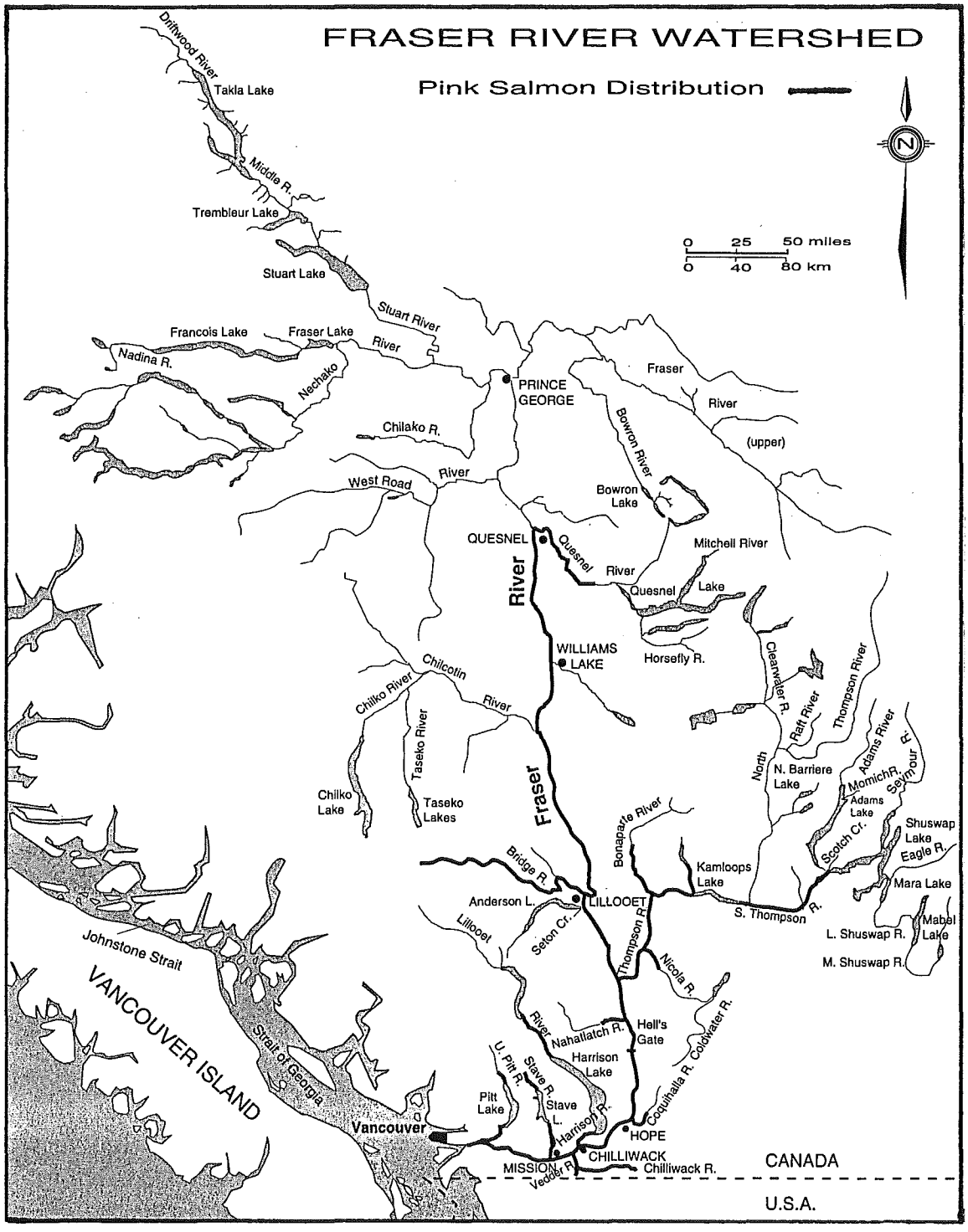


Figure 2. Main pink salmon spawning areas in the Fraser River watershed (modified from DFO 1995).

Most GSI programs involve three main steps: (1) electrophoretic analysis of tissue samples from spawning populations that may contribute to the fisheries of concern so that an accurate genetic baseline can be assembled; (2) electrophoretic analysis of tissue samples collected from fish of unknown stock-origin caught in mixed-stock fisheries; and (3) use of statistical techniques, such as the maximum likelihood estimation (MLE) model, to estimate the relative contribution of baseline stocks in mixture samples. Over the past decade, GSI has emerged as one of the primary methods of identifying Pacific salmon stocks in marine fisheries.

Beacham et al. (1985) reported that sufficient genetic differences in the allelic frequencies of Fraser River, Canadian non-Fraser, and Puget Sound pink salmon existed to allow calculation of reliable estimates of stock contributions in mixed-stock fisheries. As a result of the findings of Beacham and collaborators, the IPSFC initiated a pilot GSI program in 1985 and the PSC expanded it in 1987 to determine if the technique would be useful in achieving some of the management objectives of the Fraser River Panel. In 1987, after reviewing the results of the GSI program, the PSC introduced GSI techniques as part of its fisheries management program (replacing run reconstruction) and expanded in-season sampling to encompass northern as well as southern fisheries.

Between 1987 and 1995 the PSC, Canada Department of Fisheries and Oceans (CDFO), and the Washington Department of Fish and Wildlife (WDFW) collected numerous baseline samples from pink salmon on spawning grounds in British Columbia and Washington. The Genetics Unit of WDFW conducted the electrophoretic screening of the baseline samples. This research resulted in the identification of additional polymorphic loci, which increased the accuracy and precision of identifying Fraser River pink salmon in mixed-stock fisheries. The Genetics Unit has also improved techniques for laboratory resolution of many of the enzyme systems used in the Commission's GSI program.

The purpose of this report is to provide an overview of the major components of the Commission's pink salmon GSI program. The objectives are to describe the key aspects of the GSI program (Figure 3): (1) field collection and laboratory analysis of baseline samples; (2) statistical analysis and development of the baselines for fisheries analysis; (3) planning in-season GSI programs; (4) implementation of in-season GSI programs, including laboratory and statistical analysis of GSI samples; (5) the use of GSI stock composition estimates to help manage fisheries; and (6) post-season reviews of the in-season GSI estimates. Methods requiring further research to determine their potential utility for improving management of Fraser River pink salmon are also discussed.

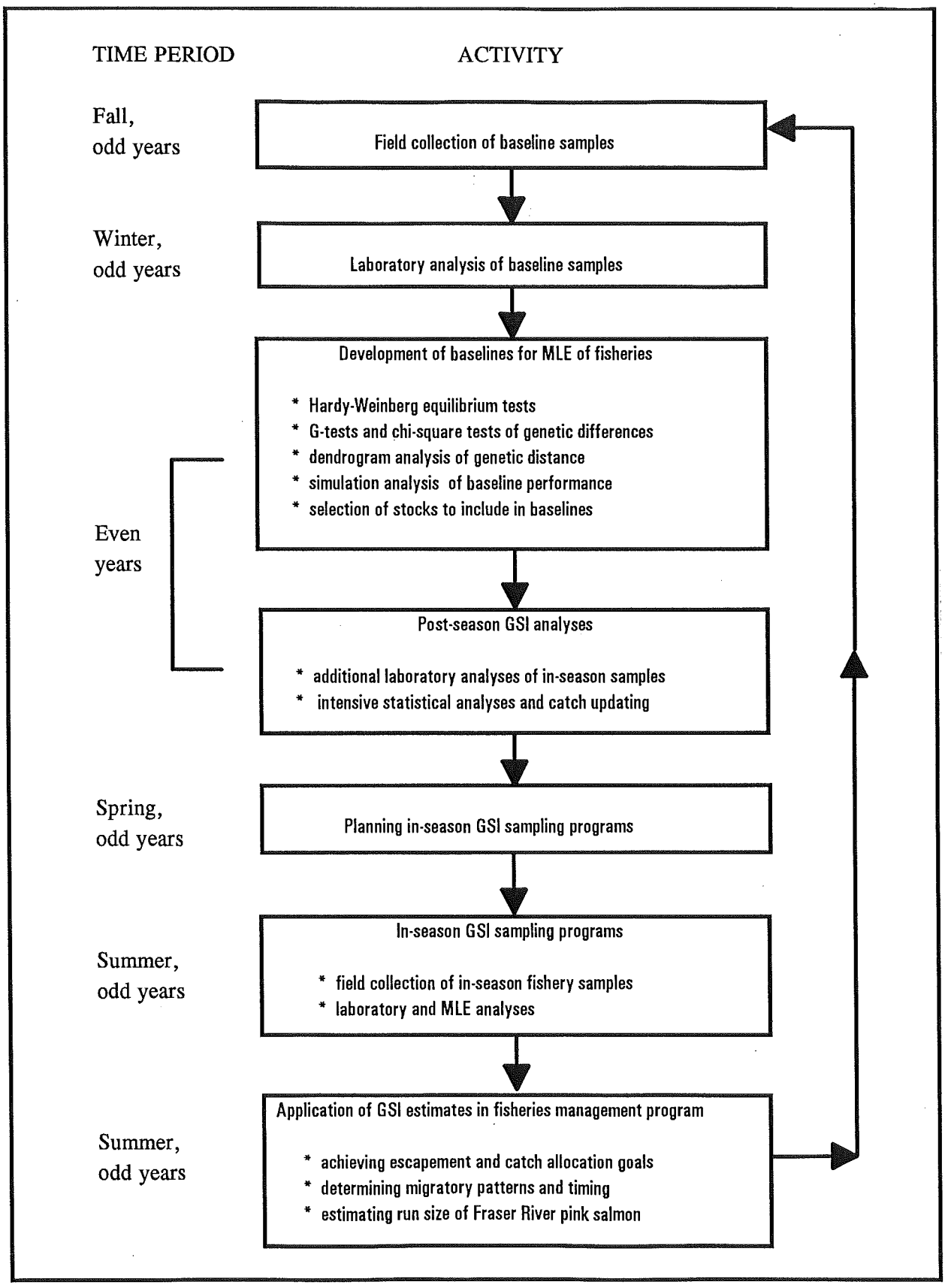


Figure 3. Chronology of activities in the PSC's pink salmon GSI program.

BASELINE FIELD SAMPLING

The collection of high quality, representative tissue samples from spawning populations of pink salmon is an important aspect of the Commission's pink salmon GSI program (Figure 3). The samples should be representative of the spawning populations they are intended to profile. The tissue samples should also be of high quality so that laboratory analyses can produce accurate genetic descriptions of each stock. Obtaining poor tissue samples from a stock and including the resulting data in a baseline can introduce error into all subsequent stock composition estimates. Unlike the laboratory and statistical analysis phases of GSI programs where analyses can be repeated, mistakes in field sampling cannot usually be corrected until the next return of pink salmon two years later.

The field component of the baselining program for pink salmon involves: (a) selecting appropriate stocks to sample; (b) sampling fish for tissues and matching biological data; and (c) freezing and transporting the tissue samples to a laboratory for analysis.

In this report the terms "collection", "stock", and "stock group" are used frequently. "Collection" refers to a group of fish (often 100-150) that have been sampled from a particular mixed-stock fishery or a river at a certain time (e.g., a collection taken in 1987 from the Harrison River in British Columbia). The term "stock" denotes a self-sustaining, inter-breeding population of fish with no (identifiable) immigration or emigration. Several collections of fish from the same river in one or more years may be used to represent the same stock (e.g., Harrison River collections taken in 1987 and 1989, combined). References to a "stock group" mean an assemblage of stocks that have been grouped based on their genetic similarity, geographic proximity, and/or because they form a natural unit for fisheries management purposes. For example, the Fraser River stock group includes pink salmon stocks from the Thompson River, Seton Creek, Bridge River, Harrison River, Vedder River, mainstem of the Fraser River, and other stocks in the Fraser River watershed.

SELECTING STOCKS TO BASELINE

The selection of stocks for baseline sampling is based on several factors.

1. Large production stocks.

The Commission's main priority is to obtain baseline data from large-production Fraser River pink salmon stocks as well as stocks from other river systems that will likely be major contributors to the mixed-stock fisheries of concern. Key stocks are selected by examining recent run size data and information on probable migration routes through marine fisheries.

2. Verifying genetic stability of the baseline.

An important assumption of GSI is that the gene frequencies are relatively constant over time, and therefore it should not be necessary to re-baseline stocks every cycle

(Shaklee and Phelps 1990). If this assumption was seriously violated, the in-season fisheries management value of the GSI estimates would be reduced because the "true" gene frequencies of each stock would not be known until post-season, i.e., after the stocks were re-baselined on the spawning grounds. The Commission often re-baselines key stocks from some of the stock groups each cycle to verify that the in-season baselines were representative of the true spawning populations and to test the assumption that the stocks are genetically stable over time.

3. Lower production stocks not previously baselined.

It is beneficial to characterize some of the moderate and lower production stocks after the large production stocks are baselined. Their inclusion in a baseline generally improves the accuracy of the stock composition estimates.

4. Re-baselining stocks where the initial samples were insufficient to provide an accurate genetic profile of the stock, or the tissue samples were of poor quality and a laboratory could not obtain adequate data for all genetic traits being screened.

Baseline collections should comprise at least 50-100 individual fish per stock (Shaklee et al. 1990a). Shaklee and Phelps (1990) recommend eliminating data for a locus in a baseline stock if less than 90% of the fish in a collection were successfully scored (genotypes identified and recorded) for that locus.

5. Stocks with other potential management applications.

Some stocks have a higher degree of genetic difference from other stocks because of their geographic isolation and evolutionary history. Those stocks may be important to include in a baseline because they could be valuable as in-season indicators of migration patterns and timing of certain stocks or stock groups. This information could be useful for modifying in-season harvest strategies.

6. Cost of obtaining the sample.

The spawning grounds of stocks that are easily accessible usually have lower sampling costs associated with them, and therefore are preferred if the other factors are equal. However, if a stock is important to include in the baseline the cost of obtaining the sample becomes less significant.

It is possible to reduce future baselining efforts substantially after sufficient baseline data have been collected. However, it is advisable to re-sample some of the large-production stocks occasionally to confirm that a baseline is genetically stable across years.

SAMPLING FISH

The fish that are sampled from spawning grounds are intended to represent the geographic and temporal characteristics of each stock. However, due to cost and other constraints, the Commission's baselining program focuses on obtaining samples from

locations where escapements are large, and from periods of peak spawning. Efforts are concentrated on constructing an accurate genetic profile of the maximum number of pink salmon that are contributing to the mixed-stock fisheries of concern.

Commission samplers usually collect approximately 100 pink salmon per stock during baselining operations. This sample size is considered adequate to provide accurate estimates of allele frequencies for all but the rarest alleles in stocks. Equipment that is commonly used in the baselining program is listed in Appendix A.

Choosing pink salmon for sampling is based on several factors.

1. The fish should be selected randomly, i.e., selecting fish based on particular phenotypic traits such as size or coloration should be avoided.

If fish are not selected randomly, it would violate the Hardy-Weinberg genetic law, which is an underlying assumption of GSI.

2. The sex ratio of the sample should be approximately 50% males and 50% females.

Although it is unlikely that sampling a disproportionate number of one sex would seriously affect the gene frequencies compiled for a stock, quite often other physical measurements taken during sampling (e.g., length) are more valuable for management applications when there are sufficient data for each sex.

3. Fish selected for baseline sampling should be fresh because their tissues yield much clearer locus-banding patterns when they are electrophoretically analyzed than those from decomposing fish.

The order of preference for sampling live fish or carcasses is as follows: (a) live spent fish; (b) gravid fish (subject to agency policies regarding killing unspawned salmon); (c) freshly dead carcasses where the gills are still red; and (d) carcasses that exhibit pink gills. Sampling tissues from carcasses where the gill colour is light pink, grey, or white is avoided because the enzymes in the tissues would likely be denatured and difficult to score in a laboratory.

The tissues most often taken during the Commission's baselining operations are muscle, eye, heart, and liver. Different tissues are taken because the banding patterns at specific loci are much clearer in some tissues than in others and thus determination of gene frequencies is more accurate. The techniques and other considerations for sampling each of the tissues are described in Appendix A.

The temperature of dry ice is approximately -76°C. It is used during field sampling to freeze all pink salmon tissue samples quickly and minimize enzyme degradation. Degraded tissues often create difficulty for laboratories trying to interpret the banding patterns of loci. Methods of field preserving tissue samples other than dry ice, such as ice packs or flake ice are not cold enough to stop protein deterioration in the tissues and are avoided. Aebersold et al. (1987) noted that preserving tissue samples at standard

freezer temperatures of -10°C to -20°C usually results in some enzyme degradation within a few weeks. Therefore, baseline tissue samples are generally stored at approximately -75°C in ultra-cold freezers prior to analysis.

LABORATORY ANALYSIS OF BASELINE SAMPLES

After baseline collections of tissue samples have been obtained in the autumn, they are electrophoretically analyzed. The resulting genetic data are used in the construction of baselines (Figure 3).

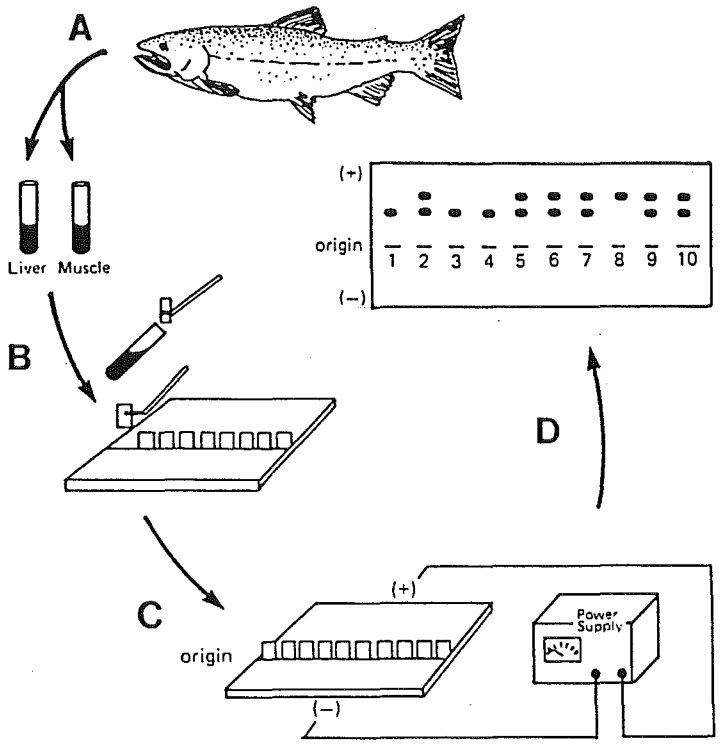
The steps employed in electrophoretic analyses have been described by many researchers (e.g., Shaklee and Keenan 1986; Aebersold et al. 1987; Utter et al. 1987; Morizot and Schmidt 1990). Figure 4 shows the steps involved in electrophoresis. Laboratories conducting electrophoretic studies have developed many precautions to minimize errors in sample processing and produce accurate genetic data (Shaklee and Phelps 1990). Some of the terms that are commonly used in electrophoresis are defined in Appendix B.

Genetic researchers (e.g., Beacham et al. 1985; Shaklee et al. 1991) have examined numerous loci in pink salmon to determine their utility for identifying pink salmon stocks in mixed-stock fisheries. Over 50 enzyme-coding loci have been reported to exhibit genetic variation in odd-year North American pink salmon (Shaklee et al. 1991). Approximately 17 of these loci have been used by the Commission for analyzing in-season collections of muscle tissue samples.

There are several reasons for not screening all known variable loci for mixed-stock fishery analysis. Loci that exhibit rare variation are excluded because they would contribute little to stock discrimination. Loci that are only scorable in liver and/or heart tissue are not included because these tissues are unavailable in samples from troll fisheries (the fish are usually eviscerated prior to sampling). Analysis of more than one tissue type increases field sampling and laboratory processing costs and also increases the time needed for laboratory analysis. Additionally, simulation analyses have indicated that many of the most informative loci for discriminating among the four major stock groups of interest to the Commission (i.e., Fraser River, Puget Sound, Canada South Coast, and Canada North Coast, see Figure 5) can be scored from skeletal muscle tissue samples. Therefore, the Commission has based its in-season fishery analysis on the 17 most variable and informative loci that are scorable from muscle tissue samples (Tables 1 and 2).

Basic Electrophoretic and Laboratory Procedures.

A. Tissue samples (e.g., muscle, heart, liver, and eye) are taken from each fish and placed in a culture tube with a small amount of water. Cellular proteins in the tissue are released into solution by freeze/thaw and mechanical agitation procedures.
B. A protein extract from each fish is individually absorbed onto a filter paper wick and placed onto the edge of a starch gel at the origin. Samples from 10 fish are shown loaded in the diagram, although typically, samples from 50 fish are loaded on one gel (i.e., with 50 wicks).
C. A direct current is applied across the gel. Protein molecules absorbed on each wick enter and move through the gel because of the molecule's net electrical charge and at a rate proportional to this charge. This charge, in turn, depends on the genetically controlled amino acid substructure of the protein molecules.
D. After about 4 hours, the gel is removed from the power source and the positions of specific proteins (usually enzymes) in the gel are identified by specific histochemical staining procedures (i.e., using general staining reagents or specific procedures involving the enzyme in the staining process). The relative migration distances of the proteins from the origin, indicated by the staining zones, are recorded as the raw data. The simplified genetic model used for interpreting electrophoretic protein variation is that one gene codes for one protein (polypeptide) chain. Therefore, electrophoretic differences between individuals in protein patterns that are based on amino acid differences are a direct reflection of genetic differences between the individuals. The simple extension of genetic differences between individuals to the evaulation of genetic differences between populations is outlined in Box B.



Steps for obtaining electrophoretic data.

Figure 4. Steps in applying electrophoresis to tissue samples (Milner et. al. 1985).

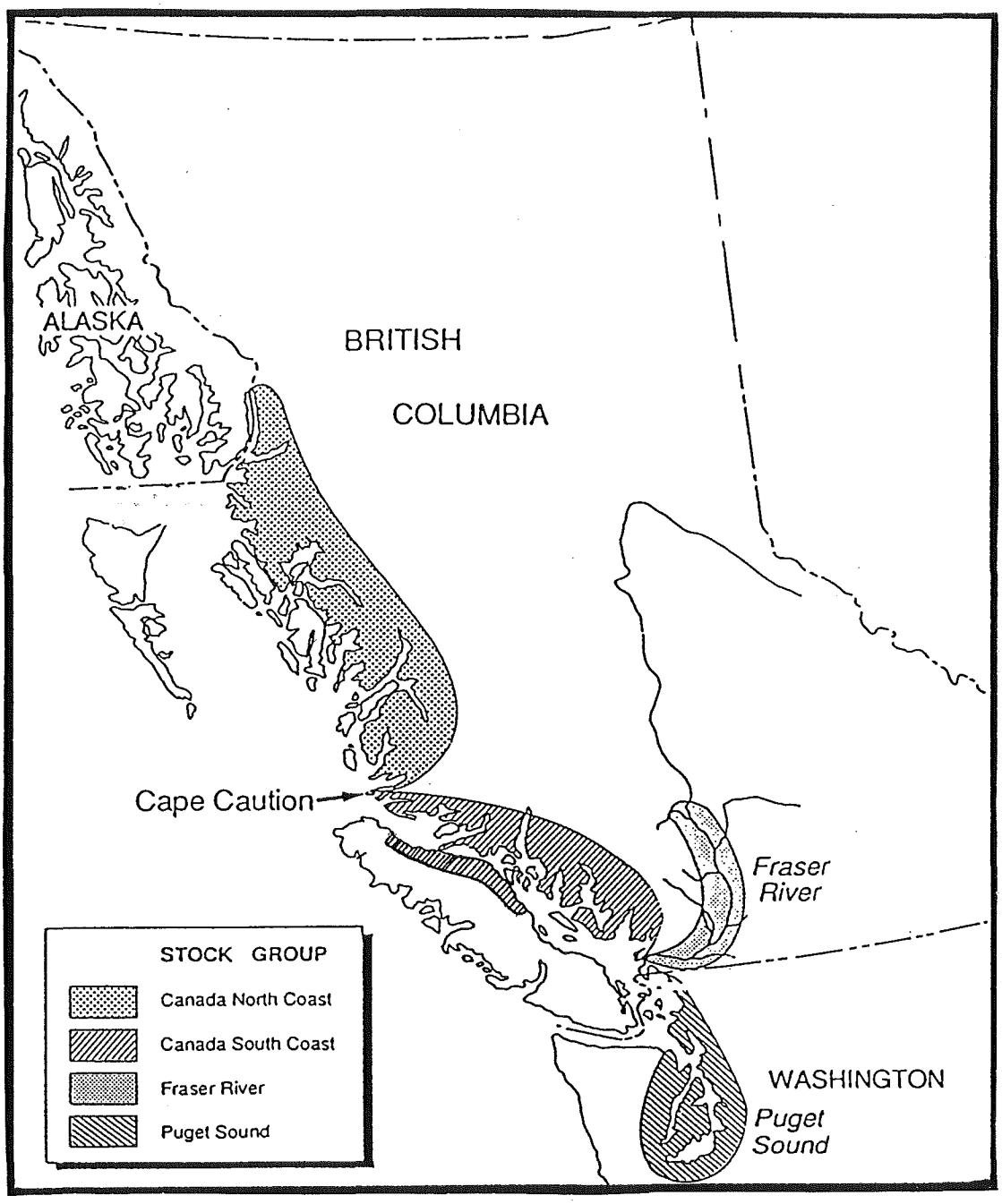


Figure 5. Location of pink salmon stock groups currently represented in the Pacific Salmon Commission's GSI baselines (after White and Shaklee 1991).

Table 1. Enzyme and locus nomenclature[a] and electrophoretic screening conditions for pink salmon GSI using muscle tissue.

LOCUS[a]	SYNONYMY[b]	ENZYME NAME (NUMBER)	SUBUNIT STRUCTURE[c]	BUFFER(S)
ADA-2	Ada (Ada-2)	adenosine deaminase (3.5.4.4)	M	CAME6.1; CAME6.8; TC-4
mAH-4		mitochondrial aconitate hydratase (4.2.1.3)	M	CAME6.8
ALAT		alanine aminotransferase (2.6.1.2)	D	TRIS-GLY; CAME6.8
CK-A1	Ck-1	creatine kinase (2.7.3.2)	D[d]	TRIS-GLY
GPI-A	Pgi-3 (Phi-3)	glucose-6-phosphate isomerase (5.3.1.9)	D	TRIS-GLY
GPI-B1,2	Pgi-1,2	glucose-6-phosphate isomerase (5.3.1.9)	D	TRIS-GLY
G3PDH-1	Agp	glycerol-3-phosphate dehydrogenase (1.1.1.8)	D	TRIS-GLY; CAME6.8
FDHG		formaldehyde dehydrogenase (glutathione) (1.2.1.1)	D	TRIS-GLY;CAME6.8
LDH-A1	Ldh-1	lactate dehydrogenase (1.1.1.27)	T	TRIS-GLY
sMDH-B1,2	Mdh-3,4	cytosolic malate dehydrogenase (1.1.1.37)	D	CAME6.3; TC-4; CAME6.8
mMEP-1	Me (MDHp-1)	mitochondrial $NADP^+$-dependent malate dehydrogenase (1.1.1.40)	T	CAME6.8
PEPD-2	Pp (PDPEP-2)	proline dipeptidase (3.4.13.9)	D	CAME6.8; TRIS-GLY
PEP-LT	Ll-1	peptidase (leucyl-tyrosine substrate) (3.4.-.-)	M	TC-4; TRIS-GLY
PGDH	6-Pg	phosphogluconate dehydrogenase (1.1.1.44)	D	CAME6.8; TC-4
PGM-2	Pgm	phosphoglucomutase (5.4.2.2)	M	TRIS-GLY CAME6.8

[a] = nomenclature according to AFS standard (Shaklee et al. 1990b).

[b] = locus symbols used in previous publications (e.g., Beacham et al. 1985).

[c] = M = monomer; D = dimer; T = tetramer.

[d] = although this enzyme is a dimer, the isozyme expressed in muscle tissue ($CK\text{-}A1_2$) exhibits monomeric patterns of variation in salmonids and other teleost fishes.

Table 2. Alleles and relative electrophoretic mobilities[a] for recognized alleles at 15 systems in pink salmon.

LOCUS	ALLELES							
ADA-2	100	110	90	105	114			
mAH-4	100	116	76	81				
ALAT	100	111	108	87	77	106		
CK-A1	100	66	110	74				
FDHG	100	138	58					
GPI-A	100	108	91	87	120			
GPI-B1,2	100	200	-64	25				
G3PDH-1	-100	-170	60	-10	20	200		
LDH-A1	-100	-250						
sMDH-B1,2	100	124	66	72	69	85	130	47
mMEP-1	100	123	115					
PEPD-2	100	120	80	110	130			
PEP-LT	100	108	90	80				
PGDH	100	108	96	86	90			
PGM-2	100	155	135	25	250			

[a] Some alleles at specific loci are pooled with the allele of closest mobility prior to MLE analysis to minimize potential laboratory scoring errors. Negative numbers associated with some alleles indicate cathodal mobility.

STATISTICAL ANALYSIS OF BASELINE DATA

Once the contract laboratory (from 1987 to 1995, WDFW Genetics Unit) has completed their electrophoretic analysis of the baseline samples, the results are tabulated into a format suitable for statistical analysis. The primary objective of the statistical analyses of these data is to provide information that enables construction of baselines that will yield accurate stock composition estimates under a wide range of potential mixtures of stocks. The following analytical tests are performed.

HARDY-WEINBERG TESTS OF BASELINE DATA

The first analysis applied to the raw genetic data obtained from each stock identifies whether the stocks satisfy Hardy-Weinberg equilibrium expectations. A population (or stock) is generally considered to remain in genetic equilibrium if the following conditions are present (May and Krueger 1990): (1) no selection against particular genetic combinations; (2) no migration between stocks; (3) random mating of individuals in the stock; (4) large stock size (i.e., > 10,000 fish with an equal sex ratio); and (5) no mutation which could change allele frequencies over a period of time.

Hardy-Weinberg equilibrium (HWE) tests are performed at each variable locus after the observed genotypes for each fish are entered into the proper data format for the BIOSYS-1 software program (Swofford and Selander 1981). The chi-square test for goodness of fit is used to determine if the observed genotype frequencies are significantly different from those expected under HWE. For each stock and locus, the program lists the probability that it conforms to HWE. If the probability (P) reported is less than 0.05, then the locus is considered to be out of HWE. The number of cases where loci significantly depart from HWE (at $P < 0.05$) is determined for each stock. The results are examined on a stock-by-stock and locus-by-locus basis to determine if there are systematic problems at certain stocks or loci. Deviations may result from non-representative field sampling, errors in laboratory gel scoring, or violation of one or more of the assumptions of HWE noted earlier. Stocks that are out of HWE at several loci should be re-sampled at a later date. If most of the departures from HWE are confined to one or a few loci, they are likely attributable to one or more of the following reasons: an incorrect model for interpreting the observed variation, inaccurate scoring of the isozyme banding patterns, or violation of one or more of the HWE assumptions (e.g., gene flow, selection, mutation). Loci that are seriously out of HWE should be excluded from the baseline, unless there is a valid reason for including them.

TESTS OF GENETIC VARIABILITY AMONG STOCKS

The genetic similarities/differences among stocks are examined after a baseline has been assembled where the stocks conform to HWE expectations. G-tests and/or chi-square tests are used to compare all possible pairs of collections for the loci of concern in the baseline, including comparisons of the same stock across several different years when collections were taken.

The null hypothesis tested is that there is no significant difference between the two collections being compared. Shaklee and Phelps (1990) suggest that as a general guideline, the G-test or chi-square test should indicate a probability of less than 1% ($P < 0.01$) that two collections are the same to retain them as separate stocks in the baseline. The criteria and significance levels applied by Commission analysts for determining how to treat collections for baseline development (i.e. retaining, pooling or omitting them) are drawn in part from Shaklee and Phelps (1990) and unpublished guidelines of the Coastwide Consortium for Pacific Salmon GSI (participants in this consortium include WDFW, PSC, CDFO, ADF&G, and the U.S. National Marine Fisheries Service laboratories in Seattle and Auke Bay).

One important feature of the Commission's GSI program is that in most of the mixed-stock fisheries sampled that are south of Cape Caution, pink salmon from the Fraser River stock group are present in much higher proportions (often 80% or more) than fish from the other stock groups. This can cause systematic underestimates of the proportion of Fraser River pink salmon in mixed-stock fisheries as a result of inherent biases in the MLE model (discussed later in this report).

Commission analysts consider misallocations between pink salmon stocks from different stock groups to be more serious errors than misallocations between stocks within the same stock group because of the adverse effect such errors could have on fisheries management applications. Therefore, specific approaches have been developed for treating GSI baseline collections of pink salmon. In situations where a decision is made to retain only one of two collections being

compared in the baseline, several factors are considered with preference given to retaining the collection that is characterized by:

a) high quality, recently collected, representative tissue samples where a large number of fish were sampled;
b) reliable electrophoretic analyses with few missing genetic data; and
c) large production size that collection would represent in the baseline (i.e. stock is probably a major contributor to the mixed-stock fisheries of most concern).

When a situation arises that it is preferable to pool two collections together the genotype frequencies of each collection may be weighted by: (1) sample size of collection; or (2) relative production size (i.e. contribution of fish to mixed-stock fisheries). The three principle situations that occur when collections are compared and the criteria and treatment prescriptions for each are described below and summarized in Table 3.

Table 3. General guidelines for treatment options[1] used by the Commission when comparing different pink salmon collections: within the same stock but collected at different times; in different stocks within the same stock group; and in different stocks that are in different stock groups.

Situations where collections are compared	Probability that two collections were drawn from the same randomly interbreeding stock; and treatment options		
	> 99% ($P < 0.01$)	95% to 99% ($0.05 > P > 0.01$)	< 95% ($P > 0.05$)
1. Same stock, but collections taken at different times	- retain one, omit the other collection	- generally pool collections - as $P = 0.05$ is approached more likely to pool	- pool both collections
2. Different stocks within same stock group	- retain both as separate collections	- generally retain both as separate collections - as $P = 0.05$ is approached more likely to retain one collection and omit other one	- generally retain one collection and omit other one
3. Different stocks in different stock groups	- retain both as separate collections	- generally retain both as separate collections only when $P < 0.01$ or 0.02 - as $P = 0.05$ is approached strongly consider retaining one collection and omitting other one	- retain one collection and omit other one

[1] The treatment options suggested here are intended only as general guidelines. The final decisions on retaining, pooling or omitting collections from a baseline depends on numerous factors, e.g. quality of genetic data representing each collection, sample sizes, stock production sizes, overall representation of collections for a stock group in a baseline, etc.

Situation 1. Testing pairs of collections from the same stock or river between years e.g., collections from the Harrison River (Fraser River stock group) taken in 1987 and 1989.

Test result (G or chi-square): $P < 0.01$ (i.e., highly improbable (< 1% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: If the result shows that there is a significant difference between two collections obtained at different times from the same stock, the collections should be examined individually (e.g., for conformance with HWE) to try to ascertain the reason. One of the collections is usually omitted from the baseline.

Test result: $0.05 > P > 0.01$ (i.e., improbable (1% to 5% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: In situations where there is marginal genetic similarity between the collections, they are usually pooled unless there is serious reason to question the data from either. However, as $P = 0.05$ is approached (i.e., higher probability they represent the same, randomly interbreeding stock) there is stronger consideration given to pooling the collections. The confidence of the analyst in the quality of the genetic data for each collection is an important element in deciding how collections should be treated. The analyst should also be careful that pooling the genotypes of two collections that have intermediate genetic distance between them does not create an "artificial stock", which does not properly characterize the stock and could lead to erroneous stock composition estimates. Examination of dendrograms and the results of simulation analyses helps to indicate whether the "pooled" stock comprised of the two collections adequately reflects the original genotypic profile of the two collections.

Test result: $P > 0.05$ (i.e., acceptable chance (> 5%) the two collections represent the same randomly interbreeding stock).

Treatment: In such cases the collections should generally be pooled.

Situation 2. Testing pairs of collections from different stocks that are within the same stock group e.g., collections from the Harrison River and Seton Creek (both are within the Fraser River stock group).

Test result: $P < 0.01$ (i.e., highly improbable (< 1% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: If the result indicates that there is a significant difference between collections from different stocks within the same stock group, they are retained separately in the baseline.

Test result: $0.05 > P > 0.01$ (i.e., improbable (1% to 5% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: If the result shows that there is intermediate difference between the two collections, then both collections are usually retained in the baseline. This is because most misallocations that might occur between the two collections during MLE would be to the other collection and because they are in the same stock group, the accuracy of the estimated total proportion of pink salmon attributable to a stock group would not usually be adversely affected. As $P = 0.05$ is approached there is greater preference to retain only one of the two collections.

Test result: $P > 0.05$ (i.e., acceptable chance (> 5%) the two collections represent the same, randomly interbreeding stock).

Treatment: As the probability that the collections represent the same, randomly interbreeding stock increases there is less rationale for retaining both in the baseline and therefore, only one collection is retained.

Situation 3. Testing pairs of collections from stocks that are in different stock groups, e.g., from the Harrison River and Skagit River (i.e. from the Fraser River and Puget Sound stock groups, respectively).

Of the three different types of comparisons discussed here, these are generally the most important tests because large misallocations that could occur between different stock groups during MLE could seriously jeopardize achieving fisheries management goals.

Test result: $P < 0.01$ (i.e., highly improbable (< 1% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: This result indicates that there is significant genetic difference between the two collections and suggests that there should be high estimation accuracy in mixed-stock fishery analyses involving them (i.e., there should be minimal bias between the two collections during MLE). Consequently, both collections are retained in the baseline.

Test result: $0.05 > P > 0.01$ (i.e., improbable (1% to 5% chance) the two collections represent the same, randomly interbreeding stock).

Treatment: One of the most difficult problems in creating a baseline is making the correct decision when there are two or more collections from different stock groups that are not significantly different at the 99% probability level, but are at the 95% level. Three options are considered. One approach is to retain both collections in the baseline, particularly as the significance level approaches 99% (i.e., $P < 0.01$ or 0.02) and the probability that the stocks are significantly different increases. This decision is more attractive when only a few of the total number of collections that may be included in the baseline have intermediate genetic distance between them. If numerous collections that have intermediate genetic difference among them are included in the baseline then the accuracy and precision of the stock contribution estimates from the MLE model would be reduced.

A second approach is to retain one of the stocks in the baseline and omit the other. This option is given stronger consideration when $P = 0.05$ is approached. The collection representing the largest production size and best genetic data is generally retained. However, it is important that doing this does not systematically bias the genetic representation of specific stock groups in the baseline. If such bias did occur, it may be necessary to adjust the resulting stock group estimates following MLE to compensate for it.

A third option is not a preferred approach because it requires excluding from the baseline all collections from different stock groups that have intermediate genetic difference among them. The main disadvantage of removing collections or stocks with this level of genetic difference between them is that more serious misallocations could occur (particularly if they were both large production stock) than if one or both of two such collections were retained in the baseline. For example, if both collections were removed from the baseline the MLE model would assign those fish to the next most similar stock in the baseline, regardless of which stock group it was in.

Test result: $P > 0.05$ (i.e., acceptable chance (> 5%) the two collections represent the same, randomly interbreeding stock).

Treatment: In cases where there is no significant genetic difference between two collections from different stock groups, one of the collections is omitted and the other retained in the baseline.

Based on the results of the tests outlined in Table 3, preliminary baselines are created that are comprised of data from various stocks. The baselines are evaluated using simulation analysis

to assess the probable accuracy and precision of stock composition estimates that might be achieved on an in-season basis.

The preceding discussion of treatment options for various between-collection comparisons indicates that numerous factors may influence decisions on which collections to retain, pool, or omit from a baseline. The initial decisions for treating collections may be modified after reviewing the results of simulation analyses.

ANALYSIS OF GENETIC RELATIONSHIPS AMONG STOCKS

Dendrograms are constructed after completion of the G-tests or chi-square tests, and the preliminary selection of collections/stocks to include in the baseline. Dendrograms depict how the genetic attributes of each stock cause them to cluster in a multi-dimensional space. The BIOSYS-1 program generates dendrograms based on different criteria of genetic similarity and distance measures. Nei's (1972) genetic distance and Cavalli-Sforza and Edwards (1967) chord distance measures are computed because they generally provide a satisfactory portrayal of genetic relationships among stocks.

Ideally the stocks cluster into distinct groups that are also convenient units for fisheries management purposes. However, as Wood (1989) pointed out, in some instances some of the stocks may not position in convenient groupings and instead form "problem clusters", where stocks from different stock groups (based on geography or management intent) have closer genetic links than stocks from the same stock group.

Dendrograms produced from preliminary baselines provide a crude, though rapid indication of how well GSI analyses will likely perform at identifying various stocks and stock groups in mixture samples. Based on the relationships among stocks that are portrayed by the dendrograms, analysts may choose to reexamine the results of particular G-tests or chi-square tests and assemble the baseline differently to reduce misallocation bias between problem clusters of stocks. The identification of problem clusters also indicates hypothetical stock composition mixtures that should be analyzed with simulation analysis. The results from simulation analyses are useful in assessing the probable accuracy and precision of MLE estimates when mixed-stock fisheries occur with similar stock compositions.

CREATION OF BASELINES FOR IN-SEASON MANAGEMENT

After field sampling, laboratory analyses and baseline statistical analyses are completed, the stocks that will form a baseline for in-season analyses are selected. This process involves examination of four main factors: (1) the genetic distinctness of the stocks as indicated by G-tests, analysis of dendrograms, etc.; (2) the average abundance of the stocks; (3) historical contribution estimates of the stocks to the mixed-stock fisheries of concern; and (4) the results of simulation analyses where the stocks included in a baseline are present in hypothetical proportions in the mixture samples. This process helps in the creation of a baseline that will generate accurate stock composition estimates from in-season mixture samples.

THE MAXIMUM LIKELIHOOD ESTIMATION MODEL

A MLE model is used to estimate stock contributions from mixed-stock fishery samples and to evaluate the accuracy of stock composition estimates from simulated mixtures. Pella and Milner (1987) define the MLE estimate of stock composition as the one which maximizes the joint probability of observing the genotype frequencies of the baseline stocks (known composition) and the stocks in the mixture or fishery sample (unknown composition). Numerous MLE algorithms have been developed (e.g., Fournier et al. 1984; and Millar 1987), however in general they provide very similar or identical estimates of stock composition if input data and model options and parameters are consistent. The Commission currently uses a MLE model based on one developed by Fournier et al. (1984) and modified by researchers from the Pacific Biological Station (PBS). Because the Commission's VAX computer system is compatible with that of PBS's, revisions of the MLE model are accessible.

The MLE program uses three input files to produce stock composition estimates: (1) startup file, (2) baseline or learning file, and (3) fishery or mixture file.

The startup file contains information that dictates how the program will interpret the baseline and mixture files. Some of the information that can be specified includes:

- number of stocks in the baseline file
- number of loci (traits) scored in baseline stocks
- maximum number of function evaluations that the program will perform to estimate the stock proportions
- convergence criterion, which specifies the numeric limit of when an acceptable solution to the mixture problem has been found
- option of bootstrapping the mixture file
- option of bootstrapping the baseline file
- number of iterations (individual MLE estimates)
- random number seed
- FORTRAN format for reading the mixture file

The baseline file contains the genotype frequencies at each locus for stocks in the baseline. The genotype data for each locus can be entered as observed (raw) or Hardy Weinberg corrected frequencies. Many researchers (e.g., Wood et al. 1987) have found that using genotype frequencies corrected to Hardy Weinberg expectations generally improves the performance of the MLE model. The PSC also uses Hardy Weinberg corrected values rather than the observed values. The number of genotypes that have been scored at each locus is listed on the first line of the baseline file. The order of listing the loci for each stock in the baseline file must be consistent, however the order of listing the stocks may be varied. The program matches the data from stocks of unknown identity contained in the mixture file to the most similar stocks in the baseline file.

The fishery or mixture file contains genetic data of unknown mixture composition. The genotypes at each locus are assigned a numeric code (i.e., usually 1, 2, 3, 4 ...). The order of listing the genotypes for each locus in the fishery file must correspond exactly with the order that they are listed in the baseline file. The identical order of listing loci and genotypes also applies

when simulation analyses are used to construct mixture files of known stock composition for examination of model performance.

The two main methods of grouping stocks during the MLE procedure are commonly called pool-allocate and allocate-sum. The pool-allocate option requires pooling the genetic data in the baseline file for several similar stocks (e.g., an entire stock group) prior to performing MLE analysis. The allocate-sum technique involves keeping the genetic data for each of the stocks separate during MLE and then combining the individual stock contribution estimates into stock groups after MLE. Commission researchers and others (e.g., Wood et al. 1987) have found that the allocate-sum method generally produce more accurate stock-group contribution estimates than the pool-allocate approach.

SELECTION OF STOCKS TO INCLUDE IN BASELINES

As noted earlier, a critical element in selecting the most appropriate stocks to include in a baseline is their genetic distinctness, and in particular, how different they are from stocks in other stock groups. Stocks that are the most genetically distinct are more attractive to incorporate in a baseline because they have lower misallocation bias associated with them during MLE.

The average production or run size of stocks is a key factor that is examined when stocks are selected for inclusion in a baseline. The run size data for all stocks that may be included in the baseline are obtained from several sources. Stocks with large run sizes that are likely to be contributing to the fisheries of concern should be represented in the baseline, unless there is a valid reason to exclude them. Information on the migration pattern and timing of the stocks that are candidates for inclusion in the baseline is available from numerous reports (e.g., Vernon et al. 1964, Hourston et al. 1965) and through communication with various fisheries agencies.

Another factor considered during assembly of a baseline stems from an inherent bias in the MLE model. The model is biased towards allocating a small proportion (generally $< 2\%$ per stock) to stocks only because they are included in a baseline rather than because their genetic profile most closely matches some of the fish electrophoretically analyzed in a mixture sample. Therefore, consideration is given to representing each of the stock groups in a baseline with a sufficient number of stocks so that the overall effect of this type of bias will be minimal, i.e., over a range of mixture analyses the misallocation bias will be spread among the stock groups and not adversely affect fisheries management applications. However, the inclusion of a stock in a baseline is still subject to the constraints of genetic distinctness discussed earlier.

For various reasons, stocks that are considered important to fishery managers may lack reliable genetic data at some loci. When the questionable data results from tissue samples of poor quality, rather than rejecting a stock or loci from inclusion in the baseline, an option is to use "surrogate" genetic data. The genotype frequencies for the problem loci are substituted from the stock within the same stock group that is most genetically similar and preferably in closest geographic proximity to the stock represented by the unreliable genetic data. For example, in the Canada south coast stock group, if there were poor data at the locus G3PDH-1 for the Wakeman River stock, the surrogate genotype scores at this locus from the "nearby" Kakweiken River stock might be substituted. This option is used on a very limited basis during baseline construction, i.e., confined to a few loci and stocks so that the stock discrimination power of the

baseline is not adversely affected. The stock(s) with missing genetic data are usually re-sampled and analyzed at the next opportunity and the true genotype scores ascertained and used.

SIMULATION ANALYSIS OF PRELIMINARY BASELINES

Preliminary baselines are evaluated after the initial selection of stocks has been made. The accuracy and precision of stock contribution estimates that would be produced from applying the MLE model to proposed baselines and hypothetical, in-season mixture samples are assessed using simulation analysis. This is an important step in developing baselines that generate accurate estimates of stock composition. Commission researchers have found that when baselines perform well during simulation analyses (i.e., high accuracy and precision of estimates) under typical stock composition scenarios (i.e., high proportions of Fraser River pink salmon in many of the mixture samples), it is probable that they will also perform well on an in-season basis.

Simulated fisheries are created by using a computer program that draws from a baseline, true (i.e., fixed) proportions of "fish" (genotypic data) from specific stocks and creates hypothetical, mixed-stock fishery collections from these known stock proportions. The MLE estimates of the simulated fisheries are bootstrapped approximately 100 times, so that variance measures of the stock contribution estimates can be calculated. The simulation tests on preliminary baselines are useful in assessing the accuracy and precision of the estimates under various stock compositions. Several types of simulation analyses are performed.

a. Simulations where stocks from each of the stock groups are present over the entire range of possible mixture combinations.

This series of simulations provides analysts with an overview of how the baseline will likely perform over a broad range of possible fishery mixture combinations. Because of the vast number of possible simulations this could involve, usually stocks or stock groups are evaluated at fixed proportion intervals of 10%, or more (i.e., true proportions of 0%, 10%, 20%, 30% ...).

b. Simulations that reflect expected stock contribution proportions in fisheries where large harvests of the stocks of most concern (e.g., stocks from the Fraser River) would occur.

These simulations give a good indication of how the baseline will perform in helping fishery managers achieve specific objectives, e.g., accurate allocation of catches of Fraser River pink salmon. From a fisheries management perspective, the results of these simulations are the most important of the four types discussed. Because emphasis can be placed on examining a relatively narrow range of mixture combinations, specific stocks (or stock groups) of interest are often evaluated at 5% contribution-interval proportions, or less, so that a thorough understanding of the bias and variance patterns can be acquired.

c. Simulations that examine stocks that are within the problem clusters as suggested by examination of dendrograms.

In situations where specific stocks from different stock groups have low or intermediate genetic difference among them, it is likely that MLE estimates produced when these stocks are present in high proportions in the mixtures, will have more bias and less precision associated with

them. These simulations are confined primarily to stocks within the problem clusters and are evaluated with the stocks set at true-proportion intervals of 5%, or less.

d. Simulations that explore the effect of varying mixture sizes (number of fish per collection from mixed-stock fisheries) and number of loci on the accuracy and precision of estimates.

Simulation studies generally indicate that when the mixture sizes and number of polymorphic loci included in a baseline are increased, the accuracy of the MLE estimates increases and the variance decreases. Simulation analysis is useful for determining appropriate in-season fishery sample sizes, as well as which loci and the number of loci to analyze in order to yield stock composition estimates that will meet specific management requirements for accuracy and precision. For example, simulation analysis might indicate that by increasing fishery sample sizes from 150 fish to 300 fish in specific fisheries, the result would be an increase in the accuracy of MLE estimates by an average of 6% per stock group.

Examination of the results of these four types of simulations provides guidance on which stocks to retain, omit, or pool in a baseline. After a baseline has been modified it is re-evaluated with similar types of simulations. This process is continued until a baseline is developed that is most suited to meet specific fisheries management goals.

NORTH AND SOUTH COAST BASELINES

Two separate baselines are currently used for analyzing the composition of mixed-stock fisheries where Fraser River pink salmon may be present (Table 4). The "South Coast baseline" is applied to mixed-stock fisheries occurring south of Cape Caution (Figure 5). This baseline includes pink salmon stocks from the following stock groups: (1) Fraser River, (2) Puget Sound, and (3) Canada South Coast. Pink salmon stocks from northern British Columbia are not included because it is unlikely that they would make significant contributions to mixed-stock fisheries occurring south of Cape Caution. By omitting stocks from northern British Columbia in this baseline the bias in the MLE estimates is reduced and the precision increased because the model analyzes a simpler mixture problem.

The "North Coast baseline" is used for analyzing samples from fisheries occurring north of Cape Caution. It comprises stocks from the three stock groups included in the South Coast baseline, as well as stocks from the Canada North Coast stock group (Figure 5). The Canada North Coast stock group includes British Columbia pink salmon stocks located from just north of Cape Caution to the Alaskan border.

Table 4. Pink salmon stocks often included in GSI baselines used by the Commission.

STOCK GROUP	STOCK
FRASER RIVER	Fraser mainstem Thompson River Bridge River Seton Creek Harrison River Vedder River Coquihalla River
PUGET SOUND	Nooksack River Skagit River Snohomish River Stillaguamish River Hamma Hamma River Duckabush River Dosewallips River Dungeness River
CANADA SOUTH COAST	Wakeman River Kakweiken River Glendale Creek Adam River Quinsam River Keogh River Phillips River Puntledge River
CANADA NORTH COAST	Johnston Creek Kilbella River Koeye River Kwatna River Bella Coola River Atnarko River Salloompt River Kainet Creek Mussel River Kemano River Quaal River Kitimat River Kumealon Creek Skeena River Andesite River Lakelse River Morice River Babine River Kispiox River Kitwanga River Khutzeymateen River Kwinamass River Ishkheenickh River Iknouk River Stagoo River

PLANNING IN-SEASON GSI SAMPLING PROGRAMS

Planning in-season GSI programs is initiated (Figure 3) after baselines for in-season mixed-stock fishery analyses are assembled. This stage is normally done in the spring and early summer of odd-years and includes: selecting appropriate fisheries to sample; arranging in-season electrophoretic analysis of the tissue samples; and training GSI port-samplers.

SELECTION OF FISHERIES TO SAMPLE

Fraser River pink salmon are harvested in purse seine, gillnet, troll, Indian food, recreational, and test fisheries over a large geographic area spanning numerous fishing areas from both Canada and the United States. Therefore, given budgetary and manpower constraints, it would not be advisable to try and sample all fisheries where Fraser River pink salmon could be caught. The following factors are considered when selecting fisheries to sample.

1. Fisheries where large numbers of Fraser River pink salmon are likely to be caught.

Assigning high priority to this factor is important to help achieve the mandate of accurately accounting for Fraser River pink salmon wherever they are caught. Information on fisheries that have historically harvested large numbers of Fraser River pink salmon is available from data collected by the Commission's GSI program in previous years as well as tagging study reports (e.g., Vernon et al. 1964; Hourston et al. 1965).

2. Fisheries where a large proportion of the catch are likely to be Fraser River pink salmon and where the sampled fish may yield additional information for fisheries management applications.

For example, in-season estimation of the run size of Fraser River pink salmon using scale pattern analysis can be aided when GSI has confirmed that fish harvested in specific fisheries are primarily of Fraser River origin.

3. Fisheries in terminal areas where the stock composition is known.

Another test of the probable accuracy of the GSI estimates can be conducted when the stock composition is known. For example, collections of tissue samples are taken from gillnet test-fisheries occurring in the Fraser River where only Fraser River pink salmon are caught. If stock composition analyses estimate mixture compositions substantially less than 100% Fraser River pink salmon, then it provides analysts with a measure of the under-estimate bias that might occur when this stock group dominates in a fishery.

4. Fisheries where political or other considerations may require additional sampling effort.

Other fisheries management agencies may request additional GSI sampling of particular fisheries to help satisfy their information requirements.

From the above list, factor number one is typically the most important because of the strong emphasis on fulfilling catch allocation and escapement goals for Fraser River pink salmon. However, in some years situations may emerge that increase the importance of the other factors.

Lists of fisheries that should be sampled are compiled, and summarized in tables. An example of a pre-season plan for sampling fisheries in Johnstone Strait and the west coast of Vancouver Island for pink salmon GSI is shown in Table 5.

Table 5. Example of a proposed sampling program for collecting pink salmon tissue samples from mixed-stock fisheries.

W/END PERIOD[1]	STATISTICAL AREAS AND SUB-AREAS TO SAMPLE					
	12-1 to 12-7 PS[2]	12-8 to 12-12 PS	13 PS	123 to 124 TR[2]	125 to 127 TR	20 PS
JULY 17					X	
JULY 24	X[3]	X		X	X	X
JULY 31	LX[4]	X	X	X	X	X
AUGUST 7	X	X	X	FX	X	X
AUGUST 14	FX[5]	FX	X	FX	FX	X
AUGUST 21	FX	FX	FX	FX	FX	FX
AUGUST 28	FX	FX	FX	X	FX	FX
SEPTEMBER 4	FX	FX	FX	X	FX	FX
SEPTEMBER 11	FX	FX	FX	X	X	X
SEPTEMBER 18	X	X	FX		X	

[1] Week-ending periods are designated here as ending on Saturdays and include samples collected from fish caught the previous Sunday through the week and including Saturday, for a catch period of seven days.

[2] PS = purse seine fishery; TR = troll fishery

[3] "X": standard fishery collection of 150 muscle tissue samples from individual pink salmon

[4] "LX": laboratory replicate collection, i.e., two or more identical collections of 150 tissue samples removed from the same 150 fish. These collections are independently analyzed by two or more different laboratories to check for consistency in the electrophoretic analyses between different laboratories.

[5] "FX": fishery replicate collection, i.e., two or more separate collections of 150 fish from the same fishery (e.g., an "FX" of 300 fish may be planned for a specific fishery, e.g., Area 12-1 to 12-7 purse seine week-ending period August 14). These collections are intended to increase the accuracy of stock composition estimates for specific fisheries.

LABORATORY ANALYSIS OF IN-SEASON TISSUE SAMPLES

Contracting laboratories to analyze tissue samples occurs in late spring of odd-years and involves: (a) sending a "Request For Proposal" (RFP) to all laboratories that may be capable of fulfilling all or part of the required electrophoretic analyses; (b) evaluating the RFPs submitted by the laboratories; (c) sending a "Letter of Intent" to the chosen laboratory(s) which identifies the amount of the total contract that the PSC is offering the laboratory to perform; and (d) preparing contracts with the laboratories that have been awarded all or part of the work.

To ensure the accuracy and consistency of laboratories performing in-season electrophoretic analyses for the Commission their results are assessed using "laboratory replicate" testing methods (White and Shaklee 1991). This testing method involves comparing the genotype scores that have been recorded by two or more laboratories from their analysis of identical pink salmon muscle tissue samples (i.e. samples taken from the same fish).

GSI PORT-SAMPLERS

Technicians are hired to collect the tissue samples at ports in Canada and the United States where significant landings of pink salmon may be off-loaded. CDFO and WDFW, as well as other agencies have helped sample landings for the Commissions's pink salmon GSI program in previous years.

THE IN-SEASON GSI PROGRAM

The baseline and in-season sampling programs require the collection of high quality, representative tissue samples to produce accurate stock contribution estimates. This is the case regardless of how well the subsequent laboratory analysis of the tissue specimens and MLE analyses are performed.

Final planning and preparation of the in-season sampling equipment and preliminary training of port samplers occurs in May and June of odd years (Figure 3). The first collections are often obtained from mixed-stock fisheries occurring around the northern portion of the Queen Charlotte Islands in July. As the season progresses into August and September, an increasing proportion of the collections are taken from southern fisheries where Fraser River pink salmon are traditionally caught in large numbers, (Johnstone Strait and Juan de Fuca Strait net fisheries, west coast Vancouver Island troll fisheries, and Salmon Banks and Point Roberts net fisheries). The final in-season collections are usually gathered from terminal gillnet test-fisheries occurring in the lower Fraser River in late September. These terminal area collections are exclusively Fraser River stocks and are used to assess the accuracy of the GSI estimates rather than for catch allocation.

IN-SEASON SAMPLING EQUIPMENT

Much of the equipment used in the collection of tissue samples from baseline or spawning stocks is also used in the in-season program. The equipment that is used for taking in-season collections is listed in Appendix C.

IN-SEASON SAMPLING

A standard collection from a mixed-stock fishery involves taking samples of muscle tissue from 150 different pink salmon. For high priority fisheries, the sample size may be increased to 300 or more fish. During the in-season sampling period, the term "collection" refers to a set of tissue samples, randomly selected from individual pink salmon caught in a fishery defined by a specific statistical area, user group (e.g., purse seiners) and week-ending date (e.g., a collection from week-ending Saturday July 24 would include fish caught from Sunday July 18, to Saturday,

July 24, inclusive). Additional details on the following sections in this report which cover: establishing good communication links during the sampling season; selecting vessels and fish to sample; and sampling tissues are discussed in greater detail in an unpublished PSC report ("GSI Sampling Procedures For Pink Salmon Caught In Mixed-Stock Fisheries", 1995). This training manual is given to port samplers prior to them collecting GSI samples.

Communication

Commission staff generally make arrangements with managers of fish processing plants where tissue sampling may be conducted several weeks before the first in-season collections are taken. Regular communication is maintained between port samplers, Commission and other government staff and managers of fish processing plants to facilitate sampling of fisheries.

Selecting vessels for catch sampling

Port samplers are given instructions on the type of fishing vessel that samples are required from, (e.g., tender, purse seiner, gillnetter, troller) prior to sampling the catch from specific fisheries. The preferred type of vessel for catch-sampling is generally a tender vessel. Tender vessels have usually collected fish from several different fishing vessels that have participated in a fishery. Consequently, fish that are randomly sampled from them are normally assumed to be representative of the overall harvest from a fishery. Samples are only taken from catches of known origin, i.e., information on the exact statistical area(s) fished and the catch date(s) is obtained from the crews of the vessels.

Two main sampling strategies are used. Selecting the appropriate strategy is influenced primarily by the availability of tender vessels. If the catches from tender vessels are available, then approximately 75 fish are sampled from each of two tenders. If only one tender vessel is suitable for sampling, all 150 tissue samples are usually taken from the same vessel, and the names of the individual vessels contributing to the catch aboard the tender vessel are recorded.

In situations where tender vessels are not available, it is necessary to sample catches from individual vessels (e.g., purse seiners). Samples are usually obtained from at least five individual vessels (approximately 30 fish/vessel). Efforts are made to avoid sampling fewer than three vessels (50 fish/vessel) so that collections will be sufficiently representative of the overall harvest from fisheries.

Selecting fish for tissue sampling

It is important to emphasize that producing accurate stock composition estimates requires that fish be randomly sampled. Bias can easily be introduced into stock identification analyses when only 150 fish are used to estimate the stock composition of mixed-stock fisheries where several hundred thousand fish have been harvested. The shoreworkers are questioned to confirm that fish in the catch selected for sampling have not been graded according to size or physical condition. Selecting fish for sampling based on physical characteristics is avoided. The fish from a catch chosen for sampling should be randomly distributed with each fish having an equal probability of being sampled.

Sampling tissues: general instructions

Several considerations are observed when GSI samples are collected. Tissue samples that have not been properly taken or were not adequately frozen are difficult or impossible to analyze electrophoretically (e.g., little staining activity, smeared banding patterns).

Port-samplers try to minimize damage to fish during tissue sampling because it may reduce the market-value of the fish for the processors and reduce their cooperation in the sampling process. This is not generally a problem because the instrument used for sampling muscle tissue removes only a minute portion of tissue and leaves a barely detectable mark on the sampled fish.

Samples are normally collected in shaded areas of processing plants to avoid exposing the tissue samples to sunlight, which could reduce tissue quality. The tissue samples are frozen immediately on dry ice to minimize denaturing of the enzymes.

Sampling muscle tissue

A sample of muscle tissue is taken from the left side (when facing in a posterior to anterior direction) of each fish (Figure 6). If the left side of the fish is not suitable, the sample is taken from the right side. The tissue samples are extracted with a metal tissue borer (often brass or stainless steel) which is approximately 10 cm long, with an inside diameter of about 0.7 cm. The quantity of muscle tissue taken is sufficient to fill the bottom one-fourth of a test tube (test tubes are usually 7.5 cm long, 1.0 cm in diameter). This is obtained by extracting a plug of muscle tissue approximately 2.0 cm long. A standard quantity of muscle tissue per sample is preferred because too little or too much tissue may cause problems during laboratory processing of the samples. The tissue borer and plunger are rinsed in freshwater after sampling each fish so that tissue fragments from the fish being sampled do not contaminate the next tissue sample. Saltwater is not used for rinsing the tissue sampling equipment because it could cause deterioration of the tissue samples and equipment.

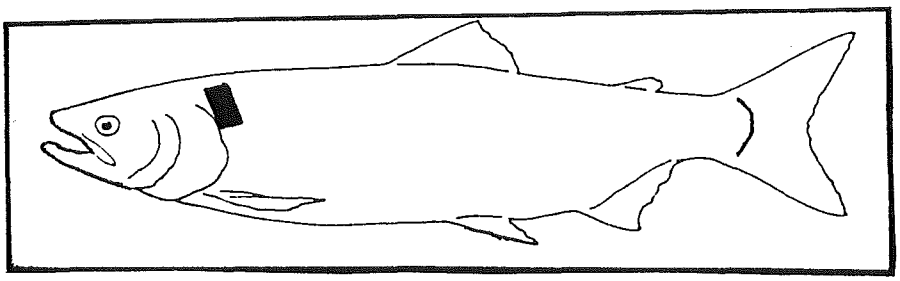


Figure 6. Nape area of pink salmon where muscle tissue samples are taken.

Laboratories conduct their most accurate electrophoretic analyses from samples of muscle tissue from pink salmon that are: consistently sized (always about 1/4 of test tube volume); contain no excess water or other foreign material; placed in the bottom of unflawed test tubes; fresh and pure pink salmon muscle tissue; and have been kept frozen on dry ice or in an ultra-cold freezer (-70°C, or colder).

Length, weight and sex data

The post-orbital fork (POF) length of pink salmon is measured (in mm) on a flat surface from the posterior edge of the eye socket to the fork of the tail. The total weight (in pounds) of the fish sampled for a collection is recorded and the average weight per fish is calculated. The sex of each fish is determined by making a neat incision along the belly and visually examining the gonads.

ELECTROPHORETIC ANALYSIS OF IN-SEASON TISSUE SAMPLES

Electrophoretic analysis of the in-season tissue samples is the last stage in the GSI data collection process (Figure 3) before the genetic data are used in stock composition analyses. The steps in this phase are: shipping the tissue samples to an electrophoretic laboratory; electrophoretically analyzing the samples; and transmitting the raw genotype scores to the PSC.

When a port sampler contacts the PSC office with shipping information for a collection, the information is relayed to the laboratory as well as the priority for analyzing the collection. The Commission requires that high-priority collections be analyzed and the results provided within 72 hours of their arrival at a laboratory. In most cases, tissue collections are in transit for two to 18 hours before arriving at a laboratory. Samples are usually shipped with enough dry ice in the transport-coolers to preserve samples for 48 hours, which helps safeguard collections that are delayed during transit from thawing.

The laboratories conduct electrophoretic analysis of the muscle tissue samples for the required loci and alleles. The laboratories use predetermined numeric codes for the genotypes at each locus so that PSC analysts can immediately apply MLE to these data upon receiving them. Laboratories also provide information regarding problems they encountered during the analyses and their mode of transmitting these data to the Commission, e.g., electronic files via e-mail, or printed files via FAX. If e-mail is used then the data file is transmitted twice so that a "file comparison" can be done to verify correct data transmission.

STATISTICAL ANALYSIS OF IN-SEASON GENETIC DATA

The electrophoretic data provided by laboratories are analyzed using a MLE program and bias-correction methodologies may be applied to increase the accuracy of the stock composition estimates.

MLE ESTIMATION OF STOCK COMPOSITION

On an in-season basis PSC analysts focus on generating accurate stock contribution estimates rapidly so that fisheries managers can develop appropriate regulations that will help achieve in-season management goals. The bootstrap re-sampling method of Efron (1982) is used for assessing the precision in stock estimates caused by sampling errors in the relative frequencies of genotypes in the baseline and mixture files. Bootstrapping involves re-sampling genetic data contained in the baseline and mixture files with replacement, to make new samples that are the

same size as the original samples. MLE is then applied to these "new" baseline and mixture files and stock composition estimates are made. Standard errors for stock or stock group composition estimates are generally calculated by bootstrapping estimates 100 times.

A preliminary stock composition estimate for a mixed-stock fishery is usually produced in less than one hour after receiving genetic data from a laboratory. The MLE program is run on a VAX computer and stock proportion estimates are sent to an ASCII file. The output file is then copied to a spreadsheet on a personal computer and contribution estimates by stock and stock group are calculated along with accompanying estimates of precision. The composition estimates by stock group are multiplied by the fishery catches, which allows the total number of pink salmon caught by stock group, area, user group, and time period to be estimated and updated regularly throughout the season.

BIAS-CORRECTION OF STOCK COMPOSITION ESTIMATES

Bias or inaccuracy in stock composition estimates is undesirable because it reduces the Panel's ability to achieve catch allocation and escapement goals for Fraser River pink salmon. Numerous factors at different stages of the GSI process can cause bias. For example, during the baseline or in-season field sampling stages, selecting fish based on external characteristics, e.g., size, may decrease the accuracy of stock composition estimates because samples were not taken randomly. During laboratory analyses, bias may occur for numerous reasons, e.g., poor quality gels that make it difficult to interpret genotypes correctly. Several measures for reducing bias imparted during field sampling and laboratory analysis have already been discussed.

MLE models have an inherent bias to underestimate the contribution of stocks that are major contributors to fisheries and overestimate the presence of those stocks that are minor or non-contributors to a fishery (Wood et al. 1987). Bias in the MLE estimates can be increased (worsened) by several factors; some of the main ones are: (1) small genetic difference between stocks or stock groups in a baseline; (2) baseline stocks or fishery samples represented by small sample sizes; and (3) large differences in relative contributions of specific stock groups to mixed-stock fisheries.

Several approaches have been suggested by researchers for reducing bias in stock composition estimates that are attributable to MLE models (e.g., Beacham et al. 1985; Pella and Milner 1987). Some of them entail using simulation analysis to determine the amount of bias in MLE estimates based on a wide range of possible stock composition scenarios. When MLE analysis is performed on a fishery mixture sample, the stock composition estimates are corrected by the amount of bias that is identified by the simulation studies. Preliminary research conducted by Commission analysts favours this approach. To date, a universally accepted method of correcting bias in stock composition estimates from MLE is not available.

APPLICATION OF STOCK COMPOSITION ESTIMATES FOR FISHERIES MANAGEMENT

Sufficient data on pink salmon catches and accompanying stock contribution estimates have usually accumulated by early August to provide fishery managers with updates of pink salmon

catches by stock grouping, statistical area, user group and country. Fraser River pink salmon are managed as a single stock because the individual stocks comprising the group cannot be reliably distinguished in mixed-stock fisheries due to low genetic difference among them and because the different Fraser stocks co-migrate through fisheries at similar times. There are three important management applications of pink salmon GSI estimates: (1) estimating the run size of Fraser River pink salmon; (2) helping to achieve escapement and catch allocation goals; and (3) projecting migration patterns and timing of Fraser River pink salmon during the in-season management period. These applications are discussed below.

The first use of GSI estimates is in the production of run size estimates. Fraser River pink salmon run size estimates are generated during the in-season management period by PSC staff. These in-season changes to run size help the Fraser River Panel in developing fishery regulations designed to achieve escapement and catch allocation objectives. The main in-season run-size estimation models used by the Commission include: the peak catch per unit effort (CPUE) of pink salmon caught in the Area 20 purse seine fishery (PSC 1990); a cumulative normal distribution (CND) model that requires current estimates of Fraser River pink salmon catches in all fisheries (PSC 1994); and scale pattern measurements, e.g. distance to the annulus (Mike Lapointe, PSC, Vancouver, pers. commun.). All of these models require accurate GSI estimates of Fraser River pink salmon caught in specific fisheries. Accurate GSI estimates are required to produce reliable in-season run-size estimates.

The Panel's predominant goal in managing Fraser River pink salmon stocks is to ensure that the gross escapement goals established by CDFO are achieved. CDFO sets the gross escapement goal for Fraser River pink salmon stocks taking into account the pre-season forecast of run size. Data that are collected in-season, including GSI estimates, are used to update the estimate of total run size and to calculate the cumulative catch of Fraser River pink salmon. Subtraction of the catch to-date from the estimated run size provides fishery managers with an estimate of the number of pink salmon remaining that are available for catch or escapement. Subsequent fishery openings or closures are modified to help achieve the gross escapement goals.

The second major goal in the management of Fraser River pink salmon by the Panel is the international allocation of the total allowable catch. The cumulative catch of Fraser River pink salmon by country provided through the GSI program gives the Panel the information necessary to help achieve international catch allocation targets. Domestic allocation is the third priority in this management process. The Canadian domestic allocation process requires a summary of catch by user group of all southerly migrating pink salmon stocks. The PSC staff collects and summarizes the data required by Canada and the United States for their domestic allocation processes. The allocation of Fraser River pink salmon both internationally and domestically, and the estimates of stock groupings that are required to meet the allocation objectives are summarized in Table 6.

Table 6. Stock composition estimates provided by the GSI program for international and domestic allocation of Fraser River pink salmon catches.

	INTERNATIONAL ALLOCATION	DOMESTIC ALLOCATION	
		UNITED STATES	CANADA
ALLOCATION BY COUNTRY AND USER GROUP	U.S. fishermen Cdn. fishermen	Treaty Indian Non-Indian	Inside Troll Outside Troll Purse Seine Gillnet
GSI ESTIMATES REQUIRED	% Fraser River	% Fraser River	% Southerly Migrating Stocks (Fraser, Puget Sound, Cdn. S. Coast)

Accounting of pink salmon catches is done by Commission analysts on a computer spreadsheet. The spreadsheet is divided into three sections where catches from different geographic areas are accounted: (1) Canada south coast (Canadian catches occurring south of Cape Caution); (2) Washington and Oregon waters; and (3) Canada north coast (Canadian catches occurring north of Cape Caution). Within each of the three sections of the spreadsheet, catches of fish from each stock group by statistical fishing area, user group and week-ending date are tabulated. Pink salmon catches from commercial, Indian food, recreational and test fisheries are compiled and entered into the spreadsheet. Catch estimates from each fishery are then multiplied by the relevant stock-group contribution estimates to calculate the catch by stock group, statistical area, user group, and week-ending date. Tables are constructed from these data that summarize catches by stock group, country, and user group within each country. The catch estimates of Fraser River pink salmon are updated regularly throughout the fishing season to help guide fisheries management decisions. Post-season estimates of the total catch made by each user group are calculated and compared to the relevant allocation goals.

Information on the run timing and migration patterns of Fraser River pink salmon in various fisheries is a third use of the GSI estimates. This information also aids in the development of fishery regulations. For example, GSI estimates in two major Canadian pink salmon fisheries (Areas 12 and 20 purse seine) are used in calculating the proportion of the run approaching the Fraser River through each migratory route. The proportion of the run approaching via Johnstone Strait is termed the "diversion rate". The estimates of diversion rate are considered during the development of fishing regulations for two main reasons. First, Canadian purse seine fisheries in Johnstone Strait and in Juan de Fuca Strait have different exploitation rates for equal fishing periods. The Johnstone Strait fishery harvests a higher proportion of the weekly abundance of available pink salmon in one day of fishing than the Juan de Fuca fishery because of the greater lineal distance and, hence, number of days of migration available for harvest. The different exploitation rates affect the number of fish escaping these fisheries and if not considered, may compromise achieving escapement goals. A second reason for considering the diversion rate is

that in years when a large proportion of the Fraser River pink salmon run migrates through Johnstone Strait, there are proportionally fewer fish available for harvest in United States Fraser River Panel Area waters. Consequently, pre-season fishing plans may have to be modified to meet international allocation goals.

Calculation of the diversion rate requires weekly estimates of the number of Fraser River pink salmon harvested in the Johnstone Strait and Juan de Fuca Strait fisheries. These catch estimates are entered into a CND model used by the Commission to estimate the run size of Fraser River pink salmon. The model uses estimates of Fraser River pink salmon catches to date, test fishing catch per unit effort (CPUE) data, and historic exploitation rates to "reconstruct" daily abundances of the run at reference areas, such as Area 20 (PSC 1990). The run-size estimates from the CND model assist in developing fisheries management strategies. The GSI program provides necessary input data for various models to allow for the in-season estimation of abundance, diversion rate and stock group timing.

POST-SEASON GSI ANALYSES

The main goal of post-season GSI analyses is to maximize the accuracy of Fraser River pink salmon catch estimates for each year that the GSI program has operated. Improvements in the accuracy of the in-season MLE estimates are gained from additions or corrections to electrophoretic data that were incomplete during in-season periods.

During the in-season electrophoretic analysis of tissue samples, laboratory staff note tissue samples and/or specific loci that might yield more accurate data if reanalysed. In addition to re-analyzing certain samples, some collections that were not analyzed in-season due to time or budget constraints are processed. Additional baseline data are available for reanalysing the in-season collections after analyses of tissues from spawning ground collections are completed.

Advances in the accuracy of the MLE estimates are also achieved by conducting more intensive statistical analyses. Updates on pink salmon catch estimates obtained post-season further improve the accuracy of Fraser River pink salmon catch estimates.

CONCLUSIONS AND FUTURE RESEARCH

This report provides an overview of the Commission's pink salmon GSI program. The program helps the Commission to meet its management obligations for Fraser River pink salmon as mandated by the Pacific Salmon Treaty. Prior to 1987, run-reconstruction techniques (applied post-season) were used to account for catches of Fraser River pink salmon. Beginning in 1985 with a pilot study, and continuing from 1987 to 1995, GSI techniques have been used to assist the Fraser River Panel achieve its mandates of: (a) accounting for Fraser River pink salmon wherever they are harvested; and (b) managing fisheries targeting on these stocks within the Fraser River Panel Area.

Stock composition estimates provided by GSI are now an integral part of the Commission's program for achieving escapement goals and international and domestic catch allocation of Fraser

River pink salmon to Canadian and United States fisheries. These estimates are also part of the in-season weekly data input into run size estimation models. The use of GSI estimates for run size estimation and in migrational analyses could increase in importance if the Fraser River Panel requires this information on an accelerated basis during the in-season management period.

Many of the Commission's run-size estimation relationships currently utilize Fraser River and non-Fraser pink salmon stocks. Quantification of these relationships would likely be improved if the contribution of non-Fraser River stocks could be estimated more accurately in the historical data. As the GSI data base expands, these historic relationships may be replaced by more accurate estimates of the probable contribution of Fraser River pink salmon.

Research to improve the genetic data collected for stock composition analyses involving Fraser River pink salmon will continue. Additional loci will be incorporated in the baselines for MLE analyses if electrophoretic and statistical analyses suggest that their inclusion would provide a cost-effective means of increasing the accuracy of the stock composition analyses.

It is probable that information provided by the Commission's GSI program will have additional applications for fisheries management. One important use may be in helping to develop fishing regulations with special considerations to minimize over-harvest of weak or endangered other stocks and species of salmon in mixed-stock fisheries so that their genetic variability is maintained (Waples et al. 1990; Allendorf et al. 1987).

Research involving the statistical algorithms used for calculating stock composition estimates will be pursued, including bias-correction methodologies for the MLE model. The newer statistical computer programs such as the Numerical Taxonomy and Multivariate Analysis System: NTSYS-pc (Rohlf 1994) will be examined to assess their utility for analyzing genetic data.

The GSI program has proven useful for aiding in the in-season management of Fraser River pink salmon. The value of the GSI program will increase as successive years of data are added to the 1987 to 1995 base years.

ACKNOWLEDGMENTS

The author thanks Jim Gable, Mike Lapointe, Jim Woodey, Ian Guthrie, and Ross Davies of the Pacific Salmon Commission; James Shaklee and Sewell Young of the Washington Department of Fish and Wildlife; and Dennis Rutherford of the Pacific Biological Station for their valuable comments on drafts of this report. The author also appreciates the help of Janice Abramson and Teri Tarita of the Pacific Salmon Commission in compiling the final report.

REFERENCES

Aebersold, P.B., G.A. Winans, D.J. Teel, G.B. Milner, and F.M. Utter. 1987. Manual for starch gel electrophoresis: a method for the detection of genetic variation. NOAA Tech. Rep. NMFS 61: 19 p.

Allendorf, F.W., N. Ryman, and F.M. Utter. 1987. Genetics and fishery management: past, present, and future, p. 1-19. In N. Ryman and F.M. Utter [ed.] Population genetics and fishery management. University of Washington Press, Seattle, WA.

Beacham, T.D., R.E. Withler, and A.P. Gould. 1985. Biochemical genetic stock identification of pink salmon (*Oncorhynchus gorbuscha*) in southern British Columbia and Puget Sound. Can. J. Fish. Aquat. Sci. 42: 1474-1483.

Cavalli-Sforza, L.L. and A.W.F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21: 550-570.

Department of Fisheries and Oceans. 1995. Fraser River pink salmon. Prep. by Fraser River Action Plan, Fishery Management Group. Vancouver, B.C., 19 p.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. National Science Foundation Conference Board of the Mathematical Sciences Monograph 38. SIAM, Philadelphia, PA., 92 p.

Fournier, D.A., T.D. Beacham, B.E. Riddell, and C.A. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Can. J. Fish. Aquat. Sci. 41: 400-408.

Hourston, A.S., E.H. Vernon, and G.A. Holland. 1965. The migration, composition, exploitation and abundance of odd-year pink salmon runs in and adjacent to the Fraser River Convention Area. Int. Pac. Salmon Fish Comm. Bull. XVII: 151 p.

May, B. 1975. Electrophoretic variation in the genus *Oncorhynchus*: the methodology, genetic basis, and practical applications to fisheries research and management. M.Sc. thesis, Univ. of Washington, Seattle, WA., 95 p.

May, B., and C.C. Krueger. 1990. Use of allozyme data for population analysis, p.157-172. In D.H. Whitmore [ed.] Electrophoretic and isoelectric focusing techniques in fisheries management. CRC Press, Boca Raton, FL.

Millar, R.B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Can. J. Fish. Aquat. Sci. 44: 583-590.

Milner, G.B., D.J. Teel, F.M. Utter, and G.A. Winans. 1985. A genetic method of stock identification in mixed populations of Pacific salmon *Oncorhynchus* spp. Marine Fish. Rev. 47(1): 1-8.

Morizot, D.C., and M.E. Schmidt. 1990. Starch gel electrophoresis and histochemical visualization of proteins, p. 223-80. In D.H. Whitmore [ed.] Electrophoretic and isoelectric focusing techniques in fisheries management. CRC Press, Boca Raton, FL.

Morris, C.G. [ed.] 1992. Academic Press dictionary of science and technology. Academic Press, San Diego, CA., 2432 p.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106: 283-292.

Pacific Salmon Commission. 1990. Report of the Fraser River Panel to the Pacific Salmon Commission on the 1989 Fraser River sockeye and pink salmon fishing season. Vancouver, B.C., 51 p.

1994. Report of the Fraser River Panel to the Pacific Salmon Commission on the 1991 Fraser River sockeye and pink salmon fishing season. Vancouver, B.C., 68 p.

Pella, J.J. and G.B. Milner. 1987. Use of genetic marks in stock composition analysis, p. 247-276. In N. Ryman and F.M. Utter [ed.] Population genetics and fishery management. University of Washington Press, Seattle, WA.

Rohlf, J.F. 1994. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Exeter Software, New York.

Shaklee, J.B. and C.P. Keenan. 1986. A practical laboratory guide to the techniques and methodology of electrophoresis and its application to fish fillet identification. CSIRO Rep., Australia 177: 59 p.

Shaklee, J.B. and S. Phelps. 1990. Operation of a large-scale, multiagency genetic stock identification program, p. 817-830. In N.C. Parker, A.E. Giorgi, R.C. Heidinger, D.B. Jester, E.D. Prince, and G.A. Winans [ed.] Fish-marking techniques. Am. Fish. Soc. Symposium 7.

Shaklee, J.B., S.R. Phelps, and J. Salini. 1990a. Analysis of fish stock structure and mixed-stock fisheries by the electrophoretic characterization of allelic isozymes, p. 173-196. In D.H. Whitmore [ed.] Electrophoretic and isoelectric focusing techniques in fisheries management. CRC Press, Boca Raton, FL.

Shaklee, J.B., F.W. Allendorf, D.C. Morizot, and G.S. Whitt. 1990b. Gene nomenclature of protein-coding loci in fish. Trans. Am. Fish. Soc. 119: 2-15.

Shaklee, J.B., D.C. Klaybor, S. Young, and B.A. White. 1991. Genetic structure of odd-year pink salmon, *Oncorhynchus gorbuscha*, from Washington and British Columbia and potential mixed-stock fishery applications. J. Fish. Biol. 39 (Suppl. 1): 21-34.

Swofford, D.L., and R.B. Selander. 1981. BIOSYS-1, Release 1.7: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Heredity 72: 281-283.

Utter, F., P. Aebersold, and G. Winans. 1987. Interpreting genetic variation detected by electrophoresis, p. 1-20. In N. Ryman and F.M. Utter [ed.] Population genetics and fishery management. University of Washington Press, Seattle, WA.

Vernon, E.H., A.S. Hourston, and G.A. Holland. 1964. The migration and exploitation of pink salmon runs in and adjacent to the Fraser River Convention Area in 1959. Int. Pac. Salmon Fish. Comm. Bull. XV: 296 p.

Waples, R.S., G.A. Winans, F.M. Utter, and C. Mahnken. 1990. Genetic approaches to the management of Pacific salmon. Fisheries 15(5): 19-25.

White, B.A., and J.B. Shaklee. 1991. Need for replicated electrophoretic analyses in multiagency genetic stock identification (GSI) programs: examples from a pink salmon *Oncorhynchus gorbuscha* GSI fisheries study. Can. J. Fish. Aquat. Sci. 48: 1396-1407.

Wood, C.C., S. McKinnell, T.J. Mulligan, and D.A. Fournier. 1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and application to complex problems. Can. J. Fish. Aquat. Sci. 44: 866-881.

Wood, C.C. 1989. Utility of similarity dendrograms in stock composition analysis. Can. J. Fish. Aquat. Sci. 46: 2121-2128.

APPENDICES

APPENDIX A

Equipment and General Methodology Employed by the Commission for Collecting Tissue Samples for Development of Pink Salmon GSI Baselines

Administrative Equipment

- pencils, permanent ink felt pens
- field notebooks (preferably waterproof pages) and clipboard
- "spawning ground collection-bag" tags
- "shipping and perishables" labels for the coolers
- duct tape for sealing coolers

Sampling Equipment

- knives and sharpeners
- forceps
- measuring board and measuring stick for taking post-orbital hypural length
- buckets
- gloves, apron, chest or hip waders
- large plywood sheet and saw-horses for sampling fish upon
- fish collecting equipment (e.g., beach seine, gillnets, pews)
- sampling bags:
  - whirl-pac or sandwich bags for individual tissues
  - numbered, zip-lock bag for containing all of the tissues from each fish
  - collection bag (e.g., garbage bag) for holding all of the bags

Freezing and Transporting Tissue Samples

- small cooler used during tissue sampling
- large cooler(s) (40 litres or > capacity) used for transporting tissue samples
- dry ice (depending on outside temperature, number of samples etc., 20 to 50 kg)

Guidelines for sampling the tissues

1. Muscle. A strip of pure muscle tissue (approximately 1 cm x 2 cm x 4 cm) is removed with a sharp knife from just behind the nape area (the musculature just posterior and dorsal to the opercle). The sample should not contain any skin, cartilage, blood, dirt etc. It is then placed in a whirl-pac or sandwich bag.

2. Eyes. Both eyes are carefully removed from their sockets using a knife and/or finger and placed in a separate whirl-pac bag. Caution should be applied to avoid rupturing the eyeballs.

3. Heart. The heart is removed (making sure the ventricle portion is included) and placed in a separate whirl-pac bag.

4. Liver. A piece of liver tissue of similar dimensions to the muscle tissue sample is removed. The liver sample should not be contaminated, which could occur if the gall bladder is ruptured. The sample is then placed loose in a numbered, zip-lock bag along with the other three "bagged tissues" and the zip-lock bag is sealed. The zip-lock bag is labelled with a fish identification number and immediately placed in a cooler with dry ice.

After samples have been collected from a stock they should be placed in a collection bag (e.g., garbage bag) and sealed with a collection-bag tag. The dry ice should be spread around the zip-lock bags so that samples near the center of the collection bag are also frozen rapidly. The collection bag should then be put in a large cooler with dry ice placed above and below the tissue specimens. The longevity of the dry ice can be increased by filling empty spaces with crumpled newspaper. If the cooler is stored somewhere other than an ultra-cold freezer overnight, the cooler can be placed inside a regular chest freezer. Tissue samples from each stock must be kept in separate collection bags to prevent confusion when the samples arrive at a laboratory.

Record keeping for spawning ground sampling is normally done in field notebooks on waterproof paper and on individual tags. For each spawning ground collection, samplers record in their field notebook at the top of the first page: the stock being sampled, where in the river system it was collected, the date that the sample was taken, the method of collecting the fish, names of the samplers and additional comments that may be useful. Data on individual fish are recorded in separate columns in field notebooks as follows: fish identification number (e.g., 1-100), sex, post-orbital hypural length (in mm from the posterior edge of the eye socket to the anterior edge of the hypural plate) and additional comments such as the condition of the fish prior to sampling (e.g., carcass, gills bright red).

Tissue samples should be preserved on dry ice when they are being transported to the laboratory. If the samples are misplaced during transportation to the laboratory they may thaw and become unsuitable for electrophoretic analysis. Basic guidelines for shipping samples should be followed to reduce the risk of this occurring. The first step in shipping the tissue specimens involves filling in collection-bag tags and shipping tags. The collection-bag tags are used to seal the collection bag as well as provide information for laboratory personnel such as: stock sampled, date(s) sampled, sampler(s) name and agency, number of fish sampled, tissue types sampled, and miscellaneous comments that might be useful for laboratory staff (e.g., regarding tissue quality). The shipping tag is securely affixed to the top of the cooler and documents: (a) name and address of the laboratory, and (b) name and phone numbers (work and home) of laboratory contact persons, because sometimes coolers cannot be delivered during normal working hours and this enables the shipper to make other arrangements. A label indicating that "the cooler contains perishable items and must be kept frozen" is also fastened to the lid of the cooler to inform the shipper of the status of the shipment.

Sufficient dry ice is added to the cooler after all of the tags have been attached to the collection bag and cooler. Depending on the volume of tissue samples and the anticipated transit time (preferably a day or less), generally, 7 kg to 15 kg of dry ice is placed in the cooler with

the samples. The cooler is then sealed with duct tape, which is fastened around the seam of the cooler lid and completely around the cooler.

The cooler(s) are normally shipped to the laboratory by courier, bus, or airplane. Samples should be sent via the highest priority shipping category so that their arrival at a laboratory is guaranteed by a specific time. Shipping the cooler(s) by air usually results in the samples arriving at a laboratory rapidly, however most airlines have designated dry ice as a "dangerous good" and have strict regulations on shipping coolers containing dry ice. Before shipping specimens by air, the airline should be contacted to confirm the maximum amount of dry ice allowed per cooler.

APPENDIX B

Glossary of Terms Commonly used During Electrophoresis (from May 1975; Aebersold et al. 1987; Morris 1992)

Allele: one of several different forms of a gene.

Dimer: enzyme composed of two polypeptide chains.

Electrophoresis: the separation of electrically charged proteins (enzymes) in an electric field to show their genetic composition.

Enzyme: a protein that is a catalyst (facilitates a chemical reaction but is not used up in it). Enzymes are involved in the multitude of biochemical reactions that, together, constitute intermediary metabolism.

Genotype: the genetic characteristics that determine the structure and functioning of an organism; often applied at a particular locus to differentiate one allele or combination of alleles from another.

Heterozygous: containing two or more different alleles at a locus.

Homozygous: containing two or more identical alleles at a locus.

Isozyme: any of the various structurally related forms of the same enzyme, having the same mechanism but differing from each other in chemical or immunological characteristics.

Locus (plural, loci): the position that a gene occupies on a chromosome.

Monomer: enzyme made up of one polypeptide chain.

Polymorphism: presence of more than one allele at a locus in a species.

Tetramer: enzyme made up of four polypeptide chains.

APPENDIX C

Equipment used by the Commission for Taking In-Season Fishery Collections of Muscle Tissue Samples for Pink Salmon GSI

Administrative Equipment

- pencils, pencil sharpener, permanent ink felt pens
- elastic bands, paper clips
- field notebooks and clipboard
- cannery report forms
- collection bag tags
- shipping tags and perishables labels for the coolers
- daily report forms
- maps of statistical fishing areas
- duct tape for fastening transport-coolers
- scale envelopes (if also taking scale samples)

Sampling Equipment

- tissue borer and plunger (for ejecting tissue plug), and sharpener
- knife and sharpener
- forceps
- measuring board and post-orbital fork length measuring stick
- thermometer (optional), approximate range -30°C to +40°C
- plastic bucket
- gloves, apron, rain jacket, rubber boots (steel toes if possible)
- PSC hat (and hair net and coveralls if required by fish processing plant)
- labelled test tubes (each with collection code and fish identification number) and caps
- test tube racks
- plastic collection bags
- extra (blank): collection bags, test tubes, test tube labels, caps

Freezing and Transporting Tissue Samples

- port-coolers (16 litre capacity, for day-sampling at fish processing plants); these coolers are not used for transporting tissue samples to laboratories, i.e., they are not large enough to hold a sufficient amount of dry ice and they are usually not well insulated
- transport-coolers (40 litre or > capacity, with additional styrofoam insulation inserts); these coolers are used for shipping tissue samples from fish processing plants to laboratories
- dry ice (approximately 2 kg is used in the port-coolers and 7 kg in the transport-coolers); if transporting by air, some airlines only allow a maximum of 2 kg of dry ice per cooler, therefore samplers should always contact the airline regarding their regulations prior to shipping samples; gloves should always be worn when handling dry ice to avoid frostbite.