



Proceedings of the

# OLYMPIC WILD FISH CONFERENCE

J.M. Walton and  
D.B. Houston, Editors

March 23-25, 1983

Fisheries Technology Program  
Peninsula College

Olympic National Park

Port Angeles, Washington  
1984

GENETIC CHARACTERIZATION OF POPULATIONS IN THE SOUTHEASTERN  
RANGE OF SOCKEYE SALMON

Fred Utter<sup>1</sup>, Paul Aebersold<sup>1</sup>,  
John Helle<sup>2</sup>, and Gary Winans<sup>1</sup>

<sup>1</sup>Coastal Zone and Estuarine Studies Division  
Northwest and Alaska Fisheries Center  
National Marine Fisheries Service  
National Oceanic and Atmospheric Administration  
2725 Montlake Boulevard East  
Seattle, Washington 98112

<sup>2</sup>Northwest and Alaska Fisheries Center  
National Marine Fisheries Service  
National Oceanic and Atmospheric Administration  
Auke Bay Laboratory  
P.O. Box 155  
Auke Bay, Alaska 99821

ABSTRACT: A biochemical genetic screening was carried out on 16 collections of sockeye salmon (Oncorhynchus nerka) from southeastern Alaska through the Columbia River drainage to determine if populations within this area were sufficiently diverged genetically to be identifiable in fisheries involving harvests of mixed stocks. Ninety loci were screened and 34 were found to have identifiable genetic variations. Full data sets collected from 12 polymorphic and 38 monomorphic loci were analyzed by hierarchical and clustering procedures. These analyses indicated a moderate structuring of populations of this region, with approximately 10% of the total genetic variation being attributable to genetic differentiation at various hierarchical levels. All comparisons among populations within areas identified significant differences except between the two collections from the mainland of southeastern Alaska. A marked distinction between the Quinault River and all other areas was found. The cluster analysis also suggested divergence of the two lower Fraser River samples from the other four collections of that drainage and a possible association between the Columbia River and upper Fraser River collections. The magnitude of the genetic distinctness within this region is similar to that observed among chinook salmon (O. tshawytscha) populations sampled over a similar range. Thus, a similar potential for identification of stocks in mixed fisheries apparently exists for sockeye salmon populations in this range.

## INTRODUCTION

The unique life history and commercial importance of the sockeye salmon (*Oncorhynchus nerka*) have resulted in considerable interest in the population structure of this species by fishery biologists and managers. The magnitude of this interest is reflected by sockeye salmon being the focal species of two international commissions (International North Pacific Fisheries Commission, International Pacific Salmon Fisheries Commission) involving extensive operations and research over the past three decades. The main goal of both commissions has been to determine the geographic origin of sockeye salmon stocks on the high seas in order to make decisions concerning fishery allocations in mixed fisheries. More recently, interest in the destination of sockeye salmon harvested in the Alaskan-Canadian boundary area has intensified, as considerable uncertainty exists concerning the national origins of sockeye salmon harvested in this area.

Methods to determine national or continental origins of sockeye salmon initially included tagging (Hartt 1962), reflections of environmental differences such as scale analyses (Mosher 1963), and searches for genetic differences through serological methods (Ridgway et al. 1962). All three approaches to separate stocks were attempted in the late 1950's with apparent success. However, serological methods were ultimately abandoned because the genetic basis for detected variations became doubtful (Hodgins 1972) and satisfactory answers were being obtained by other procedures (Anas and Murai 1969).

The search for genetic variants that could identify breeding units and measure contributions in mixed stocks has persisted to the present time. Two significant developments in this search were (1) the application of electrophoretic methods

which provided a reliable genetic basis for examining population structures (Utter et al. 1974; Allendorf and Utter 1979), and (2) the development of statistical and data processing procedures for estimating contributions of known genetic units in stock mixtures (Grant et al. 1980; Milner et al. 1981; Miller et al. 1983). These capabilities have given fishery managers a powerful alternative to traditional methods of mixed stock identification.

This study examines the electrophoretic variation in 16 collections of sockeye salmon from Prince of Wales Island, Alaska southward through the Columbia River drainage. Our purpose is to determine the feasibility of using genetic methods for stock identification as a harvest management tool in mixed fisheries involving populations from this region. Sufficient genetic differences are identified among regions, areas, and individual sites to justify collection of further data towards the ultimate application of genetic methods for mixed stock identification.

## MATERIALS AND METHODS

Sampling locations (fig. 1; table 1) ranged from Salmon Bay on Prince of Wales Island, Alaska in the north southward through the Quinault River of the Washington coast to two tributaries of the Columbia River, Lake Wenatchee and the Okanogan River. Tissues and gametes of adult fish were collected at the Fulton River in the Skeena River drainage. Fertilized eggs from these collections were raised on site through the eyed egg stage then transferred to the Stanley Hatchery in Idaho for subsequent rearing; these progeny were randomly sampled the following year. Liver, heart, muscle, and eye of adult fish and intact juvenile fish were frozen at the time of collection and remained frozen through preparation for electrophoresis.

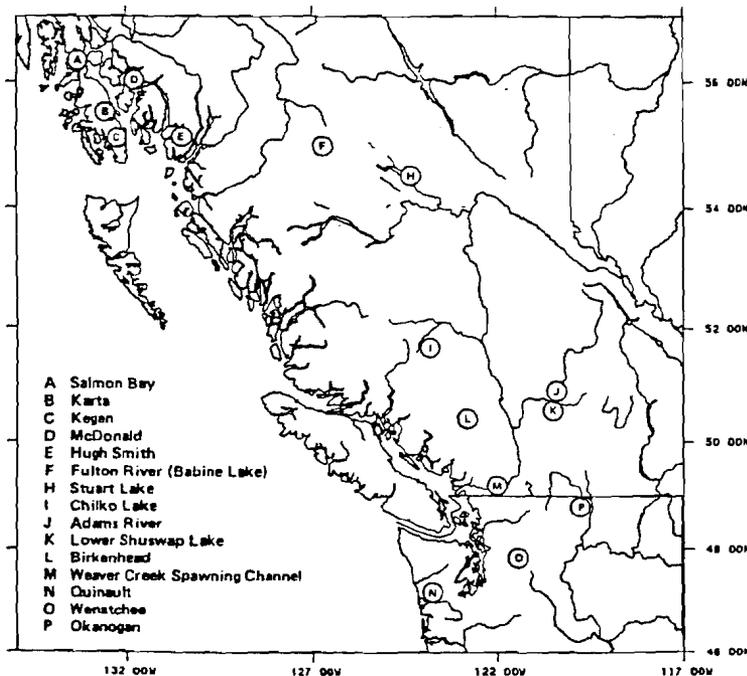


Figure 1. Southeastern range of sockeye salmon sampling locations.

Table 1--Collection data for sockeye salmon samples.

<u>Location</u>	<u>Maturity</u>	<u>Sampling date</u>	<u>No.</u>	<u>Sampling agency<sup>a</sup></u>
Salmon Bay Creek	Adult	07/26/82	69	NMFS, ADFG
Karta River	Adult	07/26/82	100	NMFS, ADGG
Kegan Creek	Adult	07/22/82	100	NMFS, ADFG
McDonald Lake	Adult	07/19/82	63	NMFS, ADFG
Hugh Smith Lake	Adult	07/19/82	100	NMFS, ADFG
Fulton River	Adult	09/03/80	99	NMFS, FO Canada, IDFG
Fulton River	Juvenile	09/06/81	61	NMFS, IDFG
Stuart Lake	Adult	08/04/82	100	IPSFC
Chilko Lake	Adult	10/05/82	100	IPSFC
Adams River	Adult	10/27/82	100	IPSFC
Lower Shuswap Lake	Adult	10/26/82	100	IPSFC
Birkenhead Lake	Adult	10/04/82	72	IPSFC
Weaver Spawning Ch.	Adult	10/29/82	100	IPSFC
Quinault River	Adult	11/01/82	51	NMFS, QIN
Wenatchee Lake	Adult	09/22-25/81	40	NMFS
Okanogan River	Adult	10/15/81	70	NMFS

a/ Sampling agency codes:

NMFS - National Marine Fisheries Service  
 ADFG - Alaska Department of Fish and Game  
 FO Canada - Canadian Department of Fisheries and Oceans  
 IDFG - Idaho Department of Fish and Game  
 IPSFC - International Pacific Salmon Fishery Commission  
 QIN - Quinault Indian Nation

Extraction procedures and electrophoretic methods followed May et al. (1979). Four buffer systems were used: RW - a discontinuous Tris, citric acid (gel pH 8.15), lithium hydroxide, boric acid (tray pH 8.0) buffer system (Ridgway et al. 1970); AC - an amine citrate buffer system, pH 6.5 (Clayton and Tretiak 1972); MF - a Tris, boric acid, EDTA buffer, pH 8.5 (Markert and Faulhaber 1965); TC - a Tris, citric acid buffer system, pH 7.3 (Shaw and Prasad 1970). Gels consisted of appropriate buffers and 13% hydrolyzed starch (Electrostarch,<sup>1/</sup> Madison, Wisconsin, Lot #307). Methods for visualizing enzyme activity followed descriptions of Siciliano and Shaw

(1976), Harris and Hopkinson (1976), Allendorf et al. (1977) or originated with this investigation (table 2). A system of nomenclature suggested by Allendorf and Utter (1979) was used to designate loci and alleles.

Tests for departures of genotypes from the expected binomial distribution (Hardy-Weinberg equilibrium) used a log likelihood ratio test (Sokal and Rohlf 1969) with df equaling the number of genotypes minus the number of alleles. Average heterozygosities were calculated on the basis of Hardy-Weinberg expectations of polymorphic loci averaged over all loci.

Table 2--Summary monomorphic and polymorphic loci examined.

Enzyme Name	Abbr- viate	Enzyme Comm.No.	Ref. <sup>a</sup>	No. of Loci	Locus	Buffer	Tissue	Observation <sup>b</sup>
Aconitase	ACO	4.2.1.3	H	3	1,2	AC	H	3
					3	AC	L	2
Adenosine deaminase	ADA	3.5.4.4	H	2	1	MF	M	1
					2	MF,AC	M	2
Adenylate kinase	AK	2.7.4.3	H	1	1	AC	M	1
Aldolase	ALD	4.1.2.13	H	4	1,2	TC	M,H	6
					3	AC	E	5
					4	MF	E	1
Aldono-lactonase	ALL		O	1	1	AC	M	1
Aspartate aminotransferase	AAT	2.6.11	H	3	1,2	AC	H,M	1
					3	RW	E	1
Creatine kinase	CK	2.7.3.2	H	3	1,2	RW	M	1
					3	RW	E	1
Enolase	ENO	4.2.1.11	S	1	1	AC	L	1
Esterase	EST	3.1.1.1	H	2	2,3	MF	M	1
Fructose diphosphatase	FDP		A+	1	1	MF	M	6
Fumarase	FUM	4.2.1.2	S	1	1	AC	L	1
N-Acetyl-B- galactosaminidase	GAM		O	2	1,2	AC	L	1
N-Acetyl-B- glucoseaminidase	HEX	3.2.1.30	H	1	1	AC	L	1
Glucose phosphate isomerase	GPI	5.3.1.9	H	3	1,2	RW	M	2
					3	RW	M	2
$\alpha$ -Glucosidase	GLU	3.2.1.20	H	2	1,2	AC	L	5
Glutamate dehydrogenase	GLUD	1.4.1.3	H+	1	1	AC	H	5
Glutamate pyruvate transaminase	GPT	2.6.1.2	H+	2	1	MF	H	2
					2	MF	M	4
Glutathione reductase	GR	1.6.4.2	H+	2	1	AC	L,H	5
					2	AC	L	6

Table 2--cont.

Enzyme Name	Abbr- viate	Enzyme Comm.No.	Ref. <sup>a</sup>	No. of Loci	Locus	Buffer	Tissue	Observation <sup>b</sup>
Glyceraldehydephosphate dehydrogenase	GAP	1.2.1.12	S	5	1,2	TC	M,H	6
					3,4	AC	H	6
					5,6	MF	E	1
Glycylleucine peptidase	GL	3.4.11.	H	2	1	MF	E	2
					2	MF	E	3
Glyoxylase I	GLO I	4.4.1.5	H	1	1	RW	L	1
Glyoxylase II	GLO II	3.1.2.6	H+	1	1	MF,AC	L	1
Guanine deaminase	GDA	3.5.4.3	H+	2	1	MF	L	6
					2	MF	H	6
Hexokinase	HK	2.7.1.1	H	1	2	AC	L	6
Inosine triphosphatase	ITP	3.6.1.19	H	2	1,2	RW	M	1
Isocitrate dehydrogenase	IDH	1.1.1.42	H	4	1	AC	M,H	1
					2	AC	H,M	5
					3	AC	L,E	1
					4	AC	L,E	3
Lactate dehydrogenase	LDH	1.1.1.27	H	5	1,2	AC,RW	M	1
					3	RW,AC	H	2
					4	AC	L	3
					5	AC	E	1
Leucylglycylglycine peptidase	LGG	3.4.11.	H	1	1	MF	E	1
Leucyltyrosine peptidase	LT	3.4.11	H	1	1	RW	M	2
Malate dehydrogenase	MDH	1.1.1.37	H	4	1,2	AC	L	2
					3,4	AC	M	2
Malic enzyme	ME	1.1.1.40	H	4	1	AC	M	2
					3	AC	M	1
					2,4	AC	L	1
Mannose phosphate isomerase	MPI	5.3.1.8	H	1	1	AC	H	1
$\alpha$ -Mannosidase	MAN	3.2.1.24	H+	1	1	AC	L	6

Table 2--cont.

Enzyme Name	Abbre- viate	Enzyme Comm.No.	Ref. <sup>a</sup>	No. of Loci	Locus	Buffer	Tissue	Observation <sup>b</sup>
Nucleoside phosphorylase	NP	2.4.2.1	H	2	1	MF	L	6
					2	AC, MF	H	6
Phenylalanylproline peptidase	PHP	3.4.11.	H	3	1,2	MF	H	1
					3	MF	E	5
Phosphoglucomutase	PGM	2.7.5.1	H	4	1	AC	H	4
					2	AC	H	4
					3,4	AC	H	5
6-Phosphogluconate dehydrogenase	6PG	1.1.1.44	H	1	1	AC	M	1
Phosphoglycerate kinase	PGK	2.7.2.3	H	2	1,2	AC	L	1
Pyruvate kinase	PK	2.7.1.40	H	2	1,2	TC	H	1
Superoxide dismutase	SOD	1.15.1.1	H	2	1	RW	L	2
					2	RW	H	1
Triokinase	TK		0	1	1	AC	M	1
Triose phosphate isomerase	TPI	5.3.1.1	H	3	1,2	MF	M	2
					3	MF	M	1

a/ H - Harris and Hopkinson (1976); A - Allendorf et al. (1977); S - Siciliano and Shaw (1976); 0 - procedure originated with this study; +- modified from original description.

b/ 1 - monomorphic, 2 - polymorphic, frequency at 0.01 or less over all populations, 3 - polymorphic, frequency at 0.05 or less over all populations, 4 - polymorphic, frequency greater than 0.05 over all populations, 5 - presumed genetic variation observed but model uncertain or resolution poor, 6 - presumed absence of genetic variation but model uncertain or resolution poor.

Two hierarchical analyses were used to identify the amount of structuring occurring among populations within the geographic range of this investigation. A nested gene diversity analysis which followed procedures described by Nei (1973) and Chakraborty (1980) examined the departure of overall allelic variation from that expected of a single panmictic unit, and measured that proportion of the total variation attributable to different hierarchical levels (outlined in fig. 2) according to the following model:

$$H_T = H_S + D_{GR} + D_{SA} + D_{AR} + D_{RT}$$

where  $H_T$  = the diversity of sites within areas,  
 $H_S$  = the diversity within sites,  
 $D_{GR}$  = the diversity over generations within regions,  
 $D_{SA}$  = the diversity of sites within areas,  
 $D_{AR}$  = the diversity of areas within regions, and  
 $D_{RT}$  = the diversity between regions.

**SOCKEYE**  
Gene diversity analysis

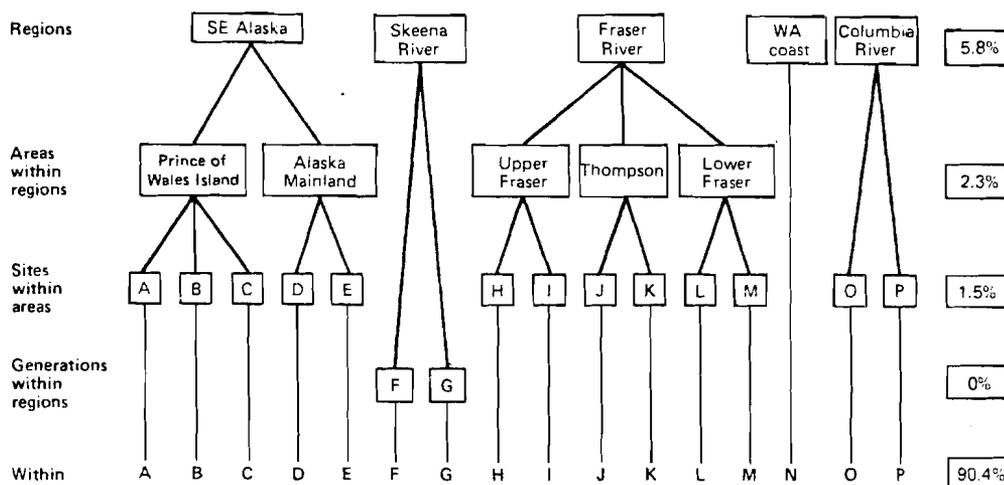


Figure 2. Hierarchy used in gene diversity analysis of sockeye salmon. Percentages in right column are proportions of total genetic variation occurring at different levels.

The same hierarchy was used in a nested contingency table analysis which measured heterogeneity among allelic frequencies at different levels by a log likelihood (G) statistic with df equaling allelic classes minus one times sites, areas, or regions minus one. Significance levels of the tests were modified by dividing the significance level by three (the number of loci used in this analysis) to account for the increase in type I error when multiple tests of the same comparison are made (Cooper 1968). Total heterogeneity was partitioned into within and among area and region components in a manner analogous to analysis of variance (Smouse and Ward 1978). This analysis was restricted to those three loci having mean allelic frequencies of the common allele less than 0.95 to avoid low expected frequencies; rare alleles were pooled with more frequent allelic classes for GPT-2 and PGM-2 loci to achieve adequate expected frequencies. A standardized measure calculated by dividing the log likelihood statistic summed over all loci by the summed degrees of freedom was used to rank the heterogeneity of individual components of the hierarchy in an overall comparison (Lancaster 1969).

Genetic identity values (Nei 1972) were calculated for each pair of collections based on the 12 polymorphic and 38 monomorphic loci where data were available for all collections. A dendrogram was constructed using the unweighted pair-group method (Sneath and Sokal 1973) based on a similarity matrix of these paired values.

#### ALLELIC VARIANTS

The relatively low frequency of polymorphisms reported in previous electrophoretic studies of sockeye salmon (Utter et al. 1973; Allendorf and Utter 1979) led to an extensive search for new

polymorphisms in the preliminary phases of this investigation. A summary of these results is outlined below, although most technical details will be presented in another paper. This screening resulted in the detection of 90 putative loci (table 2) of which 68 were adequately resolved for population studies.

Allelic variation (observed in this study) which has been previously reported for sockeye salmon includes GPT-2; LDH-4; MDH-1,2; MDH-3,4; GPI-1,2; GPI-3, PGM-2, and IDH-4 (Hodgins et al. 1969; Utter and Hodgins 1970; Grant et al. 1980; Wishard 1981). Data for five of the additional polymorphic systems involved in subsequent analyses in this paper do not warrant detailed description here. Each of these systems is expressed in a codominant fashion; variants include typically monomeric phenotypes for ACO-3 and ADA-2, dimeric variation for GL-2, and tetrameric expression for LDH-3 and ME-1.

One newly observed polymorphic system that requires further description is PGM-1. This system is identified by the presence or absence of the most cathodal bands expressed in heart extracts on an AC buffer system (fig. 3). This expression is presumably the reflection of a null or inactive allele with the absence of banding being its homozygous form. The banded phenotype

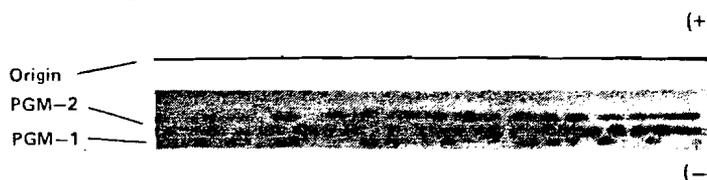


Figure 3. PGM-1 and PGM-2 phenotypes from heart extracts of 50 sockeye salmon from the Adams River, B.C.

shows no consistent differences of intensity as a reflection of dosage, thus Hardy-Weinberg estimates (that are dependent on the expression of all three genotypes) cannot be obtained. Rather, Hardy-Weinberg distribution had to be assumed; the allelic frequency of the null allele ( $\text{PGM-1(N)}$ ) was therefore estimated by the square root of its phenotypic frequency, and the estimated frequency of the active allele ( $\text{PGM-1(100)}$ ) was one minus the frequency of  $\text{PGM-1(N)}$ . The estimation procedure is regarded as valid in most instances in this study because of the good fit (reported below) in codominantly expressed loci for Hardy-Weinberg expectations, coupled with the presence of null phenotypes in substantial numbers in most populations.

Null phenotypes were absent only in the Quinault population, and the assumption of fixation of the active allele probably overestimates its actual frequency. Under Hardy-Weinberg conditions, the frequency of the active alleles would have to be 0.85 with a sample size of 50 before a null phenotype would be expected. Nevertheless, it is evident (table 3) that even 0.85 is still substantially higher than the next highest value for the active allele of  $\text{PGM-1}$ .

#### POPULATION STRUCTURING

Pertinent features of genetic structuring among

the areas of this study are revealed by examining the data in different ways. An initial step involves looking at the genotypic and allelic frequencies themselves. Only four of the genotypic frequencies among 9 loci and 16 collections listed in table 3 deviated significantly from those expected in random mating populations (i.e., Hardy-Weinberg equilibrium). Since each of these deviations was of borderline significance and such a number would be expected at the 5% level of significance, these data are consistent with the overall data set being drawn from isolated and random mating populations. Average heterozygosities among these populations (table 3) displayed no apparent geographic pattern and ranged from 0.0412 in the Salmon Bay collection to 0.0211 in the Quinault River sample.

Allelic frequencies for the nine most polymorphic loci (table 3) reveal a broad diversity among collections. The range exceeds 0.50 for  $\text{GPT-2}$  and  $\text{PGM-1}$ , and 0.25 for  $\text{PGM-2}$  and  $\text{GL-2}$ . The variation for these four loci appears to be scattered in a somewhat random fashion throughout the sampling range. However, it is of interest that the substantial extremes of frequency of the common allele occur in the Quinault sample for three of these systems. Variation for  $\text{LDH-4}$  predominates in more northerly collections (consistent with previous studies, e.g., Utter et al. 1974),

Table 3--Allelic frequencies<sup>a</sup> for the nine most polymorphic loci, and average heterozygosity values based on 12 polymorphic and 38 monomorphic loci.

Region site	Locus (allele and number of fish)																		
	GPT-2					PGM-1		PGM-2				GL-2		LDH4					
	(100)	(91)	(108)	(86)	(95)	N	(100)	N	(-100)	(-77)	(-93)	(-20)	N	(100)	N	(100)	(115)	(85)	N
Southeastern Alaska	.200	.323	0	0	.478	45	.232	69	.841	.150	0	.009	69	.888	67	.993	.007	0	69
Salmon Bay																			
Karta	.716	.011	0	0	.273	44	.200	50	.760	.240	0	0	50	.970	50	.950	.050	0	50
Kegan	.631	.303	0	0	.066	99	.360	100	.545	.450	.005	0	100	.969	98	.995	.005	0	100
McDonald	.711	.244	0	0	.045	45	.080	50	.860	.130	0	.010	50	.920	50	.980	.020	0	50
Hugh Smith	.580	.352	.035	0	.035	44*	.190	50	.850	.130	0	.020	50	.925	50	.910	.090	0	50
Skeena River																			
Fulton (adult)	.393	.539	.017	0	.051	89	.179	98	.667	.182	.025	.126	99	.935	98	.980	0	.020	99
Fulton (juvenile)	.370	.532	.009	0	.089	108	.140	111	.674	.229	.097	0	109*	.904	109	.995	.005	0	111
Fraser River																			
Stuart	.777	.162	0	0	.061	99	.571	98	.730	.270	0	0	98	1.000	99	1.000	0	0	100
Ghilko	.467	.383	0	0	.150	90	.500	100	.850	.150	0	0	100	1.000	94	1.000	0	0	100
Adams	.781	.198	0	0	.021	96	.654	100	.850	.150	0	0	100	1.000	100	.990	.010	0	100
Lower Shuswap	.637	.019	0	.093	.080	91	.615	100	.840	.160	0	0	100	1.000	100	1.000	0	0	100

Table 3--cont.

Birkenhead	.310	.563	0	0	.127	71	.253	72	.826	.174	0	0	72	1.000	72	1.000	0	0	72
Weaver Channel	.457	.495	0	0	.048	94*	.151	99	.800	.195	.005	0	99	1.000	97	1.000	0	0	100
Washington Coast																			
Quinault	.690	.270	0	0	.040	50	1.000	50	.960	.040	0	0	50	.740	48	1.000	0	0	50
Columbia River																			
Wenatchee	.813	.187	0	0	0	40	.423	34	.705	.295	0	0	39	1.000	40	.986	.014	0	35
Okanogan	.590	.285	.118	.007	0	72	.493	71	.782	.218	0	0	71	1.000	61	1.000	0	0	72

Region site	GPI-2				IDH-4				LDH-3		ADA-2		Average Hetero.
	(100)	(132)	(143)	N	(100)	(162)	(72)	N	(100)	N	(100)	N	
Southeastern Alaska													
Salmon Bay	1.000	0	0	69	.978	0	.022	69	.920	69	1.000	69	.0412
Karta	1.000	0	0	50	1.000	0	0	50	1.000	49	1.000	50	.0285
Kegan	1.000	0	0	100	1.000	0	0	100	1.000	100	1.000	100	.0381
McDonald	1.000	0	0	50	1.000	0	0	50	.990	50	1.000	50	.0213
Hugh Smith	1.000	0	0	50	1.000	0	0	50	.990	50	1.000	50	.0237
Skeena River													
Fulton (adult)	1.000	0	0	99	.965	0	.035	99	1.000	99	.985	99	.0386
Fulton (juvenile)	1.000	0	0	111	1.000	0	0	111	1.000	111	-	-	.0368
Fraser River													
Stuart	.980	.020	0	100	1.000	0	0	100	1.000	98	.970	100	.0285
Chilko	.990	.010	0	100	.995	.005	0	100	1.000	100	.955	100	.0322
Adams	.960	.040	0	100	1.000	0	0	100	1.000	100	.980	100	.0238
Lower Shuswap	.990	.010	0	100	.990	.010	0	100	1.000	100	.908	100	.0278
Bihkenhead	.777	.216	.007	72*	.889	.111	0	72	1.000	72	.990	72	.0399
Weaver Channel	1.000	0	0	98	1.000	0	0	100	1.000	100	.980	98	.0281
Washington Coast													
Quinault	1.000	0	0	50	1.000	0	0	50	1.000	50	1.000	50	.0211
Columbia River													
Wenatchee	1.000	0	0	40	.986	.014	0	35	1.000	40	-	-	.0271
Okanogan	.993	.009	0	70	.986	.014	0	72	1.000	50	-	-	.0326

a/ Only frequency of common allele given for two allele systems.

\* Significant deviation of genotype frequencies from Hardy-Weinberg expectations  $0.05 > p > 0.01$ .

- No data.

and the Fraser and Fulton River samples are distinguishable from other collections by the presence of ADA-2 variation. The Fraser River samples are further distinguished by the presence of GPI-2 variation in five of the six areas, contrasted with the virtual absence of variation at this locus in other drainages. The allelic frequencies of the juveniles of the Fulton River and those of the adult fish giving rise to this collection were generally very similar; however, it is notable that uncommon alleles were possibly lost in the juveniles for three loci: PGM-2, LDH-4, and IDH-4.

The hierarchical analyses give additional insights into population structuring. The gene diversity analysis (table 4) expectedly indicated that the three most polymorphic loci (GPT-2, PGM-1, PGM-2) contributed the greatest share of the total genetic diversity. However, other loci made substantial contributions to the actual structuring at different hierarchical levels. Approximately 90% of the total genetic variation occurred within

sites (all of the variation would occur within sites in the absence of any structuring). The locus making by far the greatest contribution to structuring was PGM-1, where only 75% of its variation was common to all populations. Only a negligible amount of the total variation was attributable to diversity between generations within regions, a result that would be expected from a single observation based on comparisons of parents and their progeny. The LDH-3 locus was the greatest contributor to the 1.5% of the total diversity observed among sites within areas, confirming previous observations made from table 3. The strong influence of the PGM-1 locus in structuring was apparent in both the 2.3% of the total diversity among areas within regions and the 5.8% of the diversity among regions. The GL-2 locus also contributed substantially to the diversity in the latter hierarchical level. Approximately 10% of the total diversity attributable to structuring lies within a moderate range of population structuring based on similar measurements from diverse organisms (Hartl 1980).

Table 4--Distribution of electrophoretically detectable gene diversity among 16 collections. The average refers to 12 polymorphic and 38 monomorphic loci.

Locus	Absolute gene diversity		Relative Gene Diversity				
	Total (H <sub>T</sub> )	Within sites (H <sub>S</sub> )	Within sites	Between years within regions	Between sites within areas	Between areas within regions	Between regions
ACO-3	0.0110	0.0109	0.9846	0.0000	0.0022	0.0077	0.0056
IDH-4	0.0259	0.0244	0.9441	0.0015	0.0294	0.0189	0.0061
PGM-1	0.4700	0.3541	0.7533	0.0002	0.0081	0.0601	0.1782
GPI-1	0.0191	0.0177	0.9272	0.0000	0.0416	0.0177	0.0149
GPI-3	0.0086	0.0084	0.9775	0.0000	0.0090	0.0070	0.0064
LDH-4	0.0256	0.0245	0.9562	0.0003	0.0182	0.0074	0.0180
MDH-1,2	0.0057	0.0056	0.9858	0.0000	0.0028	0.0076	0.0038
LDH-3	0.0124	0.0117	0.9399	0.0000	0.0430	0.0034	0.0138
PGM-2	0.3854	0.3760	0.9755	0.0000	0.0093	0.0043	0.0109
GL-2	0.0921	0.0876	0.9510	0.0003	0.0025	0.0006	0.0456
GPT-2	0.6749	0.6430	0.9528	0.0000	0.0237	0.0124	0.0111
ME-1	0.0031	0.0030	0.9883	0.0000	0.0063	0.0042	0.0013
Average	0.0347	0.0313	0.9037	0.0001	0.0151	0.0228	0.0582
Standard error	0.0178	0.0162	0.0410	0.0001	0.0043	0.0122	0.0405

A similar hierarchy used in contingency table analyses of allelic distributions of the three most polymorphic loci (table 5) gave a more focused view of specific aspects of population structuring. It is useful to first examine the standardized measures (i.e., the summed G statistic divided by df) in order to rank the relative amounts of heterogeneity at different hierarchical levels. Considerable heterogeneity was observed at all levels with by far the greatest being in the comparison of the Quinault sample with the combined regional data. It is notable that similar levels of heterogeneity were observed for

comparisons among regions, among areas within regions, and among sites within one of the areas (Prince of Wales Island). Among the contingency tests, only the two samples from mainland Alaska failed to reflect any significant heterogeneity either at individual loci or their summed G values. Highly significant variation was found for all other comparisons among locations for the GPT locus and for the summed G values. The PGM loci were more homogeneous within areas, but highly significant differences were consistently found at higher levels.

Table 5--Contingency table analysis of alleles of the three most polymorphic loci. Major elements of hierarchy are in figure 2.

Source of Variation	df <sup>a</sup>	GPT-2		PGM-1		PGM-2		Sum		Standardized Measure (G/df)
		G <sup>b</sup>	df	G	df	G	df	G		
Regions + Quinault	2	14.9**	1	201.4***	1	28.5***	4	259.7***	64.9	
Among regions	6	101.5***	3	194.0***	3	39.0***	12	334.5***	27.8	
Within regions	24	368.2***	12	304.2***	12	87.7***	48	1128.3***	23.5	
Columbia River	2	21.6***	1	0.8	1	1.6	4	24.0***	6.0	
Fulton River	2	1.2	1	0.9	1	0.0	4	2.1	0.5	
Total areas	20	345.4***	10	302.5***	10	86.1***	40	734.0***	18.3	
Among areas within regions	8	140.1***	4	280.7***	4	39.7***	16	460.5***	28.7	
Within areas	12	205.3***	6	21.8***	6	46.4***	24	273.5***	11.3	
Prince of Wales Island	4	123.2***	2	11.1**	2	37.3***	8	171.6***	21.4	
Alaska Mainland	2	3.4	1	4.3	1	0.0	4	7.7	1.9	
Upper Fraser River	2	39.4***	1	1.8	1	8.7**	4	49.9***	12.5	
Thompson River	2	27.9***	1	0.7	1	0.1	4	28.7***	7.2	
Lower Fraser River	2	11.4**	1	3.6	1	0.3	4	15.3**	3.9	
Total	32	484.6***	16	699.6***	16	155.2***	64	1339.4***	20.9	

a/ Degrees of freedom.

b/ G statistic.

\*\* Significant at the 0.01 level.

\*\*\* Significant at the 0.001 level.

A final analysis involved a matrix (table 6) of identity values (ranging from 0 to 1) based on paired comparisons of allelic frequencies for all collections. A phenogram (fig. 4) constructed from this matrix depicts groupings among collections based on varying degrees of genetic identity. The most distinct feature of the phenogram is the clear divergence of the Quinault sample from all other collections. A second feature is

the formation of two clusters at a considerably higher level of identity. Each of these clusters has some geographic consistency. Cluster A contains four of the five southeastern Alaska collections plus those of the Fulton River and lower Fraser River. Cluster B contains the collections from the upper Fraser River, Thompson River, and Columbia River, plus the Kegan River collection of southeastern Alaska as an apparent outlier.

Table 6—Matrix of genetic similarity among collections based on 12 polymorphic and 38 monomorphic loci.

Population																
Wenatchee	1	1.000														
Okanogan	2	0.997														
Quinault	3	0.974	0.979													
Weaver Chan.	4	0.989	0.991	0.955												
Birkenhead	5	0.985	0.990	0.958	0.997											
L. Shuswap	6	0.995	0.997	0.985	0.983	0.982										
Adams	7	0.996	0.996	0.989	0.980	0.979	0.999									
Chilko	8	0.994	0.998	0.981	0.993	0.954	0.997	0.994								
Stuart	9	0.999	0.998	0.983	0.985	0.982	0.998	0.999	0.995							
Fulton (juvenile)	10	0.985	0.987	0.946	0.998	0.996	0.974	0.972	0.989	0.979						
Fulton (adult)	11	0.984	0.988	0.946	0.998	0.996	0.977	0.973	0.990	0.979	0.999					
Kegan	12	0.997	0.995	0.966	0.992	0.989	0.990	0.984	0.993	0.995	0.991	0.991				
Karta	13	0.994	0.990	0.958	0.990	0.985	0.986	0.985	0.989	0.990	0.986	0.985	0.993			
Salmon Bay	14	0.978	0.983	0.955	0.991	0.992	0.977	0.973	0.990	0.978	0.991	0.989	0.984	0.988		
Hugh Smith	15	0.992	0.993	0.962	0.998	0.994	0.986	0.985	0.993	0.988	0.995	0.996	0.992	0.994	0.990	
McDonald	16	0.991	0.989	0.953	0.996	0.990	0.981	0.981	0.988	0.985	0.992	0.992	0.990	0.996	0.986	0.998

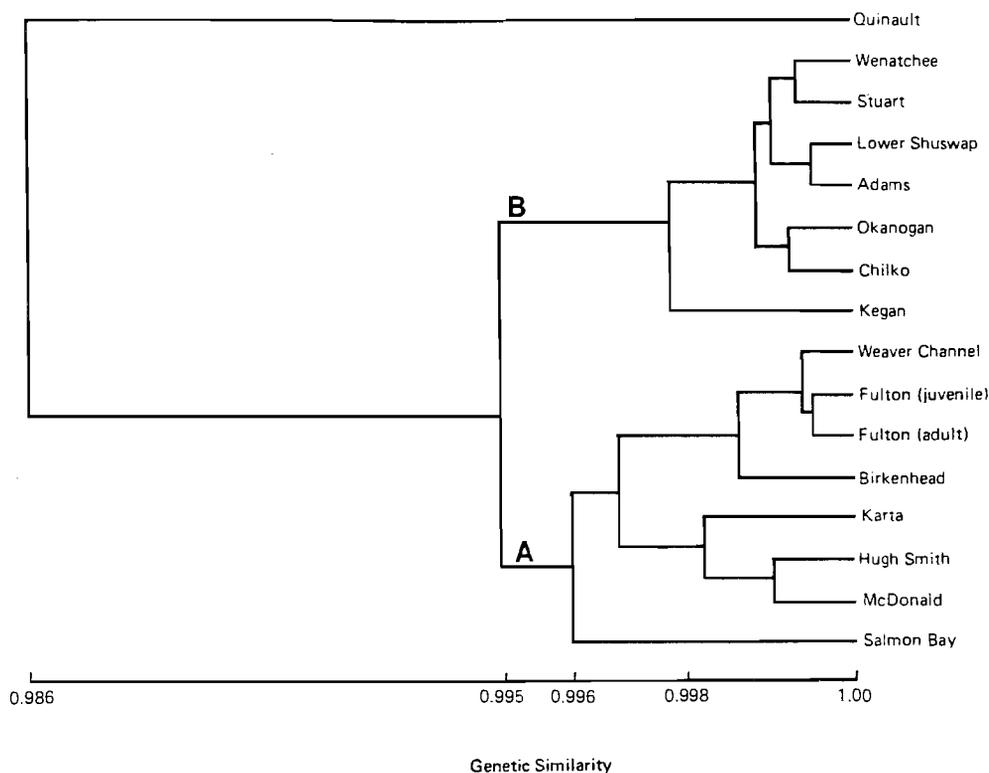


Figure 4. Phenogram based on measurements of genetic similarity (taken from data of Table 6).

## RELATIONSHIPS AMONG POPULATIONS

The analyses and descriptions of the preceding section give some insight into possible relationships among populations. The clearest feature is the distinctive nature of the Quinault collection. The possibility that much of this difference may not be a reflection of long-term isolation from other populations of this study must first be considered. The extreme frequencies of common alleles, reduced numbers of rare alleles, and the lowest average heterozygosity of all populations examined are all suggestive of one or more population "bottlenecks" or a chronically reduced effective population number (Hartl 1980). Similar attributes coupled with considerable genetic divergences from recent ancestral stock have been identified in some salmonid hatchery stocks (Allendorf and Phelps 1980; Ryman and Stahl 1980). Thus the effect of the hatchery operation on the Quinault sockeye salmon run since 1915 must be examined (Quinault Indian Nation, 1981). This operation has been supplemental rather than primary in the maintenance of the run, and the sample we obtained was from the wild segment (L. Gilbertson, biologist, Quinault Indian Nation, pers. commun.). Furthermore, the overall run size has remained consistently above a level that would constitute a bottleneck, although the current run averages about one quarter the number of historical runs (Quinault Indian Nation 1981). Thus, bottlenecks during historical times are excluded as an influence in this divergence, although earlier events of this type remain possibilities. Whatever the cause of this divergence, the

Quinault population is clearly distinct from other populations in this study.

On the other hand, the loss of alleles noted above in the juvenile collection of the Fulton River is most likely related to a founder event. A full representation of alleles was expected because the juveniles were descended only from the individuals of the adult collection. However, considerable mortality occurred during incubation and subsequent rearing to the time of sampling (R. Wahle, Environmental and Technical Services Division, NOAA, NMFS, Portland, Oregon, pers. commun.). It is apparent that the juvenile samples are not fully representative of the adult collections, and that mortality may have differentially affected specific crosses rather than acting randomly on all crosses.

Lines of distinction exclusive of the Quinault collection are less clear among the remaining populations. Utter et al. (1980) suggested that sockeye salmon populations of the Columbia River are derived from a lineage that is distinct from more northward populations. But the clustering in figure 4 of the Columbia River collections with those of the upper Fraser River does not support such a clear separation. Rather, a comparison of Clusters A and B suggests that the possible populating of the lower Fraser River with separate ancestral stocks from those of the upper Fraser River and the Columbia River. Such suggestions must await clarification with additional data from other loci, populations, and perhaps, complementary methods. Regardless of ancestral relation-

ships, the overall distinctness of the Fraser River collections from one another should encourage expanded use of protein data in identifying, monitoring, and managing the sockeye salmon populations of this drainage.

It is not surprising that a more unified picture of population structuring failed to materialize. The dependence of both juvenile and adult sockeye salmon on a lake in the environment coupled with strong homing resulting from long freshwater life histories (Okazaki 1982) has apparently resulted in abrupt discontinuities among populations of sockeye salmon in this region. In the absence of equalizing effects of migration, the isolating forces of drift and direct or indirect selection are magnified.

#### POTENTIAL FOR APPLICATION IN MIXED STOCK FISHERIES

Three conditions required for effective application of data in the analysis of mixed stock fisheries include (1) identifying a suitable number of electrophoretic variants having a genetic basis, (2) obtaining a representative sampling of populations contributing to fisheries of interest, and (3) identifying sufficient differences in frequencies of electrophoretic variants among pertinent management units to permit estimations of their contributions with reasonable precision and practical sample sizes (Milner 1977; Milner et al. 1980, 1981; Miller et al. 1983).

Condition (1) has clearly been met. The average heterozygosity of sockeye salmon remains in the lower range reported among salmonid species (Allendorf and Utter 1979). Nevertheless, the total number of polymorphic loci identified in this and other studies of sockeye salmon provides an ample amount of genetic variability to adequately characterize the breeding structure of the species.

Condition (2) is partially met. The geographic range of sampling is probably inclusive of most populations contributing to sockeye fisheries from the Alaska-Canada border southward (French et al. 1976). However, the data points are minimal and determining whether or not the present data set is representative of the total populations of this region must await the sampling from additional populations of southeastern Alaska, Canada, and the Pacific Northwest.

Possible variations of genotypic and allelic frequencies between year-classes and generations must also be considered under condition (2). Stable frequencies such as those observed in some populations of steelhead trout (Utter et al. 1980) and sockeye salmon (Grant et al. 1980) would minimize the need for resampling populations once an adequate data base had been established. However, some instability among year-classes (amid considerable stability) has been observed in chinook salmon populations of the Columbia River (Milner et al. 1980), and attributed primarily to extensive transplantations and resultant strayings. Regardless of cause, it is important to measure stability over year-classes and generations.

Condition (3) has been fulfilled. The data indicate sufficient structuring of populations to effectively estimate proportions of contributing stocks in mixed populations given an adequate data base. Realistic estimates of stock composition were obtained in a more geographically restricted study of sockeye salmon in Cook Inlet, Alaska (Grant et al. 1980) where detected levels of structuring were only half of those identified in this investigation (Ryman 1983). Genetic estimates of stock composition that were consistent with independent sets of data have also been made in mixed stocks of chum salmon (Wishard 1981; Beacham et al. 1982, Seeb et al. in press) where structuring is much reduced from that of sockeye salmon (Okazaki 1982).

The most definitive comparison at this time regarding the potential use of allelic data for mixed stock analyses of sockeye salmon populations in this study is with chinook salmon data collected over a similar geographic range. All three of the above conditions are met with the chinook salmon data (Milner et al. 1980, 1981), and the most effective use of data of the genetic method has been an analysis of chinook salmon fisheries off the coast of Washington in May 1982 (Miller et al. 1983). This study estimated the stock composition of this fishery using genetic data from 14 polymorphic loci from approximately 2,000 fish, and a data base for these loci from 28 stocks of the Columbia River and 52 stocks beyond the Columbia River from California through British Columbia. An estimated 77% of the total catch was from fall chinook salmon populations of the lower Columbia River and Bonneville Pool areas with a standard deviation on this estimate of less than 1%. Analyses by northern and southern catch areas indicated a 10-fold increase in estimated harvest of British Columbia and Puget Sound stocks in the northern area. The amount of structuring of chinook salmon populations from British Columbia through California is similar to that identified in this study (Milner, unpublished data).

Thus the genetic method for stock identification is on the threshold of application for sockeye salmon populations within the geographic region of this study. The existing data base could be extended by including allelic frequencies being collected concurrently by Canadian workers (R. Withler, Canadian Department of Fisheries and Oceans, Nanaimo, B.C., pers. commun.) from additional sites within British Columbia. Such inclusion, coupled with another season's collection of duplicated and new sites will provide a usable data base for mixed fishery analyses in this region. Subsequent attention could then be focused on actual estimations of stock compositions of mixed stock fisheries, although the data base would continually be updated as new variants were detected, additional populations sampled, and other year-classes and generations examined. These estimates could be implemented in a timely manner (e.g., within 24 hours of laboratory receipt of fishery samples; Milner et al. 1981) and thus provide a powerful and unique basis for in-season management. In addition, the cumulative data over a period of years would yield a new and detailed understanding of the temporal and spatial distribution in the North Pacific Ocean of sockeye salmon populations from this area.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of their colleagues George Milner, Carol Ranck, and David Teel. Their assistance in all phases of this study including sample collection, data gathering, and data analysis was essential to its successful and timely completion.

## LITERATURE CITED

- Allendorf, F.W., N. Mitchell, N. Ryman, and G. Stahl. 1977. Isozyme loci in brown trout (*Salmo trutta*): Detection and interpretation from population data. *Hereditas* 86:179-190.
- Allendorf, F.W., and S. Phelps. 1980. Electrophoretic detection of the loss of genetic variation in a hatchery stock of cutthroat trout (*Salmo clarki*). *Transactions of the American Fisheries Society* 109:537-543.
- Allendorf, F.W., and F.M. Utter. 1979. Population genetics. Pages 407-454 in W.S. Hoar, D.S. Randall, and J.R. Brett, editors. *Fish Physiology*. Vol. 8. Academic Press, New York, New York, USA.
- Anas, R.E., and S. Murai. 1969. Use of scale characters and a discriminant function for classifying sockeye salmon (*Oncorhynchus nerka*) by continent of origin. *International North Pacific Fisheries Commission Bulletin* 26:157-192.
- Beacham, T.D., R.E. Withler, A.P. Gould, and A.P. Stefanson. 1982. Stock identification of chum salmon in southern British Columbia through variability in electrophoretic, morphometric, and meristic characters. Unpub. Rept., Canadian Department of Fisheries and Oceans, Nanaimo, B.C. 11 pp.
- Chakraborty, R. 1980. Gene diversity analysis in nested subdivided populations. *Genetics* 96:721-726.
- Clayton, J.W., and D.N. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. *Journal of Fisheries Research Board of Canada* 29:1169-1172.
- Cooper, D.W. 1968. The significance level in multiple tests made simultaneously. *Heredity* 23:614-617.
- French, R., H. Bilton, M. Osako, and A. Hartt. 1976. Distribution and origin of sockeye salmon (*Oncorhynchus nerka*) in offshore waters of the North Pacific Ocean. *International North Pacific Fisheries Bulletin* 34.
- Grant, W.S., G.B. Milner, P. Krasnowski, and F.M. Utter. 1980. Use of biochemical genetic variants for identification of sockeye salmon (*Oncorhynchus nerka*) stocks in Cook Inlet, Alaska. *Canadian Journal of Fisheries and Aquatic Science* 37:1236-1247.
- Harris, H., and D.A. Hopkinson. 1976. *Handbook of enzyme electrophoresis in human genetics*. American Elsevier. New York, New York, USA.
- Hartl, D.L. 1980. *Principles of population genetics*. Sinauer Associates Inc., Sunderland, MA. 488 p.
- Hartt, A.C. 1962. Movement of salmon in the North Pacific Ocean and Bering Sea as determined by tagging 1956-1958. *International North Pacific Fisheries Commission Bulletin* 6:157 p.
- Hodgins, H.O. 1972. Serological and biochemical studies in racial identification of fishes. Pages 199-208 in R.C. Simon, and P.A. Larkin, editors. *The stock concept in Pacific Salmon*. H.R. MacMillan lectures in fisheries. University of British Columbia.
- Hodgins, H.O., W.E. Ames, and F.M. Utter. 1969. Variants of lactate dehydrogenase isozymes in sera of sockeye salmon (*Oncorhynchus nerka*) *Journal of Fisheries Research Board of Canada* 26:15-19.
- Lancaster, H.O. 1969. *The chi square distribution*. John Wiley and Sons, New York, New York, USA. 356 p.
- Markert, C.L., and I. Faulhaber. 1965. Lactate dehydrogenase isozyme patterns of fish. *Journal of Experimental Zoology* 156:319-332.
- May, B., J.E. Wright, and M. Stoneking. 1979. Joint segregation of biochemical loci in salmonids: Results from experiments with *Salvelinus* and review of the literature in other species. *Journal of Fisheries Research Board of Canada* 36:1114-1128.
- Miller, M., P. Pattillo, G. Milner, and D. Teel. 1983. Analysis of chinook stock composition in the May 1982 troll fishery off the Washington coast: an application of the genetic stock identification method. Technical Report #74 State of Washington Department of Fisheries. 27 p.
- Milner, G.B. 1977. Biochemical genetic variation: Its use in mixed fishery analysis. Pages 10-15 in T.J. Hassler, and R.R. VanKirk, editors. *Genetic implications of steelhead management*. Special Report 77-1, California Cooperative Fisheries Research Unit.
- Milner, G.B., D.J. Teel, and F.M. Utter. 1980. Columbia River stock identification study. Annual report of research (FY79). NOAA; NWAFC, Seattle, Washington, USA.
- Milner, G.B., D.J. Teel, F.M. Utter and C.L. Burley. 1981. Columbia River stock identification study: validation of genetic method. Annual report of research (FY80), NOAA, NWAFC, Seattle, Washington, USA.

- Mosher, K.  
1963. Racial analysis of red salmon by means of scales. International North Pacific Fisheries Commission Bulletin 11:31-56.
- Nei, M.  
1972. Genetic distance between populations. American Naturalist 106:283-292.
- Nei, M.  
1973. Analysis of gene diversity in subdivided populations. Proceeding National Academy of Science 70:3321-3323.
- Okazaki, T.  
1982. Genetic study of population structure in chum salmon (*Oncorhynchus keta*). Bull. 19.(25-116) Far Seas Fisheries Research Laboratory, Shimizu, 424, Japan.
- Quinault Indian Nation.  
1981. 1981 pre-season management report for spring and summer fisheries on the Quinault and Queets Rivers. Prepared by Harvest Management Section of Quinault Fisheries Division. 82 p.
- Ridgway, G.J., G.W. Klontz, and C. Matsumoto.  
1962. Intraspecific differences in serum antigens of red salmon demonstrated by immunochemical methods. International North Pacific Fisheries Commission Bulletin 8:14 p.
- Ridgway, G.J., S.W. Sherburne, and R.D. Lewis.  
1970. Polymorphism in the esterases of Atlantic herring. Transactions of the American Fisheries Society 99:147-151.
- Ryman, N.  
1983. Patterns of distribution of biochemical genetic variation in salmonids: differences between species. Aquaculture.
- Ryman, N., and G. Stahl.  
1980. Genetic changes in hatchery stocks of brown trout (*Salmo trutta*). Canadian Journal of Fisheries and Aquatic Sciences 37:82-87.
- Seeb, J.E., L.W. Wishard, and F.M. Utter.  
In press. Biological and management implications of marking a stock of chum salmon through manipulation of allelic frequencies. Transactions of the American Fisheries Society.
- Shaw, C.R., and R. Prasad.  
1970. Starch gel electrophoresis of enzymes - a compilation of recipes. Biochemical Genetics 4:297-320.
- Siciliano, M.R., and C.R. Shaw.  
1976. Separation and visualization of enzymes on gels. Pages 185-209 in I. Smith, editor. Chromatographic and electrophoretic techniques, Vol. 2. Zone electrophoresis. Heinemann, London England.
- Smouse, P.E., and R.H. Ward.  
1978. A comparison of the genetic infrastructure of the Ye'cuana and the Yanomama: a likelihood analysis of genotypic variation among populations. Genetics 88:611-631.
- Sneath, P.H.A., and R.R. Sokal.  
1973. Numerical taxonomy. W.H. Freeman, San Francisco, California, USA.
- Sokal, R.R. and F.J. Rohlf.  
1969. Biometry. W.H. Freeman, San Francisco, California, USA.
- Utter, F.M., and H.O. Hodgins.  
1970. Phosphoglucosomutase polymorphism in sockeye salmon. Comparative Biochemistry and Physiology 36:195-199.
- Utter, F.M., F.W. Allendorf, and H.O. Hodgins.  
1973. Genetic variability and relationships in Pacific salmon and related trout based on protein variations. Systematic Zoology 22:257-270.
- Utter, F.M., H.O. Hodgins, and F.W. Allendorf.  
1974. Biochemical genetic studies of fishes: potentialities and limitations. Pages 213-237 in D. Malins, editor. Biochemical and biophysical perspectives in marine biology. Vol. 1. Academic Press, San Francisco, California, USA.
- Utter, F.W., D. Campton, S. Grant, G. Milner, J. Seeb, and L. Wishard. 1980. Population structures of indigenous salmonid species of the Pacific Northwest. In W.J. McNeil, and D.C. Himsworth, editors. Salmonid ecosystems of the North Pacific. Oregon State University Press, Corvallis, Oregon, USA.
- Wishard, L.N.  
1981. Stock identification of Pacific salmon in western Washington using biochemical genetics. Final report to Washington Department of Fisheries Service Contract #1176 and 1276. 116 p.

1/ Reference to trade names does not imply endorsement by the National Marine Fisheries Service.