Tests of Genetic Stock Identification Using Coded Wire Tagged Fish

Jon Brodziak[1]

Institute of Theoretical Dynamics, University of California, Davis, CA 95616, USA

Boyd Bentley, Devin Bartley,[2] and Graham A. E. Gall

Department of Animal Science, University of California, Davis, CA 95616, USA

Richard Gomulkiewicz[3]

Institute of Theoretical Dynamics, University of California, Davis, CA 95616, USA

and Marc Mangel[4]

Section of Zoology and Center for Population Biology, University of California, Davis, CA 95616, USA

Genetic Stock Identification (GSI) uses allozyme variation to determine the composition of mixed-stock fisheries. The GSI method was tested using real fishery data. We report, for the first time in the primary literature, results of tests of GSI in which source stock and ocean-caught mixture samples were separately obtained and the mixture composition was known exactly because the fish used were marked by Coded Wire Tags (CWTs). The accuracy of GSI and its dependence on the quality of genetic data were studied by computer experiments. Rare alleles, which could result from poor sampling procedures, can lead to significant estimation errors. Estimation accuracy depended on the concordance between stocks present in the baseline data and the mixture sample and on the number of loci used in the analysis. Two methods for computing the contributions of groups of source stocks were found to be comparable under most, but not all, conditions. In a blind test of GSI, stock group composition estimates had absolute errors of less than 3%. This suggests that the GSI method can produce accurate stock contribution estimates using real fishery data.

La méthode d'identification génétique des stocks se sert de la variation des allozymes pour établir la composition de pêcheries pluristocks. Les auteurs l'ont appliquée à des données actuelles sur la pêche et présentent les résultats obtenus pour la première fois dans une publication primaire. Ils ont pu identifier les échantillons tirés d'un stock source et les échantillons mélangés capturés en mer, et ont déterminé exactement la composition de ceux-ci car les poissons portaient des étiquettes métalliques codées. Les auteurs ont aussi étudié la précision de cette méthode d'identification et sa dépendance sur la qualité des données génétiques par l'entremise d'expériences informatiques. Les allèles rares, qui sont peut-être le résultat de méthodes d'échantillonnage inadéquates, peuvent être à la source d'importantes erreurs d'estimation. La précision de l'estimation dépendait de l'accord entre, d'une part, les stocks représentés dans les données de base et les échantillons mélangés et, d'autre part, le nombre de loci utilisés lors de l'analyse. Deux méthodes de calcul de la contribution de groupes de stocks sources ont été identifiées comme équivalentes dans presque toutes les conditions. Lors d'une épreuve à l'insu de la méthode d'identification génétique, l'erreur absolue des estimations de la composition du groupe de stock était inférieure à 3 %. Ceci porte à croire que cette méthode peut donner des estimations précises de la contribution d'un stock à partir de données actuelles sur la pêche.

n recent years, some chinook salmon (*Oncorhynchus tshawytscha*) stocks have been depleted to the point that season closures have been imposed on commercial troll fisheries that exploit a mixture of hatchery and wild stocks (Fraidenberg and Lincoln 1985; Pacific Fishery Management Council (PFMC) 1986). Mark–recapture techniques can provide estimates of the contributions of hatchery stocks to the fishery harvest, but unless expensive and extensive marking programs are undertaken, the relative contributions of wild stocks to the harvest remain unknown. Without this information, it is difficult to assess the effectiveness of management actions designed to protect depleted wild stocks or to prevent overescapement of abundant wild stocks. Thus, developing the capacity to directly estimate the contributions of wild stocks in mixed-stock fisheries has enormous potential application.

Genetic Stock Identification (GSI) (Grant et al. 1980; Milner et al. 1981; Miller et al. 1983; Milner et al. 1983; Beacham et al. 1985; Milner et al. 1985; Beacham et al. 1987; Pella and Milner 1987; Utter et al. 1987, 1989; Shaklee et al. 1990b) has emerged as one means to distinguish stocks of salmon and directly estimate the composition of mixed-stock fisheries. In

---

[1]Current address: Northeast Fisheries Science Center, Woods Hole, MA 02543, USA.

[2]Current address: Fisheries Department, Food and Agriculture Organization, Via del Terme di Caracalla 00100, Rome, Italy.

[3]Current address: Department of Systematics and Ecology, Haworth Hall, University of Kansas, Lawrence, KS 66045, USA.

[4]Author to whom correspondence should be addressed.

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1507

TABLE 1. Sets of loci used to analyze the stock composition of mixture samples (locus nomenclature follows Shaklee et al. (1990a); an asterisk denotes that a locus is present in the set).

Locus	Locus sets		
	11-locus	15-locus	35-locus
sAAT-1,2			*
AAT-3			*
mAAT-1			*
mAAT-2			*
ADA-1			*
AH-1			*
mAH-3			*
mAH-4			*
DPEP-1	*	*	*
EST-3	*		
GPI-1	*	*	
GPI-2	*	*	*
GPI-3	*	*	
GPIH			*
HAGH			*
IDH-2			*
IDH-3			*
IDH-4	*	*	*
LDH-4	*	*	*
LDH-5			*
MDH-1,2	*	*	*
MDH-3,4	*	*	*
mMDH-1			*
MDHp-1			*
MDHp-2			*
MPI	*	*	*
PDPEP-2			*
PEPLT			*
PGDH	*		*
PGK-2			*
PGM-1	*		*
PGM-2	*		*
SOD-1	*	*	*
TAPEP-1	*	*	*
TPI-3			*
TPI-4			*
PGM-3			*
PGM-4			*

genetic data affected the accuracy of stock composition estimates.

In addition, we provide results concerning a "blind test" of GSI in which the mixture analysis was performed without prior knowledge of the composition of the sample. Such a blind test insured integrity of the analysis. This represents the first such test to be reported in the primary literature, although similar tests appear elsewhere (Milner et al. 1981).

Methods

In general, the application of GSI to determine the stock composition of a mixed-stock fishery involves three phases: (1) the collection and genetic description of baseline data from stocks that could be present in a mixed-stock fishery (baseline samples), (2) the genetic description of a representative sample collected from the mixed-stock fishery (mixture sample), and (3) the estimation of the most probable stock composition of the sample from the mixed-stock fishery using the method of maximum likelihood (mixture analysis). A detailed description of each phase as implemented in this paper follows.

Genetic Description

Allozyme analysis provided genetic descriptions of baseline and mixture samples. Samples of eye, heart, liver, and muscle tissue were collected from fall-run juvenile chinook salmon for all of the baseline samples except the Upper Sacramento River sample, which consisted of winter-run juveniles. Samples of eye and muscle tissue were collected from legal-sized adult chinook salmon for the mixture samples. The tissue samples for the baseline data were assayed at up to 124 isozyme loci using allozyme electrophoresis (Aebersold et al. 1987). We required that information on loci be present in all baseline samples and in at least some fish in the mixture sample for all loci used in stock composition analyses. Thus, due to the tissue-dependent expression of loci, only the subset of loci for which allelic variation could be detected from either muscle or eye tissue was used for GSI. Prior to 1987, this requirement limited the number of available loci to 15 loci (Table 1), but by 1987, 35 loci met this criterion (Table 1).

Baseline Samples

In the absence of prior information about which stocks in the set of baseline data are contributors, it is prudent to consider every stock to be a potential contributor. For the application of GSI to Pacific salmon (*Oncorhynchus* spp.), each "stock" should ideally represent a more or less discrete, self-sustaining breeding unit spawning at a given time and place within a lake or stream (Ricker 1972). Since we lacked conclusive evidence about whether or not each baseline sample constituted a stock under this definition, we considered samples from different collection sites to constitute samples from individual source populations (source stocks). Our use of the term "stock" also contrasts with another definition as a fish population in Hardy–Weinberg equilibrium (Booke 1981) that facilitates the computation of genotype probabilities required for stock composition analyses (cf. Fournier et al. 1984). Nonetheless, our usage is consistent with the purpose of testing whether the GSI method can accurately determine the contributions of groups of source stocks from various rivers to mixtures composed of ocean-caught fish with known stocks of origin.

The baseline data used in this paper consist of source stock samples collected from streams ranging from California to

contrast with expensive mark–recapture techniques, GSI exploits inherent genetic differences among fish stocks to estimate the relative proportion of each in a representative sample from the mixed-stock fishery (mixture sample). In order to use GSI to assess mixed-stock fishery composition, genotypic or allelic frequency data from all stocks that are believed or known to contribute to the mixture are required. The set of genetic data for all potentially contributing stocks constitutes the "baseline data" (Milner et al. 1985). Given the baseline data and corresponding genetic data for the mixture sample, a computer-based estimation procedure is applied to calculate the most likely stock composition of the mixture.

In this paper, we test the ability of GSI to determine the stock composition of mixtures of chinook salmon that were captured in the ocean troll fishery off the coasts of California and southern Oregon. In this case, the fish in the mixture were collected independently of the baseline data. Chinook salmon used in this study had been marked with Coded Wire Tags (CWTs); therefore, their stock of origin was known. Given this information, we could perform mixture analyses on mixture samples of known stock composition to explore how different uses of the

TABLE 2. Baseline data used to analyze stock composition in tests using the 1986 CWT sample (26-stock baseline) and in the blind test using the 1988 CWT sample (37-stock baseline).

River system	Source stocks	Sample size in baseline data set	
		26-stock	37-stock
Eel River	Hollow Tree Creek	—	100
	Middle Fork Eel River	105	95
	Outlet Creek	51	—
	Redwood Creek (on Eel River)	—	93
	Salmon Creek	—	96
	South Fork Eel River	140	99
	Tomki Creek	50	—
	Van Duzen River	—	100
Klamath River	Blue Creek	—	100
	Bogus Creek	40	128
	Camp Creek	—	106
	Horse Linto Creek	—	100
	Irongate Hatchery	102	99
	Omagar Creek	—	100
	Salmon River	—	99
	Scott River	50	—
	Shasta River	120	100
	South Fork Trinity River	93	100
	Trinity River Hatchery	170	120
Mad River	Mad River Hatchery	96	99
	North Fork Mad River	—	61
Mattole River	Mattole River	72	100
Redwood Creek	Prairie Creek Hatchery	99	—
	Redwood Creek	96	95
	Redwood Creek Lagoon	—	100
Sacramento River	Bear River	73	—
	Coleman Hatchery	303	100
	Feather River Hatchery	107	100
	Merced River Hatchery	114	100
	Murphy Creek	54	—
	Nimbus Hatchery	—	100
	Sacramento River	105	—
	Secret Ravine	29	—
	Stanislaus River	32	—
	Tuolomne River	96	—
	Upper Sacramento River	—	94
	Yuba River	190	—
Smith River	Rowdy Creek Hatchery	105	62
	Smith River	58	—
	Middle Fork Smith River	—	99
Alsea River	Fall Creek Hatchery	—	100
Chetco River	Chetco River Hatchery	—	100
Coos River	Millacoma River	—	100
	Morgan Creek Hatchery	—	100
Coquille River	South Fork Coquille River	—	100
Elk River	Elk River Hatchery	—	100
Rogue River	Rogue River	—	100
	Applegate River	—	100
Umpqua River	Rock Creek Hatchery	—	100

British Columbia. The data were grouped according to sample collection date. Samples collected prior to 1987 and reported by Bartley and Gall (1990) and Gall et al. (1989) are the "26-stock" baseline data (Table 2) whereas samples collected during 1987 or 1988 are the "37-stock" baseline data (Table 2; see Gall et al. 1989, 1992 for details).

In 1986, we acquired data for 85 source stocks in Oregon, Washington, and British Columbia from the U.S. National Marine Fisheries Service (NMFS). These samples will be called the "coast-wide" baseline data (see Gall et al. 1989 for details); the NMFS acts as a repository for the data. The combined data from the 26-stock and the coast-wide baseline are the "111-

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1509

stock'' baseline data, while the ''48-stock'' baseline consists of all California and coastal Oregon stocks within the 111-stock baseline. Four variable loci in the 26-stock baseline were not measured in all samples in the coast-wide baseline, and consequently there were 11 loci available for mixture analysis with the 111-stock baseline (Table 1).

Mixture Samples

Mixture samples consisted of adult chinook salmon marked with CWTs (Jefferts et al. 1963) captured during the ocean troll fisheries off the coasts of California and Oregon in 1986 and 1988 (the 1986 CWT sample and the 1988 CWT sample, respectively). The 1986 CWT sample consisted of 1628 fish; 1163 fish collected by the Oregon Department of Fish and Wildlife (ODFW) and 465 fish collected by the California Department of Fish and Game (CDFG). The stock of origin for each sample was assigned using the PFMC publication ''Pacific Salmonid Coded Wire Tag Releases Through 1986'' to interpret tag codes. A summary of the geographic origins of fish in the 1986 CWT sample is provided in the Appendix. The 1988 CWT sample contained 220 fish collected by CDFG.

Mixture Analysis

The stock composition of a given mixture sample was estimated using the method of maximum likelihood conditioned on the baseline data being known exactly (Fournier et al. 1984). The EM algorithm (Dempster et al. 1977; Milner et al. 1981) was used to compute stock composition estimates with a cutoff parameter of 10^{-5} (that is, iteration stopped when the absolute value of the maximum difference between elements of successive iterates differed by less than the cutoff parameter). Hardy–Weinberg equilibrium and independence of loci were also assumed (Bartley and Gall 1990; D. M. Bartley, unpubl. data). An individual CWT fish was excluded from the mixture analysis when the probability of observing its genotype was less than 10^{-7} for all baseline stocks (the value 10^{-7} was selected empirically to ensure that round-off error did not affect our computations). Such fish have ''improbable genotypes'' relative to the baseline data being used, and, although the likelihood that a given fish will have an improbable genotype increases as the number of loci used for analysis increases, these fish are outliers relative to the baseline data used for mixture analysis. Standard deviation estimates reported for the blind test were computed using the infinitesimal jackknife covariance estimator (Millar 1987).

When large numbers of source stocks are potential contributors to a mixed-stock fishery, it is often necessary to summarize their contribution estimates into a composite ''stock group'' estimate to be able to apply the results. In this case, the goal of mixture analysis is to estimate the stock group contribution accurately, and estimating the source stock contributions becomes an intermediary computation.

Two approaches to computing estimates for stock groups were examined: the ''allocate and sum'' and the ''pool and allocate'' procedures (Wood et al. 1987). In the allocate and sum procedure, estimates of the contribution of individual source stocks to the mixture are computed first. Individual contribution estimates are then summed to produce the estimate of each stock group's contribution. In the pool and allocate procedure, a composite stock is formed for each stock group by pooling allele frequencies of all stocks within the stock group. The stock group's contribution is then estimated using the composite stock. In this paper, allele frequencies for composite stocks were computed as the unweighted average of the allele frequencies for the constituent stocks. We chose the unweighted average, which assumes that constituent stocks contribute equally to the group, because for many stocks, estimates of total population size, spawning escapement, or other indices of relative stock abundance were not available to suggest alternative weights.

Results

Preliminary Examination of the Method

Preliminary analyses were preformed using the 1986 CWT sample with the 26-stock baseline data to examine the potential utility of the GSI method to assess the stock composition of chinook troll fisheries off the coasts of California and Oregon. To evaluate the results, each source stock in the baseline data and each fish in the 1986 CWT sample was assigned to a geographic group based on the location of its river of origin (Table 2; Appendix). The geographic groupings of stocks were the *Sacramento group*, consisting of stocks originating in the Sacramento River and San Joaquin River; the *Middle coast group*, consisting of stocks originating in the Eel, Mad, or Mattole River; the *North coast group*, consisting of stock from the Klamath or Smith River and Redwood Creek (on the California coast); the *Oregon group*, consisting of stocks originating in coastal Oregon rivers; the *Washington/Columbia group*, consisting of stocks originating in Puget Sound, Washington, coastal rivers, or the Columbia River; and the *Canadian group*, consisting of stocks originating in British Columbia rivers (Table 2). By assigning each source stock and fish to one geographic group, we could compare the accuracy of estimated and actual contributions of each geographic group to the mixture sample. In fact, some grouping by geographic proximity was essential to the interpretation of test results because several source stocks were not matched by fish in the 1986 CWT sample. Conversely, several fish in the 1986 CWT sample were not matched by source stocks in the baseline data.

Several mixture analyses were performed with the 26-stock baseline to see whether estimated stock group contributions were accurate. In one analysis, the contributions for a subsample consisting of data for 725 CWT fish all originating from California stocks were estimated using the 26-stock baseline (Table 2). This explored how the estimation procedure would perform in a best-case scenario where the mixture sample consisted of a subset of the stocks present in the baseline. Thirty loci were used (sAAT-1,2, sAAT-3, AH-1, AK-1, CK-1, CK-2, DPEP-1, EST-3, EST-4, EST-5, G3PDH-1, G3PDH-2, GPI-1, GPI-2, GPI-3, GPIH, IDH-3, IDH-4, LDH-4, LDH-C, MDH-1,2, MDH-3,4, MDHp-1, MPI, PGDH, PGK-2, PGM-1, PGM-2, SOD-1, and TAPEP-1) and the results suggested that the methodology would work well. Estimated contributions for the Sacramento, Middle Coast, and North Coast stock groups were 0.327, 0.098, and 0.575, respectively, while the actual contributions were 0.330, 0.077, and 0.593. Thus, the largest absolute error in any contribution estimate was about 2%.

Despite these promising results, the 26-stock baseline data set was limited because it contained no source stock samples collected outside California. Therefore, whenever non-California stocks were present in a mixture sample, the use of the 26-stock baseline would misallocate the contributions of non-California stocks to California stocks. To remove this limitation, we used the 111-stock baseline to perform further analyses using the 1986 CWT sample. However, some of these

1510

TABLE 3. Summary of sets of loci, baseline data, and mixture sample data used in tests involving the 1986 CWT sample (an asterisk denotes that the set of data was used in a particular test).

| Test | Locus set | | Baseline data | | | Mixture sample | |
	11-locus	15-locus	26-stock	48-stock	111-stock	All	California[a]
1	*		*	*	*	*	
2	*	*	*				*
3		*	*				*
4		*	*				*

[a] The subsample of 926 CWT fish of California origin.

TABLE 4. Results of Test 1 examining the effect of applying different portions of the 111-stock baseline on estimation accuracy (NE denotes that no estimate of this stock group's contribution was made, since it was assumed that the group does not contribute to the mixture sample).

| Stock group | Estimated contribution by baseline data set | | | |
	Actual contribution	111-stock	48-stock	26-stock
Eel	0.051	0.045	0.044	0.041
Klamath	0.352	0.332	0.300	0.449
Mad	0.000	0.000	0.000	0.000
Mattole	0.000	0.000	0.000	0.000
Redwood	0.000	0.000	0.003	0.165
Sacramento	0.166	0.036	0.148	0.346
Smith	0.000	0.000	0.000	0.000
Oregon	0.400	0.371	0.505	NE
Columbia	0.030	0.091	NE	NE
Washington	0.001	0.000	NE	NE
BC	0.000	0.125	NE	NE
Puget Sound	0.000	0.000	NE	NE

analyses produced unexpectedly poor results. In particular, one analysis that applied the 111-stock baseline to the entire 1986 CWT sample produced very large estimation errors for the Washington/Columbia and Oregon stock groups. Most of the misallocation resulted from a gross overestimate (estimate 49%, actual <1%) of the contribution of one stock, the Priest Rapids – Hanford stock, which belonged to the Washington/ Columbia stock group. The cause of this overestimate was traced to the rare LDH-C*70 allele at the LDH-C locus, which represented genetic variability now recognized to be an artifact of poor tissue quality. The LDH-C*70 allele had never been observed in baseline samples analyzed at the University of California, Davis (UCD) laboratory, although it was scored for some fish in the 1986 CWT sample and was present at low frequency in two Columbia River stocks (Priest Rapids – Hanford and Ice Harbor – Lyons Ferry). Further investigation showed that improper handling of some of the CWT fish had led to tissue degradation and misinterpretation of the common LDH-C*100 allele as the LDH-C*70 allele. Since the LDH-C*70 allele was only present in the Priest Rapids – Hanford and Ice Harbor – Lyons Ferry stocks, fish with this allele in the 1986 CWT sample could only be allocated to one of these two baseline stocks under the standard GSI model. This led to a large overestimate of the Priest Rapids – Hanford contribution in the mixture analysis. More importantly, this analysis demonstrated the necessity of consistent data interpretation and of proper handling of tissue samples for the successful application of the GSI method.

Computer Experiments Using the 1986 CWT Sample

Following a thorough reexamination of the 1986 CWT sample, a number of computer experiments (Table 3) were performed to examine how estimation accuracy depended on the selection and application of the available data. In each of our tests, stock contribution estimates were compared with actual sample composition, and the reported error was the difference between the estimated and actual stock group contribution.

The tests focussed on three targeted groups of stocks: those from the Eel, Smith, and Klamath rivers (Gall et al. 1989). Eel and Smith stocks were important because little information on ocean productivity and spawning escapement is available for these stocks (spawning stock surveys are available for only two minor tributaries of the Eel River; PFMC 1991). Klamath River stocks were important because the rebuilding, conservation, and harvest allocation among user groups of these stocks have been contentious and much debated in recent years (Fraidenberg and Lincoln 1985; McEvoy 1986). In fact, measures to reduce the ocean harvest of Klamath River stocks have included season closures as well as other restrictions on the commercial salmon troll fishery in the Klamath Management Zone (KMZ), an area of the Pacific Ocean extending from Point Delgada, California, to Cape Blanco, Oregon, where the commercial chinook harvest was well below historic levels throughout much of the 1980's (PFMC 1986; PFMC 1991, fig. II-4).

Effect of baseline stock choice (Test 1, Table 4)

The choice of stocks that are included in the baseline data is an a priori assumption that may affect estimation errors. This effect is most pronounced when the mixture sample contains stocks that are not present in the baseline data set. In Test 1, three mixture analyses applied successively smaller baseline data sets to the largest possible mixture sample. Each analysis used the entire 1986 CWT sample and the 11-locus set of loci. The Washington/Columbia stock group was subdivided into three geographic groups: Columbia, Washington Coast, and Puget Sound. Similarly, the North coast group was subdivided into the Klamath, Redwood Creek, and Smith geographic stock groups, and the Middle coast group was subdivided into the Eel, Mad, and Mattole geographic stock groups. All stock group contribution estimates were summarized using the allocate and sum procedure.

First, we used the 111-stock baseline data to analyze the composition of the entire 1986 CWT sample (thereby effectively assuming that any of the 111 stocks were potential contributors to the mixture). Second, we excluded the British Columbia (BC), Washington, and Columbia River baseline data and used the 48-stock baseline to analyze the entire 1986 CWT sample. This analysis simulated the assumption that only stocks originating from river systems south of the Columbia River were potential contributors to the mixture and led to the deliberate

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1511

TABLE 5. Results of Test 2 comparing the effect of using the 11-locus and the 15-locus sets of loci with the 26-stock baseline.

| Stock group | Actual contribution | Set of loci used for mixture analysis | | | |
| | | 11-locus | | 15-locus | |
		Estimate	Error	Estimate	Error
Eel	0.090	0.113	0.023	0.089	−0.001
Klamath	0.617	0.600	−0.017	0.606	−0.011
Mad	0.000	0.000	0.000	0.000	0.000
Mattole	0.000	0.000	0.000	0.000	0.000
Redwood	0.000	0.021	0.021	0.031	0.031
Sacramento	0.293	0.266	−0.027	0.273	−0.020
Smith	0.000	0.000	0.000	0.000	0.000

underestimation of the Columbia River and Washington Coast contributions. In reality, a few fish from these groups were included in the mixture sample. This analysis explored the trade-offs in estimation accuracy that resulted from the assumed absence of the more northerly stock groups. Third, we explored how the contribution of non-California stocks would be misallocated to California-origin stocks when only the 26-stock baseline was used to analyze the comparison of the entire 1986 CWT sample.

The results (Table 4) demonstrate the importance of trying to match the baseline data with the mixture sample. The total error (the sum of the absolute estimation errors for all stock groups) was smallest (0.22) when the 48-stock baseline was used. In contrast, the use of the 111-stock and 26-stock baseline data sets led to total errors of 0.37 and 0.88, respectively. While the GSI method will misallocate the contributions of stocks present in the mixture sample but not present in the baseline data, the exclusion of relatively minor contributors may improve the match between stocks present in the baseline data and the mixture sample and can, in some instances, actually improve estimation accuracy.

Comparing the 11-locus and 15-locus sets of loci (Test 2, Table 5)

In Test 2, we examined whether the exclusion of genetic data would affect the accuracy of stock composition estimates. For this, we used a mixture consisting of fish originating in California. As described above, when the coast-wide baseline data were added to our 26-stock baseline data, the number of loci available for analysis was reduced from 15 to 11. The four loci no longer available were EST-3, PGDH, PGM-1, and PGM-2 (Table 1). Of these four, only PGM-2 was both a highly variable and a easily resolvable locus. A priori, it was not clear whether estimation accuracy would be lost through the use of this 11-locus loci when only California-origin stocks were present in both the baseline and the mixture sample. Thus, we analyzed the mixture consisting of all 926 California-origin fish in the 1986 CWT sample using the 26-stock baseline and both sets of 15 or 11 loci.

The test results (Table 5) show that the total estimation error was only slightly larger (0.09 versus 0.06) when the 11-locus set of loci was used. Note, however, that the contribution of Redwood Creek was actually estimated less accurately when the larger set of loci was used. Thus, the inclusion of more loci need not uniformly reduce estimation accuracy.

Effect of stock grouping on the pool and allocate procedure (Test 3, Table 6)

In Test 3, we examined whether different stock groupings would affect estimation accuracy when the pool and allocate

TABLE 6. Results of Test 3 comparing the effect of using different stock groupings with the pool and allocate procedure.

Stock group	Actual contribution	Estimate	Error
Using 3 stock groups			
Coastal$_5$	0.090	0.125	0.035
Klamath	0.617	0.602	−0.015
Sacramento	0.293	0.273	−0.020
Using 5 stock groups			
Coastal$_3$	0.000	0.000	0.000
Eel	0.090	0.081	−0.009
Smith	0.000	0.080	0.080
Klamath	0.617	0.576	−0.041
Sacramento	0.293	0.262	−0.031
Using 7 stock groups			
Eel	0.090	0.027	−0.063
Mad	0.000	0.018	0.018
Mattole	0.000	0.000	0.000
Redwood Creek	0.000	0.026	0.026
Smith	0.000	0.090	0.090
Klamath	0.617	0.576	−0.041
Sacramento	0.293	0.264	−0.029

procedure was applied (Table 6). The stock groups were chosen to reflect the importance of the Eel, Klamath, and Smith stocks, as well as other stocks originating in California coastal rivers. In particular, the three analyses in this test focussed on stocks originating in northern California coastal streams which are important contributors to the KMZ (PFMC 1991) and for which an exploration of the effects of stock grouping was considered important. Each analysis used a subsample of the 1986 CWT sample consisting of California-origin fish so that the 15-locus loci set could be used. All stock groups were formed by pooling allele frequency data. (A corresponding analysis using the allocate and sum procedure was applied in Test 2 (Table 5).)

For the first analysis, the California-origin source stocks were grouped into three stock groups: Coastal$_5$, Klamath, and Sacramento. These three stock groups represented a coarse level of pooling for the 26-stock baseline and eliminated the possibility of estimating the Eel and Smith contributions separately. The Coastal$_5$ composite stock consisted of the pooled allelic data for the source stocks from the Eel, Mad, Mattole, and Smith rivers and Redwood Creek (Table 2). The Klamath com-

1512

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

posite stock consisted of the pooled allelic data for the source stocks from the Klamath–Trinity drainage. The Sacramento composite stock consisted of the pooled allelic data for the source stocks from the Sacramento – San Joaquin drainage. We next grouped the California source stocks into five stock groups so that estimates of the Eel and Smith contributions could be made. Two additional stock groups were defined by dividing the Coastal$_5$ stock group into three smaller groups (Coastal$_3$, Eel, and Smith). The Eel and Smith stock groups consisted of source stocks from the Eel and Smith rivers, respectively, while the Coastal$_3$ group included source stocks from the Mad and Mattole rivers and Redwood Creek. Finally, we formed seven stock groups by splitting the Coastal$_3$ group into the Mad and Mattole River and Redwood Creek stock groups. This analysis with seven groups explored whether estimation accuracy would be increased by attempting to estimate the contributions of the remaining coastal rivers separately.

The smallest total error (0.07) was obtained when three stock groups were used (Table 7). Total error more than doubled to 0.16 and 0.27 when five and seven stock groups were used, respectively. To some extent, the increase in estimation error is due to the additional number of parameters that must be estimated when five or seven stock groups are used. The trade-off, however, was that no direct estimates of the relative contributions of source stocks originating in the Eel and Smith rivers could be made when only three stock groups were used.

Comparing the allocate and sum and pool and allocate procedures (Test 4, Table 8)

We performed three comparisons of the accuracy of the allocate and sum and the pool and allocate procedures (Table 8). Each comparison used the 15-locus set of loci. First, we compared the allocate and sum and pool and allocate procedures under the best circumstances, that is, when stocks present in the mixture sample matched the stocks present in the baseline.

This comparison used the subsample of the 1986 CWT sample consisting of California-origin fish and the 26-stock baseline. Second, we compared the two procedures when the Eel River stock group was absent from the mixture sample but was present in the baseline. In this case, we analyzed the subsample of the 1986 CWT sample consisting of California-origin fish with the samples originating in the Eel River drainage removed using the 26-stock baseline. Third, we compared the procedures when the Klamath River stock group was absent from the mixture sample but was present in the baseline. For this, we used the subsample of the 1986 CWT sample consisting of California-origin fish with the samples originating in the Klamath River drainage removed and the 26-stock baseline. The objectives of these two comparisons were to determine if the GSI method could detect the absence of Eel or Klamath River stocks and to compare relative performance of the allocate and sum and pool and allocate procedures under these circumstances. A corresponding comparison could not be made using data for the Smith River, since there were no CWT samples from this river.

We found that, under the best circumstances, the estimated stock group contributions were extremely accurate (total errors <0.04) and that either method of pooling data worked well. When the Eel River stocks were excluded from the mixture sample but included in the baseline data, the allocate and sum procedure (total error 0.09) was more accurate than the pool and allocate procedure (total error 0.22) but more severely overestimated the contribution of the Eel River stock group. Similarly, when the Klamath River stocks were excluded from the mixture sample but included in the baseline data, the allocate and sum procedure (total error 0.28) was more accurate than the pool and allocate procedure (total error 0.35), but only marginally so and once again more severely overestimated the Klamath stock group contribution. Obviously, a stock presumed to be in a mixture sample which is actually not present may be

TABLE 7. Results of Test 4 comparing the allocate and sum and the pool and allocate procedures for summarizing stock group contributions.

| Stock group | Actual contribution | Pooling procedure | | | |
| | | Allocate and sum | | Pool and allocate | |
		Estimate	Error	Estimate	Error
Best circumstances: stocks in the baseline data and mixture sample match					
Eel	0.090	0.106	0.016	0.109	0.019
Klamath	0.617	0.617	0.000	0.614	−0.003
Sacramento	0.293	0.277	−0.016	0.277	−0.016
Eel River stocks absent from the mixture but present in the baseline					
Coastal$_3$	0.000	0.013	0.013	0.000	0.000
Eel	0.000	0.030	0.030	0.009	0.009
Smith	0.000	0.003	0.003	0.100	0.100
Klamath	0.678	0.664	−0.014	0.616	−0.062
Sacramento	0.322	0.291	−0.031	0.275	−0.047
Klamath River stocks absent from the mixture but present in the baseline					
Coastal$_3$	0.000	0.049	0.049	0.025	0.025
Eel	0.235	0.169	−0.066	0.133	−0.102
Smith	0.000	0.004	0.004	0.127	0.127
Klamath	0.000	0.087	0.087	0.025	0.025
Sacramento	0.765	0.691	−0.074	0.690	−0.075

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1513

TABLE 8. Results of the blind test of GSI with the set of 34 loci used to form stock groups. Estimated standard deviations (SD) accompany stock group contribution estimates.

Stock group	Actual contribution	Pooling procedure			
		Allocate and sum		Pool and allocate	
		Estimate ($\pm$ SD)	Error	Estimate ($\pm$ SD)	Error
Sacramento	0.223	0.272 ($\pm$0.057)	0.049	0.242 ($\pm$0.026)	0.019
Coastal	0.036	0.028 ($\pm$0.086)	-0.008	0.021 ($\pm$0.014)	-0.015
Klamath	0.486	0.503 ($\pm$0.066)	0.017	0.542 ($\pm$0.058)	0.056
S. Oregon	0.246	0.198 ($\pm$0.140)	-0.048	0.195 ($\pm$0.033)	-0.061
Mid-Oregon	0.009	0.000 ($\pm$0.034)	-0.009	0.000 ($\pm$0.027)	-0.009

overestimated, and these results show that the allocate and sum procedure can exaggerate this overestimation. These results also show that the trade-off between total estimation error and accurate estimation of a single stock group contribution depends on the particular stock group of concern.

A Blind Test of GSI

The tests using the 1986 CWT data suggested that the GSI method was limited by the degree of concordance between stocks present in the baseline data and the mixture sample and by the number of variable loci available for mixture analyses. To increase the number of available loci, baseline samples were collected again in 1987 and 1988. This resampling more than doubled the number of loci available for mixture analyses. A blind test of the GSI method was then conducted using the 1988 CWT sample to gauge the performance of this new baseline data. To ensure the integrity of the test, the contribution estimates were computed without any prior knowledge of the sample's composition, which was known only to personnel at CDFG.

To summarize the estimation results, five genetic stock groups were defined (Gall et al. 1989) based on a dendrogram of genetic similarity (Wood 1989). The dendrogram was produced using Nei's genetic distance measure (Nei 1972) calculated from the allele frequency estimates for 34 loci in the 35-locus loci set (Table 1), excluding MDHp-2. The MDHp-2 locus was excluded from the construction of the dendrogram because heterozygotes could not be scored directly so that allele frequencies for the jMDHp-2 locus could only be estimated from homozygote frequencies (see Gall et al. 1989 for details). The five prominent genetic stock groups present in the dendrogram were (1) the *Sacramento group* (Coleman Hatchery, Feather River Hatchery, Upper Sacramento River, Merced River Hatchery, and Nimbus Hatchery); (2) the *California coastal group* (Middle Fork Eel River, South Fork Eel River, Salmon Creek, Redwood Creek (on Eel River), Hollow Tree Creek, Van Duzen River, Mad River Hatchery, North Fork Mad River, Mattole River, Redwood Creek, and Redwood Creek Lagoon); (3) the *Klamath group* (Salmon River, Irongate Hatchery, Bogus Creek, Camp Creek, Horse Linto Creek, Shasta River, Trinity River Hatchery, and the South Fork Trinity River); (4) the *Southern Oregon coastal and Smith River group* (Blue Creek, Omagar Creek, Rowdy Creek Hatchery, Middle Fork Smith River, Chetco River Hatchery, Rogue River, Applegate River, and Rock Creek Hatchery); and (5) the *Mid-Oregon coastal group* (Millacoma River, Morgan Creek Hatchery, Fall Creek Hatchery, Elk River Hatchery, and the South Fork Coquille River).

Both the allocate and sum and the pool and allocate procedures were employed to summarize the stock group estimates

TABLE 9. Results of the blind test of GSI using the 35-locus set of loci. Estimated standard deviations (SD) accompany stock group contribution estimates.

Using the allocate and sum procedure			
Stock group	Actual contribution	Estimate ($\pm$ SD)	Error
Sacramento	0.223	0.246 ($\pm$0.162)	0.024
Coastal	0.036	0.048 ($\pm$0.079)	0.012
Klamath	0.486	0.498 ($\pm$0.050)	0.012
S. Oregon	0.246	0.208 ($\pm$0.124)	-0.038
Mid-Oregon	0.009	0.000 ($\pm$0.131)	-0.009
Using the pool and allocate procedure			
Sacramento[a]	0.219	0.242 ($\pm$0.028)	0.023
Coastal	0.037	0.038 ($\pm$0.021)	0.001
Klamath	0.489	0.485 ($\pm$0.051)	-0.004
S. Oregon	0.247	0.236 ($\pm$0.037)	-0.011
Mid-Oregon	0.009	0.000 ($\pm$0.031)	-0.009

[a]One sample from the Sacramento stock group had an improbable genotype relative to this baseline data and set of loci.

(Table 8). Mixture analyses were performed with the set of 34 loci used to construct the dendrogram and with the entire 35-locus set of loci (Tables 8 and 9). Data for the MDHp-2 locus were included to see how this affected the accuracy and precision of the estimates (Table 9). Since the MDHp-2 locus was known to be highly variable for the Klamath stocks, variable in some coastal Oregon stocks, and almost monomorphic in other baseline stocks (Gall et al. 1989), it was expected that the use of 35-locus set of loci would improve the accuracy and precision of estimates for the Klamath and possibly other stock groups.

Indeed this was the case, for estimates of the Klamath stock group contribution were more accurate and precise when the 35-locus set of loci was used (Tables 8 and 9). Total estimation errors were also lower for both summarizing procedures when the 35-locus set of loci was used. Further, when the pool and allocate procedure was used with the 35-locus set of loci (Table 9), we estimated that the Sacramento stock group contributed 24.2 $\pm$ 2.8% of the fish (actual 21.9%), the California coastal stock group contributed 3.7 $\pm$ 2.1% (actual 3.7%), the Klamath stock group contributed 48.5 $\pm$ 5.1% (actual 48.9%), the Southern Oregon coastal and Smith River stock group contributed 23.6 $\pm$ 3.7% (actual 24.7%), and that the Mid-Oregon coastal stock group contributed 0.0 $\pm$ 3.1% (actual 0.9%). Thus, we were able to estimate very accurately the contributions of five major stock groups to the mixture sample.

1514

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

Discussion

The results of the blind test show that GSI analysis can produce accurate stock group contribution estimates using actual fishery data collected in ports along the Oregon and California coasts. With the number of polymorphic loci available for analysis of chinook salmon and appropriate baseline data, the estimated contributions of stock groups to the mixture sample had absolute errors no greater than 6.2% and had total estimation errors no greater than 16.0%. In addition, the blind test results compared favorably with other test results, despite the fact that the blind test was performed on the smallest mixture sample used in this study. Indeed, with 35 loci using pool and allocate, the estimation error for each stock group was less than 2.3% based on a sample of only 220 fish.

The results of the blind test also corroborate the validation study of Milner et al. (1981) (performed on mixtures composed of tagged juvenile fish) as well as the findings of several simulation studies showing that the GSI method can produce accurate results using genetic data for Pacific salmon (Fournier et al. 1984; Beacham et al. 1985; Pella and Milner 1987; Wood et al. 1987; Brodziak 1990). In particular, the blind test results are consistent with the findings of Wood et al. (1987, fig. 9) that show that small mixture samples are sufficient for mixture analyses when enough genetic variability exists between stock groups. While simulation studies are an extremely useful tool to examine the sensitivity of the GSI method to various assumptions, such studies cannot anticipate some difficulties that can occur in practice, even when resampling techniques such as the bootstrap are applied to simulate the effect of sampling variation. Indeed, the difficulties with the LDH*C locus that occurred during our preliminary analyses using the 1986 CWT sample show the importance of proper handling of tissue samples, data standards, and thorough screening of all variable loci to the GSI method. Spurious genetic variation was observed in the 1986 CWT sample at the LDH-C locus because some tissue samples were not stored in $-80°C$ freezers. Electrophoresis of the LDH enzyme using the degraded tissues produced misleading banding patterns. Thus, tissue degradation altered data interpretation, which in turn led to inaccurate stock composition estimates. At present, GSI laboratories at UCD, Washington Department of Fisheries, and NMFS have exchanged standards and tissue samples to assure consistent allele scoring for chinook salmon. In fact, the LDH-C*70 allele was eliminated from the coast-wide database for chinook salmon in 1989 because this allele was determined to be an artifact of poor sample quality. At present, there is also an interlaboratory loci sponsorship program that excludes a locus from the coast-wide baseline if any of the participating laboratories cannot reliably score the alleles for that locus. Such cooperation has greatly enhanced and accelerated the advancement of GSI technology for fisheries management. Additionally, many loci now recognized as important for GSI analyses of chinook salmon were not available for the preliminary analyses or tests using the 1986 CWT sample because of primitive laboratory techniques that impaired the detection of variable loci.

Nonetheless, the tests using the 1986 CWT sample lead to a number of observations that have broader implications to the general application of the GSI method. First, the results of the tests on the effect of choosing stocks included in the baseline data show that the magnitude of estimation error for a particular stock group depends on which stock groups are included in the baseline, if the set of loci and mixture sample are fixed. Accurate estimates for the targeted Eel and Smith stock groups were accompanied by inaccurate estimates for the Klamath and Sacramento stock groups. On the one hand, by including as many source stocks and loci as possible, substantial misallocations to noncontributing stocks can occur. On the other hand, estimation errors can also be substantial if contributing source stocks are excluded based on prior assumptions about fishery composition. Both situations are undesirable. However, a lack of detectable genetic differences between stocks may preclude accurate estimation even if the baseline data can be judiciously chosen to exclude stocks whose contribution to the mixture sample is insignificant. This highlights the need to examine existing data and hypotheses concerning the mixing and relative abundance of stocks when testing the adequacy of baseline data. Such tests can be performed using a combination of validation studies and exploratory simulation and genetic similarity analyses (Wood et al. 1987; Wood et al. 1989) before a final selection of stocks to include in the baseline data is made.

The results of the test comparing the 11-locus and 15-locus sets of loci show that changing the loci used in the analysis may improve estimation accuracy for some stock groups and simultaneously decrease the estimation accuracy for others. In particular, with the four additional loci the overestimate of the Redwood Creek stock group increased by 1%, while the estimation errors for the Eel, Klamath, and Sacramento stock groups decreased by less than 2%. In general, adding more loci will cost more money, but might not appreciably improve estimation accuracy for important stock groups. On the other hand, the blind test results show how the addition of a single variable locus can increase the accuracy of the estimates. In principle, loci should to be selected to decrease estimation error, and their relative effect on estimation error should be quantified before the GSI method is applied (cf. Gomulkiewicz et al. 1990).

The results of the test on the effect of stock grouping on the pool and allocate procedure show that enlarging the number of groups can decrease estimation accuracy for the pool and allocate procedure for a fixed mixture sample and set of loci. In fact, the lowest maximum absolute error was obtained using three stock groups at the expense of combining the Eel and Smith contributions. Creating two additional stock groups by splitting up the Coastal$_5$ group increased estimation errors for the Eel, Klamath, Sacramento, and Smith groups. However, progressing from five to seven stock groups did not appreciably affect the results and increased the largest absolute error by only 1%. Although the contributions of the three additional stock groups (Mad and Mattole rivers and Redwood Creek) were estimated accurately with seven groups, their summed estimation error was 3.9% larger when they were combined as a single unit ("Coastal$_3$" group). Thus, the choice of the best level of stock grouping depends on the accuracy required for each group's contribution.

Overall, the definition of stock groups should reflect the relative degree of genetic similarity among stocks within the baseline data to reduce misallocation to noncontributing stocks that exhibit genetic affinity with contributing stocks. Genetic similarity analyses can be useful for defining stock groups (Wood et al. 1989), but for stocks originating in northern California coastal rivers, the resulting groups might conflict with management programs that use geographic proximity to group stocks (Gall et al. 1989). In particular, by using seven instead of three stock groups, the assumption of independence of stock groups required by the standard GSI model may be untenable given the amount of detectable genetic variation. Such attempts to obtain more information from a limited set of genetic data may decrease estimation accuracy. This underscores the point

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1515

that attempting to accurately estimate the contribution of a particular stock group may be limited by genetic similarity between stocks.

The results of the test comparing the allocate and sum and the pool and allocate procedures show that the choice of a summarizing procedure may be important to the accurate estimation of particular stock group contributions. However, under the best circumstances, the estimation errors for all stock groups were small (less than 2%) using either the allocate and sum or the pool and allocate procedure. Thus, the summarizing procedure does not always affect the accuracy of the estimation.

Similarly, the absence of the Eel and Klamath stock groups was detectable by both summarizing procedures, if one accepts relatively small contribution estimates to indicate absence. However, while the pool and allocate procedure more accurately detected the absence of the targeted stock group in this test, it also increased the error for the nontargeted stock groups. The application of a variant of the standard GSI model to detect the number of source stocks present in the mixture might be helpful when the detection of an individual stock or stock group is important (Smouse et al. 1990).

A comparison of the results of Tests 2 and 3 also shows that the allocate and sum procedure can dominate the pool and allocate procedure with respect to total estimation error. In addition, a comparison of the precision of the two summarizing procedures in the blind test shows that, for these data, the pool and allocate procedure produces variance estimates for stock group contributions that are smaller than those using the allocate and sum procedure. So, although the allocate and sum procedure is more flexible and easier to apply to summarize contribution estimates (Wood et al. 1987), direct comparisons of the performance of the procedures should be made before the decision to apply either is made. These comparisons can be explored using validation or simulation studies.

GSI has emerged as a powerful tool in fisheries management (Shaklee et al. 1990b; Waples et al. 1990), and we have demonstrated that GSI has tremendous potential application to accurately assess the composition of chinook salmon fisheries off the coasts of California and Oregon. Regardless of the particular application, fisheries managers who want to use GSI to analyze the composition of mixed-stock fisheries must make practical choices in the face of the uncertain, but potentially important, effects of manipulating the genetic data. In general, the loci used for mixture analyses should reflect both the cost limitations of management programs and the accuracy requirements for improving the quality of stock assessments. However, the manipulation of baseline data to reflect groupings of stocks into management units may preclude the possibility of accurately estimating some stock groups' contributions. In such cases, the use of other population characteristics may improve the discriminatory power of stock identification analyses (Fournier et al. 1984; Wood et al. 1989). Moreover, even with prior information about the composition of a mixed-stock fishery, it is recommended that the adequacy of baseline data be examined through validation and simulation studies before GSI analyses are used to provide stock assessment advice. Nonetheless, our results are further evidence that GSI can provide accurate estimates of stock contributions in mixed-stock fisheries when the method is carefully applied.

Acknowledgments

References

AEBERSOLD, P. A., 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. U.S. Natl. Mar. Fish. Tech. Rep. 61.

BARTLEY, D. M., AND G. A. E. GALL. 1990. Genetic structure and gene flow in chinook salmon populations in California. Trans. Am. Fish. Soc. 119: 55–71.

BEACHAM, T. D., A. P. GOULD, R. E. WITHLER, C. B. MURRAY, AND L. W. BARNER. 1987. Biochemical genetic survey and stock identification of chum salmon (*Oncorhynchus keta*) in British Columbia. Can. J. Fish. Aquat. Sci. 44: 1702–1713.

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.

BOOKE, H. E. 1981. The conundrum of the stock concept — are nature and nurture definable in fishery science? Can. J. Fish. Aquat. Sci. 38: 1479–1480.

BRODZIAK, J. K. T. 1990. Theoretical aspects of genetic stock identification. Ph.D. dissertation, University of California, Davis, CA.

DEMPSTER, A. P., N. M. LAIRD, AND D. B. RUBIN. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 39: 1–38.

FOURNIER, D., 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.

FRAIDENBERG, M. E., AND R. H. LINCOLN. 1985. Wild chinook salmon management: an international conservation challenge. N. Am. J. Fish. Manage. 5: 311–329.

GALL, G. A. E., BARTLEY, D.M., B. BENTLEY, J. BRODZIAK, R. GOMULKIEWICZ, AND M. MANGEL. 1992. Geographical variation in the population genetic structure of Chinook salmon from California and Oregon. Fish. Bull. U.S. 90: 77–100.

GALL, G. A. E., B. BENTLEY, C. PANATTONI, E. CHILDS, C. QI, M. MANGEL, J. K. T. BRODZIAK, AND R. GOMULKIEWICZ. 1989. Chinook mixed fishery project 1986–1989. California Department of Fish and Game, Rancho Cordova, CA 95670.

GOMULKIEWICZ, R., J. K. T. BRODZIAK, AND M. MANGEL. 1990. Ranking loci for genetic stock identification by curvature methods. Can. J. Fish. Aquat. Sci. 47: 611–619.

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*) in Cook Inlet, Alaska. Can. J. Fish. Aquat. Sci. 37: 1236–1247.

JEFFERTS, K. B., P. K. BERGMAN, AND H. F. FISCUS. 1963. A coded wire identification system for macro-organisms. Nature (Lond.) 198: 460–462.

MCEVOY, A. F. 1986. The fisherman's problem: ecology and the law in the California fisheries, 1850–1980. Cambridge University Press, New York, NY.

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

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. Wash. Dep. Fish. Tech. Rep. No. 74.

MILNER, G. B., D. J. TEEL AND F. M. UTTER. 1983. Genetic stock identification study. NWAFC, NOAA, Seattle, WA 98112.

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), NWAFC, NOAA, Seattle, WA 98112.

MILNER, G. B., D. J. TEEL, F. M. UTTER, AND G. A. WINANS. 1985. A genetic method stock identification in mixed populations of Pacific salmon, *Oncorhynchus* spp. Mar. Fish. Rev. 47: 1–8.

NEI, M. 1972. Genetic distance between populations. Am. Nat. 106: 283–292.

PACIFIC FISHERY MANAGEMENT COUNCIL (PFMC). 1986. Review of 1985 ocean salmon fisheries. Pacific Fishery Management Council, Portland, OR 97201.

1516

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1991. Review of 1990 ocean salmon fisheries. Pacific Fishery Management Council, Portland, OR 97201.

RICKER, W. E. 1972. Heredity and environmental factors affecting certain salmonid populations, p. 19–160. *In* The stock concept of Pacific salmon. H. R. MacMillan Lectures in Fisheries, University of British Columbia, Vancouver, B.C.

PELLA, J. J., AND G. B. MILNER. 1987. Use of genetic marks in stock composition analysis, p. 247–276. *In* N. Ryman and F. Utter [ed.] Population genetics and fisheries management. University of Washington Press, Seattle, WA.

SHAKLEE, J. B., F. W. ALLENDORF, D. C. MORIZOT, AND G. S. WHITT. 1990a. Gene nomenclature for protein-coding loci in fish. Trans. Am. Fish. Soc. 119: 2–15.

SHAKLEE, J. B., C. BUSACK, A. MARSHALL, M. MILLER, AND S. R. PHELPS. 1990b. The electrophoretic analysis of mixed-stock fisheries of Pacific salmon, p. 235–265. *In* Z-I. Ogita and C. L. Markert [ed.] Isozymes: structure, function and use in biology and medicine. Progress in clinical and biological research. Vol. 344. Wiley-Liss, Inc., New York, NY.

SMOUSE, P. E., R. S. WAPLES, AND J. A. TWOREK. 1990. A genetic mixture analysis for use with incomplete source population data. Can. J. Fish. Aquat. Sci. 47: 620–634.

UTTER, F. M., G. B. MILNER, G. STAHL, AND D. J. TEEL. 1989. Genetic population structure of chinook salmon, *Oncorhynchus tshawytscha*, in the Pacific Northwest. Fish. Bull. U.S. 87: 239–264.

UTTER, F. M., D. J. TEEL, G. B. MILNER, AND D. McISAAC. 1987. Genetic estimates of stock compositions of 1983 chinook salmon harvests off the Washington coast and the Columbia River. Fish. Bull. U.S. 85: 13–23.

WAPLES, R. S., G. A. WINANS, F. M. UTTER, AND C. MAHNKEN. 1990. Genetic approaches to the management of Pacific salmon. Fisheries (Bethesda) 15: 19–25.

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

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., D. T. RUTHERFORD, AND S. McKINNELL. 1989. Identification of sockeye salmon (*Oncorhynchus nerka*) stock in mixed-stock fisheries in British Columbia and southeast Alaska using biological markers. Can. J. Fish. Aquat. Sci. 46: 2108–2120.

Appendix

TABLE A.1. Summary of stock groupings and stocks of origin for all chinook salmon in the 1986 CWT sample.

Stock group	Stock of origin	Sample size (percentage of total)
Sacramento	Coleman Hatchery	64 (3.9%)
	Feather River	80 (4.9%)
	Merced River	67 (4.1%)
	Mokelumne River	3 (0.2%)
	Nimbus Hatchery	50 (3.1%)
	Tehama Colusa	1 (0.1%)
	Yuba River	5 (0.3%)
Middle coast	Eel River	1 (0.1%)
	Redwood Creek (on Eel River)[a]	18 (1.1%)
	Silverado[a]	33 (2.0%)
	Silverking[a]	1 (0.1%)
	Sprowel Creek[a]	4 (0.2%)
	Van Arsdale[a]	26 (1.6%)
North coast	Bogus Creek[b]	4 (0.2%)
	Iron Gate Hatchery[b]	115 (7.1%)
	Trinity River[b]	454 (27.9%)
Oregon coast	Anadromous Inc.	123 (7.6%)
	Butte Falls	5 (0.3%)
	Cole Rivers	293 (18.0%)
	Elk River	50 (3.1%)
	Oregon Pacific	25 (1.5%)
	Oregon Aqua Foods	35 (2.2%)
	Rock Creek	32 (2.0%)
	Rogue River	40 (2.5%)
	Salmon River	4 (0.2%)
	Stayton Pond	2 (0.1%)
	Trask River	43 (2.6%)
Columbia River	Bonneville	3 (0.2%)
	Cowlitz	13 (0.8%)
	Hagerman	2 (0.1%)
	Lewis River	1 (0.1%)
	Little White Salmon	1 (0.1%)
	Lyons Ferry	12 (0.7%)
	McNary	1 (0.1%)
	Priest Rapids	1 (0.1%)
	Rocky Reach	8 (0.5%)
	Spring Creek	1 (0.1%)
	Washougal	6 (0.4%)
Washington coast	Nemah	1 (0.1%)
	Quinault	1 (0.1%)

[a]Samples originating in the Eel River drainage (see Test 4).
[b]Samples originating in the Klamath River drainage (see Test 4).

Can. J. Fish. Aquat. Sci., Vol. 49, 1992

1517