Applications of Population Genetics to Conservation of Chinook Salmon Diversity in the Central Valley

D. Hedgecock, M.A. Banks, V.K. Rashbrook, C.A. Dean, and S.M. Blankenship

Uses of Modern Population Genetics in Conservation

Population genetics is playing an increasingly important role in the conservation of salmonid resources in the Pacific Northwest. The National Marine Fisheries Service considers a salmon population worthy of conservation under the U.S. Endangered Species Act if it represents an Evolutionary Significant Unit (ESU), "…a population (or group of populations) that (1) is substantially reproductively isolated from other conspecific population units, and (2) represents an important component in the evolutionary legacy of the species" (Waples 1991, 1995). Genetic data provide an important, though indirect means for establishing the degree of reproductive isolation between conspecific populations. Indeed, numerous studies of electrophoretically detectable protein polymorphisms carried out over the past 30 years on Pacific salmon species have shown that a high degree of spatial substructure and reproductive isolation results from their homing behavior (Utter 1991). With the advent of DNA markers, particularly mitochondrial and microsatellite DNA markers, resolution of reproductively isolated or partially isolated populations has become more precise. Here, we describe progress resolving chinook salmon diversity and stock structure in the Central Valley of California.

Modern tools of population genetics, for example, using polymorphic protein markers, also allow us to address problems that could not formerly be approached. Whereas protein markers had long supported the statistical allocation of catch in mixed ocean fisheries to contributing spawning populations (Utter and Ryman 1993), highly polymorphic microsatellite DNA markers now enable us to ascertain the origins of individual fish. We describe how individual assignment of salmon, first achieved for Central Valley chinook salmon, has become an integral part of a hatchery supplementation program for the endangered Sacramento River winter chinook salmon. Individual assignment is also being used to the identify winter-run juvenile migration patterns through the Sacramento-San Joaquin Delta and in assigning ocean catches to various Central Valley stocks, some of which are threatened or endangered.

The use of highly polymorphic DNA markers has also enabled tremendous improvements in identifying parentage and kinship. Indeed, determining the parentage of hatchery-reared winter chinook in the supplementation program was our original motivation for developing microsatellite DNA markers. Since then, microsatellite markers have provided, aside from a description of genetic diversity within and among Central Valley stocks, an important validation of the demographic model used to assess the genetic effect of the hatchery supplementation program for winter run. Microsatellite markers also allow the assessment of kinship in juvenile samples, which often are the only material that can be collected from small, threatened, or endangered populations. In the past, population geneticists advised against using juvenile samples because of the potentially confounding effects of family structure on the estimation of adult allele frequencies (Allendorf and Phelps 1981). Now, however, highly polymorphic markers enable the kinship of juveniles to be detected and the effects of family structure to be removed. Such data not only allow the genotypes and allele frequencies of the unobserved adult population to be reconstructed but also shed light on the reproductive behavioral ecology of salmon populations.

Genetic Diversity of Chinook Salmon in the Central Valley

Phenotypic Diversity

Major spawning subpopulations among California's Central Valley chinook salmon have very similar anatomical and morphological features but marked differences in timing of spawning, juvenile emergence, early rearing and migration from the freshwater habitat to the ocean. Four runs have been named—winter, spring, fall, and late-fall—based on the season when most individuals from a subpopulation return to freshwater for spawning (Stone 1874; Fry 1961). Spawning not only occurs at a distinct time for each run, with only partial overlap between temporally adjacent runs, but, historically at least, often in a distinct habitat (for example, major rivers compared with higher elevation streams; see Fisher 1994). This natural, spatial and temporal isolation of the various spawning habitats has been greatly perturbed by human activity. For example, 150 years ago, spring and fall runs overlapped in spawning time but were geographically isolated; spring run spawned in the upper headwaters and fall run, in rivers and major streams of the lower valley floor. Forced co-existence of these two runs caused by substantial damming and loss of habitat in recent years, however, has lead to concern for their genetic integrity (Cope and Slater 1957; Banks and others 2000).

Several studies have focused on genetic characterization of California's Central Valley chinook salmon using of a variety of genetic marker types. Results will be presented for each type, separately, followed by a synthesis across marker types. Wright's (1931) standardized variance of allele frequencies among subpopulations, F_{ST}, is used to measure genetic diversity among runs and to compare results from different marker classes.
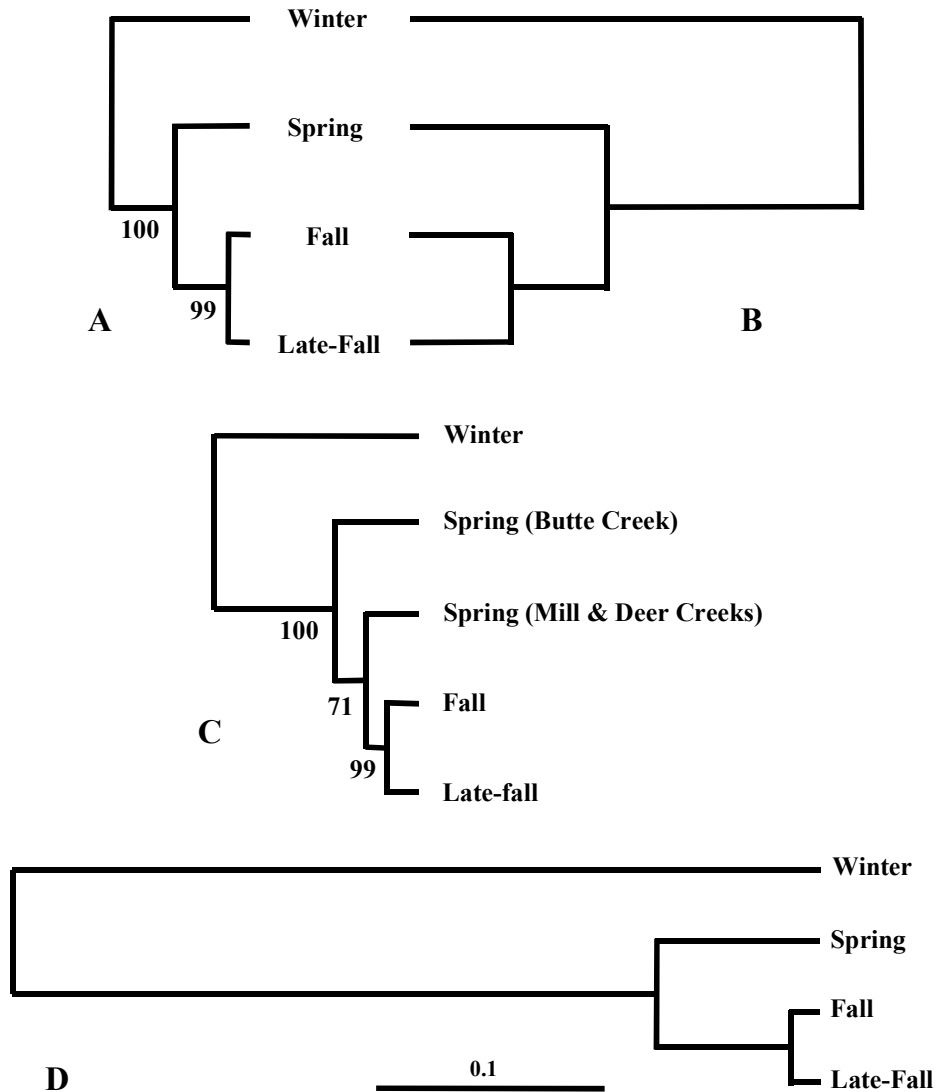
Allozymes

A study of 39 allozyme loci (Bartley and others 1992) revealed little divergence between fall and winter-run chinook salmon, with Wright's standardized allele-frequency variance, F_{ST} = 0.01. However, the authenticity of winter-run samples used in this study has been questioned (D. Teel and G. Winans, personal communication, see "Notes"). A more recent study (Winans and others forthcoming), based on more extensive sampling, indicates significant genetic structure among Central Valley chinook runs, in accord with results based on other marker types (Figure 1A).

Mitochondrial DNA

Nielsen and others (1994) reported substantial divergence in frequencies of six mtDNA haplotypes (F_{ST} = 0.24) among recognized Central Valley chinook stocks (Figure 1B). However, the probability that any two Central Valley chinook haplotypes are identical is 0.7, precluding use of this marker alone for individual identification. Further, maternal inheritance of mtDNA limits use of this marker type for genetic inference related to family structure.

Microsatellites

The listing of winter run under the federal and California endangered species acts increased the need to discriminate among subpopulations of Central Valley chinook. Banks and others (1999) cloned and developed ten new microsatellites for this task, verifying that their inheritance was Mendelian. A subsequent study used these and other microsatellites from the literature to characterize 41 population samples taken throughout the valley between 1991 and 1997 (Banks and others 2000; Figure 2). Samples encompassed geographic and temporal variation within subpopulations. Maximum likelihood methods were used to correct for family structure among samples comprised of juveniles (see "Parentage and Kinship"), as well as to correct for run admixture in adult samples (see "Avoiding Hybridization in the Winter Run Supplementation Program"). This extensive sampling and sample adjustments established a database of accurate and precise estimates of microsatellite allele frequencies for Central Valley chinook.

Figure 1 Genetic distances among subpopulations of Central Valley chinook salmon calculated from data on four genetic markers, (A) allozymes or proteins (after Winans and Teal, unpublished); (B) control region sequences of mitochondrial DNA (after Nielsen and others 1994); (C) microsatellite DNA markers (after Banks and others 2000); (D) a class II member of the Major Histocompatibility Complex (after Kim and others 1999). Numbers next to nodes in (A) and (C) are the percentages of 1,000 bootstrapped trees showing that same node.

The most important finding of this study is that chinook salmon of the Central Valley in California have substantial genetic diversity and structure (Figure 1C). Except for discovery of two distinct lineages of spring run, this study revealed a genetic structure congruent with the recognized winter, spring, fall and late-fall spawning runs (Fisher 1994). It is, perhaps, surprising but encouraging that such biological diversity has survived more than 100 years of massive habitat destruction, exploitation, and artificial propagation (Yoshiyama and others 1998, this volume). Moreover, the data retrospectively support the designation of winter run and spring run as Evolutionary Significant Units protected under the U.S. Endangered Species Act (Waples 1995; NMFS 1994, 1999). Winter run, whose blend of ocean- and stream-type life-history characteristics is unique in the species (Healey 1991), is the most distinctive of the subpopulations in the Central Valley. The next most distinctive subpopulations are the spring runs, particularly those in Butte Creek, which have unique life-history adaptations (Yoshiyama and others 1996). Formerly the most abundant chinook salmon throughout the Central Valley, spring chinook are presently found in only a few tributaries of the Sacramento River, primarily those considered in this study (Fisher 1994; Yoshiyama and others 1996, 1998). Finally, fall and late-fall runs, though closely related, are significantly different at 10 microsatellite markers (Figure 1C) and differ in geographic range, run timing, and size at maturity (Fisher 1994).

Winter run, and to a lesser extent spring run from Butte Creek, show lower levels of allelic diversity than other runs, suggesting that these populations experienced past reductions in size (bottlenecks). This may also explain a part of their divergence from the other runs in the Central Valley (Hedrick 1999). Despite spatial and temporal overlap of chinook salmon spawning runs in the Central Valley, no evidence for natural hybridization among runs was found by Banks and others (2000). A commonly held view is that most spring-run populations have hybridized with fall run and that Butte Creek spring run, in particular, has hybridized with the Feather River fall hatchery stock (Yoshiyama and others 1998). However, two observations contradict this hypothesis. First, genotypic proportions in the Butte Creek spring run mostly conform to random mating expectations. Second, Butte Creek spring clusters farther from the fall run than does spring run from Deer and Mill creeks (Figure 1C), not closer as expected under the hybridization hypothesis. Run-admixture can nevertheless occur and appears a likely cause for significant linkage disequilibrium in hatchery populations (see "Avoiding Hybridization in the Winter Run Supplementation Program") and, to a lesser extent, in samples from certain populations spawning in the wild.
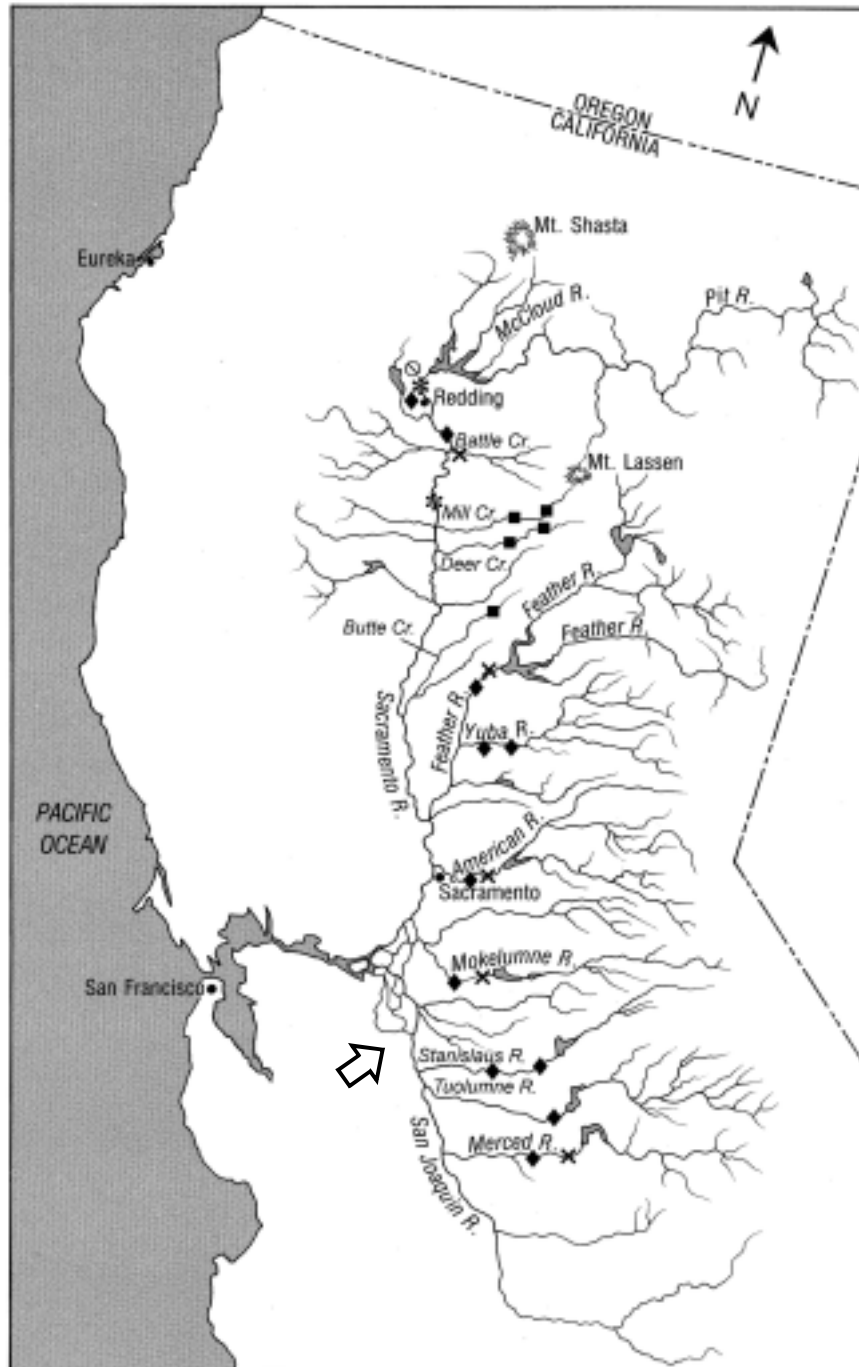
Figure 2 Map of the Central Valley, showing the localities from which chinook salmon were sampled for genetic analysis (from Banks and others 2000). The open arrow indicates the general location of the SWP and CVP water pumping plants in the Sacramento-San Joaquin Delta.

Nielsen and others (2000) also characterized Central Valley chinook using 10 microsatellites, five of which were in common with those used by Banks and others (2000). Overall relationships between major subpopulations revealed by this study were the same as described by Banks and others (2000), the two studies both verifying the distinctiveness of winter and spring runs. In contrast to Banks and others (2000), however, Nielsen and others (2000) found that year-to-year variation within runs was substantial (nearly 11% of the total variance) though not significant. Moreover, they found significant heterogeneity within fall-run hatchery samples as well as within spring run samples from Mill, Deer, and Butte creeks. However, Nielsen and others (2000) used samples of juveniles and did not correct for the potential effects of kinship within such samples.

A Class II Gene of the Major Histocompatibility Complex

Characterization of class II MHC variation for Central Valley chinook salmon also found significant frequency differences among runs (F_{ST} = 0.129) except between fall and late-fall (Kim and others 1999). Thus, in consensus with other marker types, MHC variation demonstrates the distinctiveness of the endangered winter run (Figure 1D), with no evidence for significant variation among winter run samples from different years.

Concordance Across Marker Types

The pictures of divergence among chinook salmon runs in the Central Valley painted by the above marker types are concordant. Winter run is the most distinctive subpopulation, followed by spring run, then fall and late fall. There is substantially less variation among the geographic samples within a subpopulation than among subpopulations, even for the fall run, which is presently the most widely distributed. Finally, most studies have not detected significant temporal variation within a spawning population.

Mitochondrial DNA and MHC appear, at first glance, to show greater divergence among runs than do microsatellite markers (compare Figures 1B and 1D with 1C). The average 0.078 F_{ST} estimate for 10 microsatellite loci from Banks and others (2000) is less than the F_{ST} of 0.24 from the mtDNA data of Nielsen and others (1994) or the 0.129 estimate from the MHC class II b1 exon (Kim and others 1999). However, some microsatellite markers do show comparable levels of divergence (for example, Ots-2 with F_{ST} of 0.169). Another difference among microsatellites, MHC, and mitochondrial DNA, which may account for different levels of among-subpopulation divergence, is in numbers of alleles. The last two marker types have substantially fewer alleles than is typical of microsatellites. Several researchers (Hedrick 1999 and references therein) have shown that, for highly variable loci such as microsatellites, F_{ST} is constrained by high within-population diversity. This problem can be overcome

to some extent by using different distance metrics, including the percentage of individuals correctly assigned to their sample of origin, as discussed in the next section.

Mixed Stock Analysis and Individual Assignment

Mixed Stock Analysis vs. Individual Assignment

Distinguishing among the five morphologically similar subpopulations (fall, winter, late fall, Butte Creek, and Mill and Deer Creek springs) of chinook salmon in the Central Valley is important in fisheries management and conservation, particularly because some stocks are protected and others are not. Population genetics has been applied to this problem, in several different contexts, involving adult and juvenile phases of the life cycle. Run identification is made possible by the baseline survey of microsatellite DNA variation in population samples from the Central Valley (Banks and others 2000). Two population genetic methods are used to distinguish among the different spawning runs: mixed stock analysis (MSA) and individual assignment to population of origin. MSA is a population-based method that has been widely used to estimate the relative contributions of salmon stocks to random samples of adults taken in mixed ocean harvests (Milner and others 1985; Utter and Ryman 1993). In the Central Valley, MSA can be applied to mixtures of chinook salmon juveniles from different spawning populations, which comingle in the Sacramento-San Joaquin Delta during emigration from the freshwater habitat. Individual assignment, on the other hand, estimates the most likely population of origin for an individual, based on the odds that its genotype belongs to one rather than to another subpopulation (Paetkau and others 1994; Banks and Eichert 2000). Individual assignment is useful when adults are collected for hatchery propagation or when the presence of protected runs must be ascertained in small samples from fish salvage operations at Delta pumping facilities. Actually, as we shall illustrate, a combination of the two methods is needed to analyze mixtures in the Delta and the ocean fishery.

The Central Valley chinook baseline can be used in computer simulations to illustrate the two methods and to demonstrate their relative merits and effectiveness. The baseline data are randomly permuted to produce 200 individuals from each of the five populations: winter, spring from Mill and Deer creeks, spring from Butte Creek, fall and late fall. Each individual has been genotyped for seven of the 10 markers studied by Banks and others (2000). This creates a mixed stock of 1,000 individuals of known population descent, with which to evaluate the characteristics and performance of each method. MSA uses the Statistical Package for Analysis of Mixture (SPAM, version 3.2, available at http://www.cf.adfg.state.ak.us/geninfo/research/genetics/soft-

ware/spamPage.htm). Individual assignment is performed following procedures described by Banks and Eichert (2000). Statistical power of assignment is then assessed through population simulations (Banks and others forthcoming). Results of both MSA and individual assignment are presented in Table 1. MSA accurately estimates the contributions from all runs; the actual contribution of each subpopulation, 0.2, lies within two standard errors of the estimated contribution. On the other hand, although 99.7% of simulated winter-run individuals are correctly assigned, only 60% to 80% of non-winter individuals are correctly assigned. The poorer assignment of non-winter fish is attributable to the smaller genetic distances separating the non-winter runs from one another. MSA is better at identifying the contributions of all runs because it uses not only the information present in the baseline but also the information in the mixed population sample. Individual assignment, like MSA, uses the baseline information but has only the limited information from the single individual being assigned.

Although the five subpopulations contribute equally to our example mixture, they are likely to contribute very unequally to most samples from natural populations. The accuracy of individual assignment based strictly on the likelihood of genotypes in baseline populations is affected by the relative contribution from source populations. If genotype A is relatively common in run 1 but quite rare in run 2, individuals with genotype A will be assigned to run 1 in the absence of information on the relative abundance of the two runs. However, if run 2 is 1000 times more abundant than run 1, then the likelihood that genotype A belongs to run 2 increases. Prior information on the relative abundance of runs can be used to correct the individual assignment, using Bayesian statistical methods (Shoemaker and others 1999). We shall show that MSA can provide estimates of relative run abundance that are, in turn, used to adjust the assignment.

Individual assignment for spring, fall, and late-fall populations could be improved with additional markers that increase the genetic distance among these runs. New microsatellite markers have been developed for spring-run characterization (Greig and Banks forthcoming), and additional markers for Pacific salmon are being developed by West Coast laboratories at an increasing rate. A program for evaluating the power of alternate sets of markers through re-sampling simulations (WHICHLOCI, Banks and others forthcoming) now facilitates the choice of markers needed to reach a given level of accuracy and precision of individual assignment. The cost of assigning individuals to non-winter runs will be greater, of course, than the cost of assigning winter run individuals, because more markers will be required.

Table 1 Results for assigning components of a mixed stock to population origin using mixed stock analysis and individual assignment[a]

Mixed Stock Analysis

Population	Expected	Estimate	Standard Error
Winter	0.2000	0.2009	0.0126
SP-MD	0.2000	0.2185	0.0122
SP-B	0.2000	0.1899	0.0122
Fall	0.2000	0.1874	0.0093
Late fall	0.2000	0.2033	0.0122

Individual Assignment

Population	% Correct	Standard Deviation
Winter	99.7226	0.508
SP-MD	77.5115	4.0626
SP-B	90.4935	2.9192
Fall	69.7285	4.5995
Late fall	80.0215	4.0677

[a] A mixed stock was composed of 200 individuals from each of five populations created through permutation of baseline populations. The mean, standard deviation, and standard error estimated from 1,000 bootstrap samples.

Avoiding Hybridization in the Winter-run Supplementation Program

In 1991, the U.S. Fish and Wildlife Service initiated a hatchery supplementation program aimed at helping to prevent the Sacramento River winter chinook salmon from going extinct. Research on the genetic effect of the program is described in the next section. Here, we consider a problem that became apparent in 1995, namely, how to distinguish winter run from non-winter run in selecting broodstock for the hatchery supplementation program.

In 1995, 38 of 85 fish collected by the USFWS for the winter-run supplementation program failed to mature in the hatchery. These non-maturing fish appeared to have phenotypic and genotypic affinities with spring chinook. A re-investigation of 140 winter-run brood stock that had been used for winter-run supplementation from 1991 to 1995 revealed strong, non-random associations (called gametic-phase or linkage disequilibria or LD) of allelic combinations at pairs of microsatellite loci. Typically, adults from naturally spawning populations show random associations of allelic combinations at pairs of loci, because mating of Pacific salmon occurs randomly with respect to genetic

markers (Figure 3). One significant cause of LD in samples from salmon populations, particularly hatchery populations, is admixture of non-interbreeding populations (Waples and Smouse 1991). Mixture was already evident from the spring-run affinities of non-maturing brood fish captured in 1995. The implication of finding significant levels of LD in the spawning fish was that spring run had been hybridized with winter run in the supplementation program and that possibly all samples of winter-run had actually been mixtures of two or more distinct runs.

By identifying and removing individuals with multiple, pairwise allelic combinations typical of spring run, it was possible to divide the mixture into winter and spring components, each of which is in linkage equilibrium. A multi-factorial analysis of individual genotypes confirms the separation based on analysis of LD (Figure 4). Nineteen of the 140 winter brood fish clearly cluster with 37 of 38 non-maturing 1995 brood fish (one of the non-maturing fish clusters with the true winter-run fish). The remaining 121 "true" winters show only 2% of loci-pairs with significant gametic-phase disequilibria when 5% are expected by chance (Figure 4). The winter-run baseline population now comprises these "true" winters plus samples of carcasses obtained from the Sacramento River, which were similarly purged of a few, admixed non-winters.
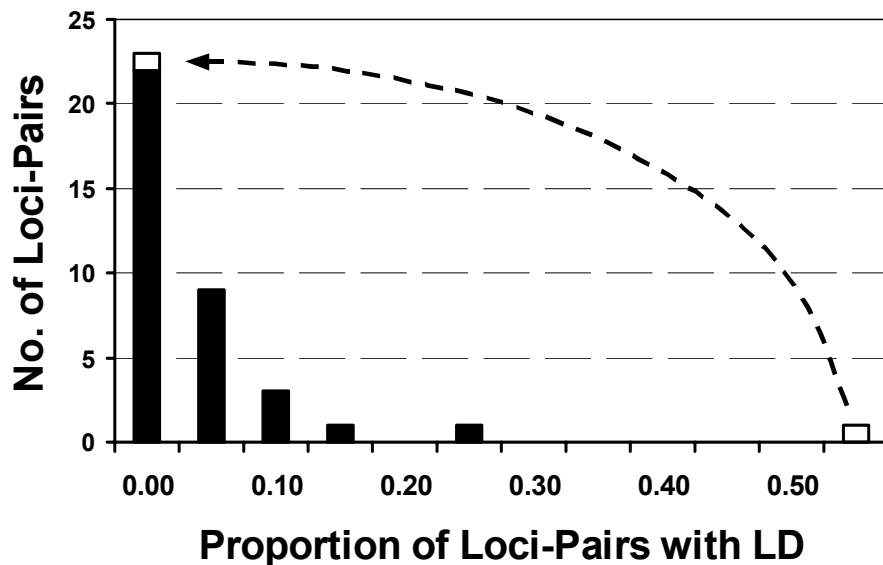
Figure 3 The proportion of loci-pairs with significant associations (linkage disequilibrium or LD) in 36 samples of non-winter chinook salmon from the Central Valley (black bars). The extremely high proportion of significant associations in winter chinook captured for a hatchery supplementation program (white bar) is greatly reduced after likely non-winter fish are removed from the sample (dotted arrow).
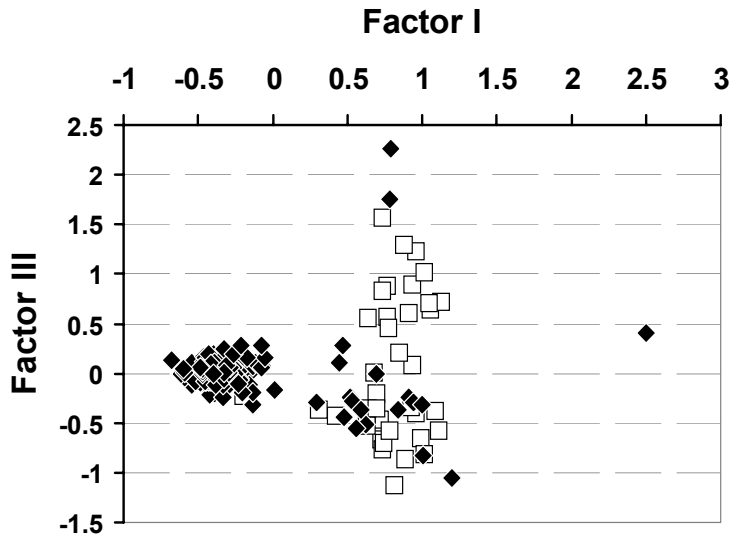
Factor I

Figure 4 Genetic clustering of chinook salmon captured for hatchery-propagation of the winter run. Scores of each fish on the first and third factors derived from factorial correspondence analysis of genotypes at 13 loci are plotted. Black diamonds denote the 140, putative winter run spawned from 1991 through 1995. White boxes denote adults captured in 1995 that did not mature and that clustered closely with spring-run populations (not shown). Note that 19 of the putative winter run adults cluster with the non-maturing, spring-run fish, while one of the non-maturing fish clusters with the true winter run.

The discovery of unwitting winter-spring hybridization in 1995, together with the observation in the same year that hatchery-spawned fish were returning to Battle Creek rather than the Sacramento River (where they had been released as fry), caused the USFWS to temporarily halt the supplementation program. The program resumed in 1998, after construction of the Livingston Stone Fish Culture Facility on the Sacramento River solved the imprinting problem and development of sufficient microsatellite markers and baseline data permitted accurate assignment of brood stock. A "rapid response" program was implemented in 1998 to genotype potential brood stock caught at the fish traps at the Keswick and Red Bluff diversion dams on the Sacramento River, as well as fish returning to the Coleman National Fish Hatchery on Battle Creek. A caudal fin clip is taken from each trapped fish and sent to the Bodega Marine Laboratory for analysis of seven microsatellite markers. Simulation results suggest that 99.1% (s.d. = 0.91%) of true winter run are correctly identified when the criterion for assignment is 10:1 or greater odds that a given genotype belongs to the winter run. More importantly, the percentage of non-winter run incorrectly assigned to winter run under this criterion is 0.02% (s.d. = 0.16%). Thus, a threshold of 10:1 or greater odds provides ample protection against incorporating non-winter run adults into the hatchery supplementa-

tion program for winter run. We typed 356 fish from the winter spawning runs of 1998 to 2000, of which 240 were assigned to the winter run (Table 2). From 1997 to 2000, we continued to monitor fish returning to the Coleman National Fish Hatchery; out of 357 examined, 108 were winters, most of which were relocated to spawning habitat in the Sacramento River.

Table 2 Numbers of chinook adults caught at the Keswick Dam (Sacramento River) and at the Coleman National Fish Hatchery (Battle Creek) subsequently genotyped and assigned to winter run

Year	Keswick Dam (Sacramento River)		Coleman National Fish Hatchery (Battle Creek)	
	Number genotyped	Number winter run	Number genotyped	Number winter run
1997	---	---	116	89
1998	152	107	117	15
1999	42	24	70	0
2000	162	109	54	4

Juvenile Emigration and Delta Salvage

We have applied both MSA and individual assignment methods to juvenile chinook emigrating from California's Central Valley. Though peak times of emigration for the different subpopulations differ, all five populations potentially intermix in the Sacramento-San Joaquin Delta (Fisher 1994). Understanding the timing of winter-run emigration and their occurrence at the State Water Project (SWP) and the Central Valley Project (CVP) is essential to evaluating the effects of these water-pumping facilities on the endangered Sacramento River winter-run chinook salmon. More than 5,000 samples were collected and genotyped over five consecutive seasons (1995–2000) at two large water pumping facilities in the Sacramento-San Joaquin Delta. In this application, in contrast to the selection of hatchery brood stock, we use an assignment criterion of even or better odds, rather than 10:1 odds, that a given genotype belongs to the winter run. The aim of this criterion is to protect all winter run at the expense of also protecting some non-winter run fish incorrectly assigned by the inclusive criterion.

The contributions of various spawning populations to the mixture of juveniles in the Delta are expected to be unequal and variable with the season. Winter run contributes a large number of samples early in the season and fewer samples later, when fall run dominates fish salvage. However, as mentioned above, the relative abundance of the various subpopulations can have a sub-

stantial effect on individual assignment. To correct for this, we use MSA to estimate the relative abundance of the runs among juveniles of similar size caught around the same time as each individual whose genotype alone suggests winter-run provenance. In other words, MSA establishes the prior probability for the runs and, using a Bayesian statistical approach, serves to correct the individual population assignment for unequal relative frequencies of subpopulation (Dean and others forthcoming). In practice, the assignment of relatively few individuals is affected by this correction (Figure 5). Having thus identified which emigrating juveniles are winter run, we see that the results do not accord with the growth model predicting the relationship of juvenile size and provenance. Winter juveniles are caught at similar sizes throughout the season of emigration, in contrast to the growth curves that presently define the subpopulations for purposes of determining take of protected winter run (Figure 5). The growth curves clearly overestimate the losses of winter-run in the Delta. These results further suggest the hypothesis that the winter run does not use the lower Delta as rearing habitat.
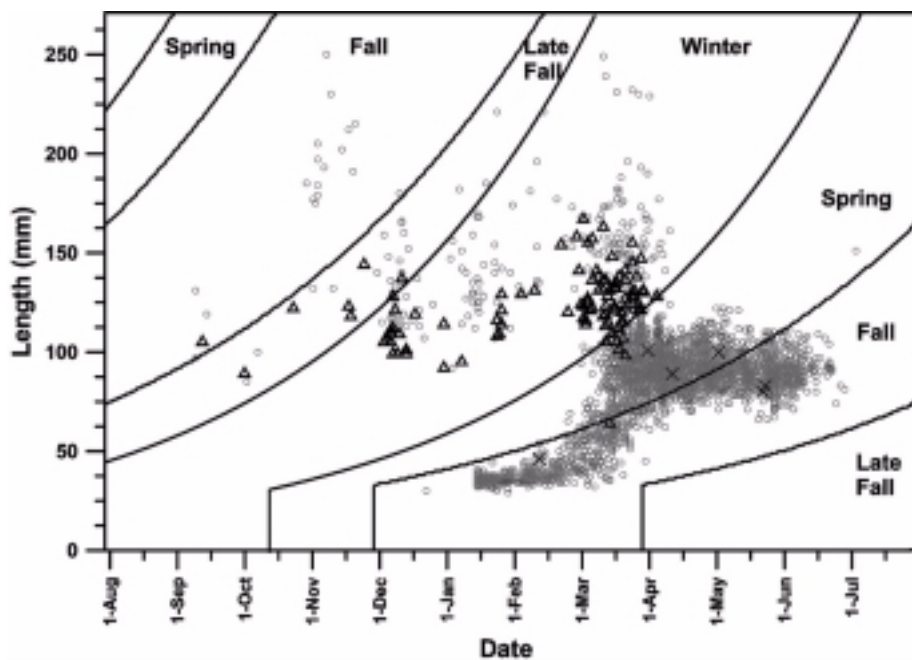
Figure 5 Size and date of salvage for 4,045 chinook juveniles genotyped between 1995 and 1999. Those individuals with greater than even odds of being assigned to the winter run, adjusted for the abundance of all runs at the time of sampling, are indicated with triangles. Six individuals, whose assignments to winter were overturned by adjustment for relative run-abundance, are indicated with an "X." All other genotyped samples are indicated with small open circles. Curved lines represent the confidence limits around the expected growth curves for each of the named runs.

Ocean Catch

Another area where the use of genetic stock identification can help protect threatened stocks is in the monitoring of ocean catches. A recent study considers data from an experimental fishery conducted for seven days (April 15–21, 1997) between Lopez and Magu points in southern California (Banks and others forthcoming). As above, both MSA and individual assignment were applied in this study, as was the Bayesian correction of individual assignment for the actual abundance of contributing stocks. Three data sources were used, microsatellites, allozymes and coded-wire tag recoveries, and all indicated a surprisingly large harvest of the endangered winter run in this short fishery (about 2%). Precise identification of protected subpopulations within watersheds, such as winter run from the Central Valley, could lead to more refined fishery management. For example, it should be possible to determine the specific conditions and/or locations that minimize the harvest of protected runs, so that a more targeted fishery on non-threatened stocks could be sustained. Real-time genetic monitoring could be used to verify run composition of harvest, and effort could be re-directed as necessary to ensure maximum harvest of chosen runs. Such use of population genetics for adaptive fisheries management could facilitate sustainable salmon harvests even in areas where threatened stocks exist.

Genetic Impact of Supplementation

Ryman-Laikre Models

Having plummeted from annual runs of nearly 100,000 fish in the late 1960s to less than 200 fish in 1991, the winter chinook was protected under both California and federal endangered species laws in the early 1990s. A hatchery supplementation program was initiated with broodstock captured from the Sacramento River and taken to the Coleman National Fish Hatchery on Battle Creek for maturation and spawning. Progeny were tagged internally with coded-wire tags, marked externally by clipping of adipose fins, and released into the Sacramento River as juveniles (smolts). Hedrick and others (1995, 2000a, 2000b) have used a demographic population genetics model (Ryman and Laikre 1991) to evaluate the potential genetic effect of this hatchery supplementation program from 1991 through 1995. One danger of hatchery supplementation is that it could dilute the gene pool by flooding the natural population with the offspring of a few individuals. However, this dilution need not occur.

The effect of hatchery supplementation on genetic diversity is mediated through effects on the effective size (N_e) of the natural population. N_e is the size of a mathematically ideal population that has rates of genetic drift and

inbreeding equivalent to those in the actual population under study. In the mathematically ideal population, there are equal numbers of both sexes, adults mate at random, and variance in number of offspring per adult is binomial or Poisson. The number of adults N in the ideal population is, by definition, equal to the effective size, and the ratio of $N_e:N = 1.0$. In actual populations, the sexes may not be in equal numbers, mating may not be at random, or the variance in offspring number may be larger than binomial or Poisson.

For a hatchery-supplemented population, N_e depends on the effective sizes of the hatchery and wild components of the population and on the relative proportion of hatchery origin fish (after Ryman and Laikre 1991):

$$N_e = \frac{N_{eh} \times N_{ew}}{x^2 N_{ew} + y^2 N_{eh}}$$

N_{eh} and N_{ew} are the effective sizes of the hatchery and wild components of the population, respectively, while x and y are their relative contributions to the total ($x + y = 1.0$). For each year, we calculate N_{eh} from data on the number of progeny contributed by each male and female brood fish to the release of juveniles. The $N_e:N$ ratio for the naturally spawning population is assumed to have a lower bound of 0.10 (Bartley and others 1992) and an upper bound of 0.33 (R.S. Waples, personal communication, see Notes). These ratios are multiplied by the run-size estimate in any year to obtain N_e before capture of adults (that is, what the effective size would have been without supplementation). The N_{ew} after capture of adults for supplementation discounts N_e by the number of adults taken to the hatchery. Estimates of the Ryman-Laikre model parameters from 1991 through 1995 for the winter-run supplementation program are given in Table 3. There are four important points to note:

1. The supplementation program likely had little, or perhaps a slightly positive effect on winter-run effective population size in all years. N_e with supplementation is higher than N_e without supplementation in all years, if $N_e:N = 0.1$; N_e with supplementation is higher than without in three of five years at $N_e:N = 0.33$ (Table 3).

2. The proportion of fish contributed by the hatchery, x, tends to be high in years when the run size was low (1994), and low when the run size was high (1992, 1995). Estimates of x are based on numbers of females, their egg production, and the survival of these progeny from egg to smolt stages. For hatchery stocks, the egg to smolt survival is esti-

mated to be 28.5%, about twice as high as estimates for egg to smolt survival in the wild, 14.7% (Hedrick and others 2000). Of course, this boost in early survival is precisely what makes hatchery supplementation such an attractive recovery option in the first place.

3. The genetic effect of supplementation depends critically on x, unless run size is very small. For example, if x in 1995 had been 10% higher, the effect would have been negative, at $N_e/N = 0.33$, rather than positive. On the other hand, in years of low run size, the hatchery program increases effective population size over a broad range of parameter combinations.

4. Ratios of effective to actual numbers of captive broodstock, $N_{eh}:N_h$, ranged from 0.47 to 0.8, much higher than the $N_e:N$ ratio assumed for the naturally spawning population (0.1 to 0.33). This boost in $N_e:N$ ratio of the hatchery component is what counterbalances the dilution of natural genetic diversity that seemingly ought to occur in a simple view of supplementation.

Table 3 Effect of hatchery supplementation on the effective size of Sacramento River winter chinook salmon, 1991-1995

Parameter	1991	1992	1993	1994	1995
Naturally spawning run size	191	1180	341	189	1361
No. taken captive (N_h)	23	29	18	29	47
No. of breeding parents ($N_f + N_m$)	15	26	12	26	42
Hatchery effective size (N_{eh})	7.02	19.07	7.74	23.2	29.2
(95% confidence interval)	(3.58, 12.22)	(12.67, 26.68)	(3.20, 13.34)	(15.9, 30.8)	(21.3, 37.8)
N_{eh}/N_h ratio	0.468	0.733	0.645	0.8	0.62
Relative contribution from hatchery (x)	0.159	0.061	0.130	0.407	0.083
N_e without hatchery (lower & upper bounds)	19.1 – 63.7	118 – 393.3	34.1 – 113.7	18.9 – 63.0	136.1 – 453.7
N_{ew} (lower & upper bounds)	16.8 – 56.0	115.1 – 383.6	32.3 – 107.7	16.0 – 53.3	131.4 - 438
N_e with hatchery (lower & upper bounds)	21.9 – 61.6	127.3 – 401.0	39 – 108.6	34.3 – 72.8	150.7 – 463.6

Since N_{eh} is based on adult contributions at release rather than at return and spawning, the above calculations are predictions of N_e. By typing microsatellite DNA markers (Banks and others 1999, 2000) on all returning, adipose fin-clipped adults, we were able to assign 93 fish from the 1994 year class to family (Hedrick and others 2000b). We found that the contributions at release of each fish spawned in the hatchery remained approximately the same at return (Table 4) and that the N_{eh} calculated for spawning adults was within the predicted 95% confidence intervals (Hedrick and others forthcoming).

As illustrated in this example, higher survival and higher N_e:N ratios of hatchery offspring, combined with contributions that are inversely proportional to the wild stock size, can increase variance effective size and conserve more of the natural biodiversity than would have been conserved in the absence of supplementation. Hatchery enhancement does not necessarily constitute a threat to genetic resources; indeed, hatchery supplementation can help to retain biodiversity that would otherwise be lost from threatened and endangered populations without intervention. However, we agree with Waples and Do (1994) that supplementation programs are likely to succeed only when the initial environmental causes of population decline are ameliorated.

Table 4 The proportions of progeny released and returning from the different female and male parents of the 1994 brood year

Female	Releases	Returns	Male	Releases	Returns
3	0.080	0.108	B	0.102	0.097
4	0.070	0.054	C	0.073	0.097
5	0.058	0.075	D	0.107	0.172
6	0.054	0.065	E	0.139	0.086
7	0.056	0.032	F	0.120	0.161
8	0.053	0.022	G	0.128	0.065
9	0.054	0.054	H	0.102	0.108
11	0.054	0.075	I	0.147	0.172
12	0.062	0.032	J	0.070	0.032
13	0.092	0.086	K	0.029	0.011
14	0.032	0.022			
15	0.079	0.108			
16	0.066	0.108			
17	0.064	0.075			
18	0.071	0.043			
19	0.057	0.043			
Total	43,346	93	Total	43,346	93

Domestication Selection?

Concern is often expressed about genetic changes in supplemented populations resulting from artificial or domestication selection for survival in the hatchery environment or from shielding of adults or hatchery-reared progeny from natural selection (for example, Waples 1999). While this is undoubtedly true for production hatchery stocks of fall chinook salmon in the Central Valley, conservation hatcheries get brood stock continually from the wild and do not typically use hatchery-reared progeny to propagate the next generation. In this case, the efficiency of selection on a single pass through a hatchery is likely to be low, especially if differential survival among families is minimized.

Equivalence in the relative proportions of winter-run families at spawning, release, and return suggests low additive genetic variance for survival in the hatchery or at sea. Moreover, data on the relative numbers of naturally spawned and hatchery fish returning to the Sacramento River, though subject to large uncertainty (Botsford and Brittnacher 1998), suggest the hatchery contribution is not consistently less at return than at release. For 1995 through 1998, the proportions of hatchery-origin winter run, at return compared to release (three years before), are 6.1% vs. 6.1%, 20.1% vs. 16.1%, and 25.0% vs. 41.7% (addendum to USFWS ESA Section 10 Permit Supplement, dated February 20, 1998). These data suggest the relative survival of hatchery and wild fish in the wild is not grossly different, given the large uncertainty in the escapement estimates. In this one example, at least, we see little evidence for selection as the result of a single pass through a supplementation hatchery. The long-term risk to diversity from over-propagating a few adults appears to far outweigh the risk from artificial selection, at least in the winter-run propagation program. This need not be the finding in other programs, however. The important point is that data on family proportions at spawning, release, and adult stages allow evaluation of the relative strengths of selection and random drift and should be required for supplementation programs.

Hybridization in Production Hatcheries

Hybridization in the Coleman National Fish Hatchery Fall Stock

Analyses of linkage disequilibrium in samples of fall and late-fall chinook stocks propagated or heavily influenced by hatcheries show higher levels of LD than typically observed in naturally spawning stocks of chinook salmon (Figure 6). The median proportion of pairwise combinations of loci showing significant LD is 0.069 for hatchery stocks and 0.025 for naturally spawning adult chinook populations. A likely explanation for this slight elevation of LD in production hatchery stocks is recent admixture and hybridization between

fall and spring or between fall and late-fall stocks in the hatchery programs for fall and late-fall chinook. Because of the high genetic similarity of these stocks, however, information from many more loci will likely be needed to test this hypothesis. New microsatellite loci being developed for the diagnosis of spring chinook may help resolve the causes of LD in production hatchery stocks.
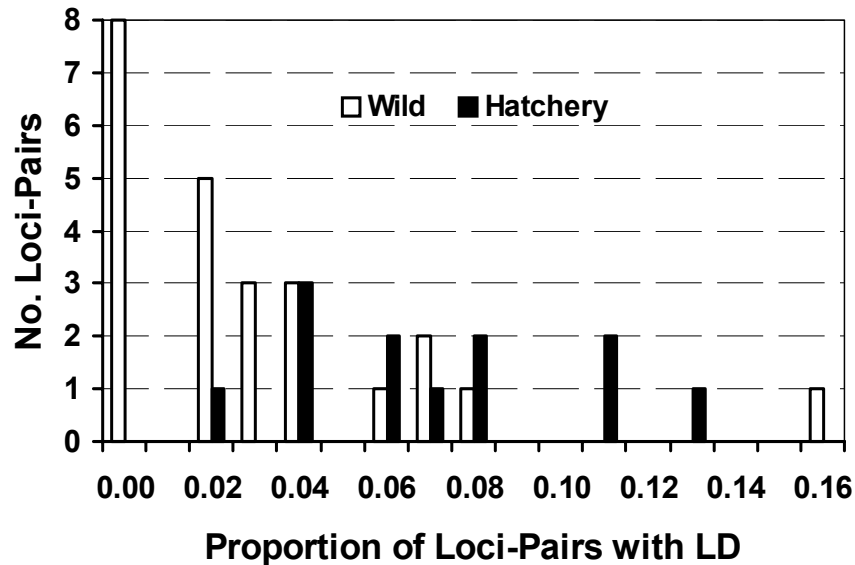

Figure 6 The proportion of loci-pairs with significant linkage disequilibrium in non-winter chinook stocks of the Central Valley. Hatchery populations (black bars) appear to have higher levels of linkage disequilibrium than naturally spawning populations (white bars). Hatchery populations include hatchery stocks as well as populations likely to be heavily affected by hatchery operations, such as late-fall in the Sacramento River. The wild population with significant LD at about one-sixth of the loci-pairs is a sample of spring run from Butte Creek that may have been contaminated with a few fall-run fish.

Hybridization of Fall-run and Spring-run in the Feather River Hatchery?

Hybridization of fall and spring run is thought to have occurred in the Feather River Hatchery, based on returns of tagged fall progeny during the spring-run spawning season and vice versa. Our analyses of samples from hatchery and naturally spawning chinook populations in the Feather River do not support this hypothesis, however. First, none of these populations shows significant linkage disequilibrium, unlike the winter and fall chinook stocks discussed above. Lack of LD suggests either that hybridization of fall with spring runs, such as those observed in Mill, Deer, and Butte creeks, has not occurred or that it has not occurred recently. Several generations of random mating fol-

lowing some past hybridization event could have reduced initial LD to non-detectable levels. Second, chinook in the Feather River cluster with the fall-run lineage in the Central Valley (Figure 7), not with the spring chinook lineages observed in Mill, Deer, and Butte creeks. This proximity of Feather River chinook to the fall-run lineage is observed when samples, whose origin is marked "unknown" by DFG collectors, are pooled after testing for and failing to find any significant heterogeneity among these samples. Still, few of the "unknown" samples can be included in the homogeneous pool of fall samples, so some slight but statistically significant genetic differentiation does exist between many of these unknown samples and fall chinook populations. The nature of this differentiation is still under investigation, but it seems not to be the result of hybridization. Finally, under the hypothesis of past hybridization followed by random mating, one might expect to see Feather River populations occupying a genetically intermediate position between fall and spring runs. Yet, there is no consistent tendency for Feather River "unknown" samples to have frequencies intermediate to fall and spring frequencies.
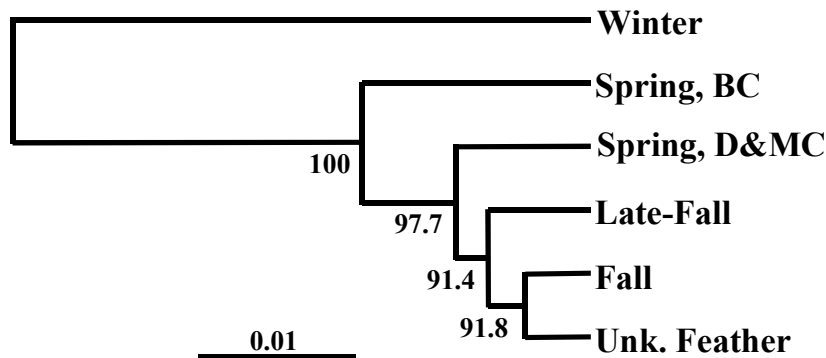
Figure 7 Clustering of Central Valley chinook samples by similarity at seven microsatellite loci shows chinook of unknown (spring?) race in the Feather River to be most closely related with fall chinook

Parentage and Kinship

One of the exciting new areas in population genetics is the application of highly polymorphic microsatellite DNA markers to questions of parentage and kinship in natural populations (O'Reilly and others 1998; Goodnight and Queller 1999; Bentzen and others 2000). These methods and markers are equally applicable to hatchery populations, in which the parents or potential parents are often known, as in the case of the winter-run hatchery supplementation program. In this case, the parents of any given progeny can be identi-

fied by simple matching algorithms; WHICHPARENT, a program facilitating such matching of progeny and parents, is available at http://www-bml.ucdavis.edu/imc/Software.html.

More difficult is ascertaining kinship when parents are unknown. In the course of our survey of variation in the Central Valley, for example, we had several samples of the threatened spring run that comprised only juveniles. In the past, population geneticists advised against using such samples because the presence of full- or half-sibs could bias allele-frequency estimates (Allendorf and Phelps 1981). Indeed, these samples showed significant departures from single locus and pairwise linkage equilibrium, compared to samples from naturally spawning adult populations. We investigated kinship in these spring-run chinook juvenile samples and attempted to estimate the allele frequencies of the adult spawning population from which they were derived (Banks and others 2000). This was done by first identifying groups of individuals showing significant odds of being full-sibs. Of the 206 individuals in these samples with sufficient genotypic information, 114 were involved in pairwise comparisons for which the hypothesis of a full-sib relationship was significantly more likely ($P < 0.01$) than the hypothesis that they were unrelated. Next, we determined the mating type or combination of parental genotypes at each locus with the maximum likelihood of producing the array of genotypes in each full-sib group. We then replaced these 114 individuals with 86 inferred parents. After adjustment of juvenile samples for kinship, the proportions of single- and multiple-locus genotypes within each conformed to random mating expectations. This procedure allowed us to use the information gained from juvenile samples in our Central Valley baseline data set.

These procedures for adjusting estimates of allele frequencies for kinship should be generally applicable to salmon molecular ecological studies. This is an active area of research, and several laboratories, including ours, are presently refining statistical approaches that will accurately recover parental genotypes from juvenile samples.

Conclusions

Population genetic analysis of highly polymorphic microsatellite DNA markers confirms the existence of genetically diverse subpopulations of chinook salmon in the Central Valley. These subpopulations correspond to the traditional seasonal runs, winter, spring, fall, and late-fall, though two distinct lineages of spring run have been identified, one in Mill and Deer creeks, the other in Butte Creek. The availability of a high quality genetic database for Central Valley chinook populations now enables identification of the run-composition of mixtures, which can occur at all stages of the life-cycle, using the traditional method of Mixed Stock Analysis (MSA). Moreover, the high

level of diversity among runs at microsatellite DNA markers enables the assignment of individuals to run with an unprecedented degree of accuracy and precision. Individual identification is useful in determining the presence of winter run at all phases of the life cycle. Confirming the run-origin of putative winter chinook brood stock is essential for the hatchery supplementation program. Identifying protected runs in the fish salvage operations at the CVP and SWP in the Delta and in ocean harvests are other important application of microsatellite DNA markers. Thus, the development and application of microsatellite DNA markers has significantly advanced knowledge of winter-run biology as well as conservation efforts. Extension of the methods developed for winter-run identification to threatened spring-run populations should now be straightforward.

References

Allendorf FW, Phelps SR. 1981. Use of allelic frequencies to describe population structure. Can J Fish Aquat Sci 38:1507–14.

Banks MA, Eichert W. 2000. WHICHRUN (Version 3.2): a computer program for population assignment of individuals based on multilocus genotype data. J Hered 91:87–9.

Banks MA, Blouin MS, Baldwin BA, Rashbrook VK, Fitzgerald HA, Blankenship SM, Hedgecock D. 1999. Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tshawytscha*). J Hered 90:281–8.

Banks MA, Rashbrook VK, Calavetta M J, Dean CA, Hedgecock D. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of chinook salmon (*Oncorhynchus tshawytscha*) in California's Central Valley. Can J Fish Aquat Sci 57:915–27.

Bartley D, Gall GAE, Bentley B, Brodziak J, Gomulkiewicz R, Mangel M. 1992. Geographic variation in population genetic structure of chinook salmon from California and Oregon. Fish Bull US 90:77–100.

Bentzen P, Olsen JB, McLean J, Seamons TR, and Quinn TP. 2000. Kinship analysis of Pacific salmon: insights into mating, homing and timing of reproduction. J Heredity. Forthcoming.

Botsford LW, Brittnacher JG. 1998. Viability of Sacramento River winter-run chinook salmon. Conserv Biol 12:65–79.

Fisher FW. 1994. Past and present status of Central Valley chinook salmon. Conserv Biol 8:870–3.

Fry DH. 1961. King salmon spawning stocks of the California Central Valley, 1949-1959. California Fish and Game 47(1):55-71.

Goodnight KF, Queller DC. 1999. Computer software for performing likelihood tests of pedigree relationship using genetic markers. Mol Ecol 8:1231–4.

Healey MC. 1991. The life history of chinook salmon. In: Groot C, Margolis L, editors. Pacific salmon life histories. Vancouver, BC: University of British Columbia Press. p 39-52.

Hedrick PW. 1999. Perspective: highly variable genetic loci and their interpretation in evolution and conservation. Evolution 53:313–8.

Hedrick PW, Hedgecock D, Hamelberg S. 1995. Effective population size in winter-run chinook salmon. Conserv Biol 9:615–24.

Hedrick PW, Hedgecock D, Hamelberg S, Croci SJ. 2000. The impact of supplementation in winter-run chinook salmon on effective population size. J Hered 9:112–6.

Hedrick PW, Rashbrook VK, Hedgecock D. Effective population size of winter-run chinook salmon based on microsatellite analysis of returning spawners. Can J Fish Aquat Sci. Forthcoming.

Kim JT, Parker KM, Hedrick PW. 1999. MHC differentiation in Sacramento River chinook salmon. Genetics 151:1115–22.

Milner GB, Teel DJ, Utter FM, Winans GA. 1985. A genetic method of stock identification in mixed populations of Pacific salmon, *Oncorhynchus* spp. Mar Fish Rev 47:1–8.

[NMFS] National Marine Fisheries Service. 1994. Endangered and threatened species; status of Sacramento River winter-run chinook salmon. Federal Register 59(2):440–50.

[NMFS] National Marine Fisheries Service. 1999. Endangered and threatened species; Threatened status for two chinook salmon evolutionary significant units (ESUs) in California. Federal Register 64(179):50394–415.

Nielsen JL, Fountain MC, Sundermeyer DR, Heine EL, Malone CL, Avelino OL. 2000. Microsatellite variation in chinook salmon spawning runs from the Central Valley, California, 1992-1997. Report to the California Department of Fish and Game. 38 p.

Nielsen JL, Tupper D, Thomas WK. 1994. Mitochondrial DNA polymorphism in unique runs of chinook salmon (*Oncorhynchus tshawytscha*) from the Sacramento-San Joaquin river basin. Conserv Biol 8:882–4.

O'Reilly PT, Herbinger C, Wright JM. 1998. Analysis of parentage determination in Atlantic salmon. Anim Genet 29:363–70.

Paetkau D, Waits L, Clarkson P, Craighead L, Strobeck C. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics 147:1943–57.

Ryman N, Laikre L. 1991. Effects of supportive breeding on the genetically effective population size. Conserv Biol. 5:325–9.

Shoemaker JS, Painer IS, Weir BS. 1999. Bayesian statistics in genetics: a guide for the uninitiated. Trends in Genetics 15:354–8.

Stone L. 1874. Report of operations during 1872 at the United States salmon hatching establishment on the McCloud River. U.S. Commission on Fish and Fisheries, Report for 1872 and 1873, Part II: 168–215. Washington, D.C.

Utter FM, Ryman N. 1993. Genetic markers and mixed stock fisheries. Fisheries 18:11–21.

Utter FM. 1991. Biochemical genetics and fishery management: an historical perspective. J Fish Biol 39(Suppl A):1–20.

Waples RS. 1991. Definition of "species" under the Endangered Species Act: application to Pacific salmon. National Oceanic and Atmospheric Administration Technical Memorandum, National Marine Fisheries Service F/NWC-194. Seattle (WA): Northwest Fisheries Science Center.

Waples RS. 1995. Evolutionary significant units and the conservation of biological diversity under the Endangered Species Act. Amer Fish Soc Symp 17:8–27.

Waples RS. 1999. Dispelling some myths about hatcheries. Fisheries 13:12–21.

Waples RS, Do C. 1994. Genetic risk associated with supplementation of Pacific salmonids: Captive broodstock programs. Can J Fish Aquatic Sci 51:310–29.

Waples RS, Smouse PE. 1990. Gametic disequilibrium analysis as a means of identifying mixtures of salmon populations. Amer Fish Soc Symp 7:439–58.

Wright S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Yoshiyama RM, Gerstung ER, Fisher FW, Moyle PB. 1996. Historical and present distribution of chinook salmon in the Central Valley drainage of California. Sierra Nevada Ecosystem Project: final report to Congress. In: University of California, Center for Water and Wildlife Resources, Davis. Assessments, commissioned reports, and background information. p 309-362.

Yoshiyama RM, Fisher FW, Moyle PB. 1998. Historical abundance and decline of chinook salmon in the Central Valley region of California. N Am J Fish Manage 18:487–521.

Yoshiyama RM, Gerstung ER, Fisher FW, Moyle PB. 2001. Historical and present distribution of chinook salmon in the Central Valley drainage of California. In: Brown RL, editor. Fish Bulletin 179: Contributions to the biology of Central Valley salmonids. Volume 1. Sacramento (CA): California Department of Fish and Game.

Notes

Teel D. (National Marine Fisheries Service, NW Fisheries Science Center, Seattle, Washington). June 1999 poster at the Coastwide Salmon Genetics meeting, University of Montana, Missoula, Montana.

Waples RS. (National Marine Fisheries Service, NW Fisheries Science Center, Seattle, Washington.) Personal communication with D. Hedgecock at December 1992 meeting of the Genetics subcommittee of the winter chinook captive breeding committee, UC Davis, with P. Hedrick, L. Botsford, and D. Hedgecock attending.

Winans G. (National Marine Fisheries Service, NW Fisheries Science Center, Seattle, Washington). June 1999 poster at the Coastwide Salmon Genetics meeting, University of Montana, Missoula, Montana.