TESTING COMPONENT CONTRIBUTIONS IN FINITE DISCRETE MIXTURES: DETECTING SPECIFIC POPULATIONS IN MIXED STOCK FISHERIES

Joel Howard Reynolds & William David Templin
Gene Conservation Lab, Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, Alaska, 99518-1599, USA

Keywords: mixed stock analysis, genetic stock identification, finite mixture analysis, component testing, compositional data analysis, power.

Synopsis

Mixed stock analysis (MSA) is used to estimate the relative contributions of distinct populations in a mixture of organisms, generally via conditional maximum likelihood estimation using a baseline of learning samples from all potentially contributing populations. MSA is increasingly used to judge the presence or absence of specific populations in specific mixture samples. This is commonly done by inspecting the marginal bootstrap confidence interval of the contribution of interest. This method suffers from a number of major statistical deficiencies, including zero power to detect even a perfectly identifiable population at the low contribution levels of interest. In contrast, the likelihood ratio test has 100% power to detect any positive contribution from this ideal population. Both methods are compared in a power analysis using a 17-population baseline of sockeye salmon (*Oncorhynchus nerka*) from the Kenai River, Alaska, watershed.

Power to detect a contribution varies with the population(s) relative identifiability, contribution level, mixture sample size, and analysis method. The power analysis shows the likelihood ratio is more powerful than the bootstrap method, with equality only at 100% power. Power declines for both methods as contribution declines, but the bootstrap method declines faster and goes to zero. Power quickly declines for both methods as population identifiability declines, though the likelihood ratio test is able to capitalize on the presence of 'perfect identification' characteristics, such as private alleles in genetic markers. Given the baseline-specific nature of detection power, MSA researchers are encouraged to conduct a priori power analyses.

Introduction

Mixed stock analysis (MSA) is used to estimate the relative contributions of distinct populations in a mixture of organisms. This is an important tool in wildlife management and research, with genotypes commonly used as natural markers to distinguish major populations or stocks (e.g., genetic stock identification) (Begg et al. 1999, Pearce et al. 2000). Increasingly, MSA is used to judge the presence or absence of specific stocks in specific mixture samples. For example, management of an interception fishery may be heavily influenced by the presence of a specific threatened, weakened, or politically high profile stock (e.g., salmon of Canadian origin harvested by Alaskan fisheries).

MSA can overestimate stocks contributing little or nothing to a mixture (Pella & Milner 1987). Managers and researchers require a method for testing whether a specific nonzero stock contribution is really a biased estimate of zero. In practice, one checks the limit of the contribution's 95% bootstrap lower confidence interval: a stock contribution with zero lower interval limit is deemed 'statistically indistinguishable from zero' (Seeb & Crane 1999).

This method is statistically flawed. Using the interval as a test assumes that the contribution estimate is a pivotal statistic, which it is not. Further, such marginal tests implicitly employ an inappropriate measure of distance between compositions (Aitchison 1992). Of greatest practical importance, the method is shown, below, to have very low or even zero power exactly in the settings of interest.

Consider an ideal marker and an ideal population: a gene for which Population A is fixed for an allele that is unique among the other populations in the baseline – a private allele. Population A is perfectly identifiable, so a mixture sample of size n containing π_A times n Population A individuals will produce a nonzero contribution estimate, $\hat{\pi}_A > 0$. In this scenario, bootstrap resampling to estimate the marginal confidence interval of π_A is equivalent to sampling from a Binomial (n, π_A). As $\pi_A \Rightarrow 0$, Prob($\hat{\pi}_A = 0$) = Prob($X = 0 \mid X \sim Bin(n, \pi_A)$) $\Rightarrow 1$, hence Prob(bootstrap confidence interval's lower limit = 0) $\Rightarrow 1$ (**Appendix**). Thus, even in the ideal case this method has at most moderate power to detect small contributions (Table 1); the power will be much lower in applications with less than ideal identifiability (**Results**).

One might consider abandoning MSA for this question and using an analysis that directly assigns each observation in the mixture sample to the 'most probable' population of origin (for example, Banks & Eichert 2000, Cornuet et al. 1999). Such individual assignment methods ignore information

coming from the joint distribution of characteristics observed in the mixture, are therefore less powerful than MSA methods (Millar 1987) in estimating population contributions, and hence likely to be less powerful in detecting nonzero population contributions.

Table 1. Maximum power to detect nonzero contributions in MSA using the bootstrap confidence interval method, as a function of true population contribution π and mixture sample size. Calculations assume individuals from the population of interest are perfectly identifiable in the mixture sample (**Appendix**). Power will decline as identifiability declines (Fig. 2). Calculations are based on 1000 nonparametric bootstrap resamples and a 95% one-sided percentile bootstrap confidence interval (Davison & Hinkley 1997).

	Mixture Sample Size						
π	50	75	100	150	200	300	350
0.10	1	1	1	1	1	1	1
0.05	0	1	1	1	1	1	1
0.04	0	0.66	1	1	1	1	1
0.03	0	0	0.62	1	1	1	1
0.02	0	0	0	0.58	1	1	1
0.01	0	0	0	0	0	0.54	1

The likelihood ratio provides a more powerful test of $H_0: \pi_A = 0$ vrs $H_A: \pi_A > 0$. We briefly review the standard MSA model and estimation method, conditional maximum likelihood estimation (Pella & Milner 1987, Millar 1987), develop the likelihood ratio test (**Methods**), and describe how to estimate P values. The method is demonstrated in a simulation study of mixtures of sockeye salmon (*Oncorhynchus nerka*) from the Kenai River, Alaska (**Demonstration**). The likelihood ratio test and the confidence interval method are compared in terms of power to detect a nonzero contribution from a specific population or group of populations. The comparison demonstrates how to conduct a priori power analyses for a given baseline, a specific stock of interest, and a range of stock contributions and mixture sample sizes.

The likelihood ratio test is more powerful than the bootstrap confidence interval, though both methods display lower power than desired. Both methods lose power as population identifiability and contribution decline, but the bootstrap method loses power faster, displaying zero power in some scenarios. In the absence of perfect identifiability, the likelihood method's low power limits application to those problems involving small to moderate sized baselines of potentially contributing populations.

Methods
The finite mixture model
The following model describes mixtures of contributions from a finite number of source populations (see, for example, Pella & Milner 1987 or Millar 1987). Although the presentation assumes discrete characteristics are observed on each individual, such as a genotype, this is not essential; the model holds for continuous characteristics as well.

Randomly sample n individuals from a mixture of J populations. Let the jth population contribute an unknown proportion $\pi_j >= 0$ to the mixture, $\Sigma \pi_j = 1$; $\Pi = (\pi_1, ..., \pi_J)$. If the characteristic measured on the ith sample observation is denoted by x_i, then the probability of observing the sample $\mathbf{X} = \{x_1, x_2, ..., x_n\}$ is:

$$\Pr(\mathbf{X}|\mathbf{\Pi}, \Phi) = \prod_{i=1}^{n} \Pr(x_i|\mathbf{\Pi}, \Phi) = \prod_{i=1}^{n} \left\{ \sum_{j=1}^{J} \pi_j \Pr(x_i|\mathbf{\Phi}_j) \right\}$$

with Φ_j the column vector of parameters specifying the probability density function of the characteristic in population j and Φ the matrix $[\Phi_1|...|\Phi_J]$. For a discrete characteristic with k possible outcomes, $\Phi_j = (\phi_1^j, ..., \phi_k^j)$. This assumes that the set {Pop. 1, Pop. 2,..., Pop. J} includes all potentially contributing populations (see Smouse et al. 1990). Expanding $\Pr(x_i|\Phi_j)$ allows for multivariate characteristics.

Identifiability of the mixture requires that the probability density functions of the characteristics differ across the contributing populations (Redner & Walker 1984). Characteristics commonly used in fisheries include parasite assemblages (Moles & Jensen 2000), scale patterns (Marshall et al. 1987), morphometrics and meristics (Fournier et al. 1984), artificial tags such as thermal marks, coded wire tags, or fin clips (Ihssen et al. 1981), and, increasingly, genetic markers (Seeb & Crane 1999; Ruzzante et al. 2000).

Estimation
Estimating the mixture proportions, Π, requires information regarding the characteristic probability density function, Φ_j, for each contributing population. This is generally available in the form of a sample from each baseline population. In most fisheries applications researchers fix the nuisance parameters, Φ_j, at their estimates from the baseline samples, $\hat{\Phi}_j$ (Millar 1987). Maximum likelihood is then used to estimate the unknown Π conditional on $\Phi_j = \hat{\Phi}_j$. This is justified by the relatively small amount of information on Φ_j in the mixture sample relative to the baseline sample (Milner et al. 1981). Bayesian methods also exist (Pella & Masuda 2001).

Uncertainty in the mixture proportion estimates, $\hat{\Pi}$, arises from sampling uncertainty in both the mixture and the population baselines. In practice, these sampling uncertainties can be accounted for by nonparametric bootstrap resampling of the mixture sample and parametric bootstrap

resampling of the baseline characteristic distributions, $\hat{\Phi}_j$. Bootstrap resampling the baseline widens the confidence intervals, reducing the power to detect nonzero contributions. The following demonstration only resamples the mixture sample.

Testing population contribution

Assume a sample from a mixture consisting of contributions from a known set of baseline populations and specific interest in testing $H_o: \pi_A = 0$ for Population A. The likelihood ratio test compares the likelihood of the observed sample under the general model in which $\pi_A > 0$ to the likelihood under the null model in which $\pi_A = 0$. The likelihood ratio test statistic, conditional on $\Phi_j = \hat{\Phi}_j$, is:

$$LR = \frac{L(\{\pi_1, \pi_2, ..., \pi_J\} \mid \mathbf{X}, \hat{\Phi})}{L(\{\pi_1', \pi_2', ..., \pi_A' = 0, ..., \pi_J'\} \mid \mathbf{X}, \hat{\Phi})} =$$

$$\prod_{i=1}^{n} \left\{ \sum_{j=1}^{J} \pi_j \Pr(x_i \mid \hat{\Phi}_j) \right\} \Big/ \prod_{i=1}^{n} \left\{ \sum_{\substack{j=1 \\ j \neq A}}^{J} \pi_j' \Pr_j(x_i \mid \hat{\Phi}_j) \right\}$$

with $\{\pi_1, ..., \pi_J\}$ and $\{\pi_1', ..., \pi_J'\}$ replaced by their conditional maximum likelihood estimates under their respective models. The observed ratio, LR^{obs}, is calculated by (i) fitting the mixture sample using the full baseline (the general model, the numerator), then (ii) fitting the mixture sample using the reduced baseline with Population A dropped to force $\pi_A = 0$ (the null model, the denominator). The test can be extended to the joint contribution of a specific group of populations, $H_o: \pi_1 = \pi_2 = ... \pi_v = 0$ vrs $H_A: \pi_1$ or π_2 or $...\pi_v \neq 0$.

In applications using genetic markers, if Population A has private alleles that also occur in the mixture sample, the likelihood under the reduced baseline will be zero, giving a likelihood ratio of ∞. Commonly used MSA software (Debevec et al. 2000) generally assigns such individuals to an 'unknown' baseline component, clearly identifying the nonzero contribution of Population A.

Under the null model and regularity conditions, -2 times $\ln(LR^{obs})$ is asymptotically distributed as a χ^2 with degree of freedom equal to the number of populations being simultaneously tested for zero contribution (Stuart et al. 1999). Regularity conditions breakdown when any population in the null model contributes 0, which is often the case. Even when all null model mixture populations are expected to have nonzero contributions, experience suggests that the asymptotic results may be less than reliable in this setting. The null reference distribution can be approximated by Monte Carlo simulation under H_o, conditional on Π_0 as estimated from fitting the null model (Davison & Hinkley 1997).

Demonstration

The two methods were compared in terms of their power to detect a nonzero population contribution. The simulation study used a baseline of nineteen allozyme markers for the sockeye salmon populations of Kenai River, Alaska (details in Seeb et al. 2000). To explore how population identifiability influences detection power, we explore three population sets that display declining, though relatively high, identifiability.

Kenai River sockeye baseline

The Kenai River watershed is the major producer of sockeye salmon in Cook Inlet, supporting a commercial fishery (in the inlet), a personal use fishery (at the river mouth), and a recreational fishery (within the river itself). The inriver fisheries are managed to allow a set range of individuals at each spawning ground. Resource managers and researchers are interested in detecting the presence of specific populations at time points throughout the fishing season.

The Kenai River baseline consists of seventeen populations aggregated into five inriver reporting regions based on their identifiability for mixed stock analyses[1] (Fig. 1). Three scenarios were investigated: detecting the highly identifiable Upper Russian River region (two populations), the moderately identifiable Trail Lakes region (three populations), and the somewhat less identifiable Tern Lake region (Fig. 1).

The Upper Russian River drainage is above a waterfall, a partial barrier to upstream movement. The populations spawning above the falls are relatively genetically distinct (Seeb et al. 2000), though they do not exhibit private alleles at the allozyme markers considered. The Railroad Creek population in the Trail Lakes region exhibits a private allele at one locus (relative frequency 0.0125); the Tern Lake population exhibits a private allele at another locus (relative frequency 0.01).

Simulated mixtures

For each reporting region of interest, mixture samples of genotypes from n=200 individuals were simulated over a range of regional contributions, $\pi = \{10\%, 5\%, 4\%, 3\%, 2\%, 1\%\}$. For the Upper Russian River or Trail Lakes, π was evenly split among the region's populations. The remaining

---

[1] Reporting regions must demonstrate a 90% or better mean contribution estimate for simulated mixtures constructed 100% from populations in the reporting region. For example, see Seeb et al. (2000).

populations in the baseline evenly contributed the rest of the mixture. A contribution from a given population was simulated by randomly generating a genotype from that population's allele frequencies for each of the nineteen allozyme markers. Baseline allele frequencies are available in Seeb et al. (2000). Fifty mixtures were simulated for each region of interest by contribution scenario.

Analyses

Each mixture sample was analyzed to estimate: the reporting region contributions under both the full and reduced baseline models, the 95% one-sided percentile bootstrap confidence interval for π under the full baseline model, and all quantities required to conduct the likelihood ratio test of H_0: $\pi = 0$ vrs H_A: $\pi > 0$. The bootstrap confidence interval used 1000 resamples; for equal numerical accuracy, the Monte Carlo approximation to the null reference distribution used 1000 simulations.

The bootstrap and likelihood ratio test were compared in terms of their power to detect the nonzero contribution of the region of interest. For a given scenario – method by region of interest by contribution, power was estimated as the percentage of the fifty simulated mixture samples for which the method detected a nonzero contribution from the region of interest. Detection was defined as: bootstrap method - nonzero limit, when rounded to two significant digits, on the 95% one-sided confidence interval, likelihood ratio – P value of $\leq$ 0.05 or nonzero contribution assignment to the 'unknown' category when fitting the mixture using the reduced baseline model. Baseline allele frequencies were not resampled.

Mixture samples were generated using S-Plus 2000 (Insightful, Inc., Seattle, WA., USA) and locally written functions. Mixture analyses were conducted using the freeware package SPAM 3.5 (Reynolds 2001).

Demonstration Results

The likelihood ratio test was as or more powerful than the bootstrap confidence interval method (Fig. 2), detecting at least every contribution the bootstrap method detected. Equality occurred only when both displayed 100% power. Both methods displayed less than 'ideal identifiability' power (versus Table 1 or 100% power for the likelihood ratio). However, the likelihood ratio always displayed positive power.

Both methods lost power as identifiability declined (Fig. 2, roughly decreasing down the column), or true contribution declined. The likelihood ratio's power did not decline as quickly as that of the bootstrap confidence interval.

Discussion

Increased usage of mixed stock analysis (MSA) has increased demand for methods of detecting small nonzero contributions from a specific population. The likelihood ratio test is a more powerful method than the bootstrap approach, detecting every contribution the bootstrap method detected and more (Fig. 2) while maintaining a positive power in every scenario.

When the population(s) of interest is highly identifiable, both methods are capable of 100% power even for moderately small contributions (Upper Russian, Fig. 2). The likelihood ratio retains this power to detect even a single individual in a mixture when that individual displays a characteristic seen only in its source population, e.g., a private allele. A substantial portion of the Trail Lakes and Tern Lake contribution detections involved such rare alleles. The current demonstration does not allow us to judge whether or not a contribution would have been detected in the absence of a rare allele, though the Russian River results clearly demonstrate that private alleles are not required for high power.

Power can be greatly reduced by a reduction in identifiability (Russian River vrs Trail Lakes or Tern Lake, Figs. 1 & 2). As the population of interest, call it Population A, becomes similar to other members of the baseline, the few individuals actually contributed by Population A may be adequately explained as having originated from the other populations. When these populations themselves contribute to the mixture, such as in the current demonstration, parsimony leads to absorption of Population A's contribution into that of the other populations.

Absorption is prevented if Population A has either sufficiently distinct characteristics, allowing clear detection of even a single contribution (e.g., private alleles or the Russian River scenario – Fig. 2), or has a large enough contribution to the mixture sample such that the likelihood ratio test detects the signal in the sample's joint distribution of characteristics. It is sobering to note how easily a contribution can be absorbed, that is, how low the power to detect a nonzero contribution can be. The power observed here, with a baseline of only seventeen populations, will likely decrease for larger baselines. This suggests (i) neither method will be effective in applications involving large baselines, such as samples from high seas mixtures or large-scale interception fisheries; (ii) power may improve using markers with improved identifiability, i.e., more polymorphic loci.

Bootstrap confidence intervals are obviously highly flexible and informative tools. However, this flexibility does not eliminate the need to consider the

underlying statistical issues of the application at hand. In fact, this flexibility makes it even more important that the user consider these issues.

Researchers can investigate the power to detect specific nonzero contributions in a mixture sample of size n using a baseline of interest by repeating the above process. Such a priori analyses allow one to determine the sample size required to detect a given contribution with a given power, as well as compare methods in a specific context. Note that posterior power analyses, while unfortunately rather common, are generally uninformative and, therefore, misleading (Hoenig & Heisey 2001). A priori power analysis methods in the general context of MSA are discussed elsewhere[2].

Acknowledgements

Fig. 1 – UPGMA dendrogram of Cavalli-Sforza and Edwards genetic distance among the Kenai River sockeye salmon baseline populations (Weir 1996). Populations were aggregated into five reporting regions for mixed stock analysis, where a reporting region is the smallest set of populations that achieves, on average, a 90% or greater contribution estimate for simulated mixtures consisting 100% of individuals from the region's populations. Left to right: Upper Russian River (2 populations, 99.9% contribution estimate), Hidden Creek (1 population, 99%), Trail Lakes (3 populations, 95%), Tern Lake (1 population, 90%), Kenai / Skilak Lakes (10 populations, 94%).
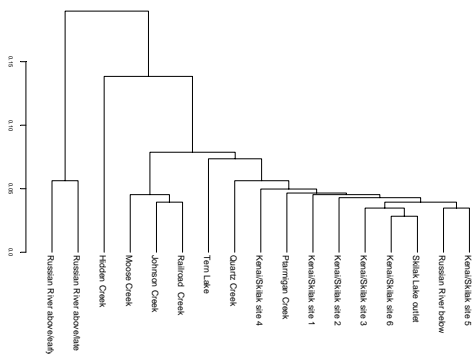

Fig. 2 - Power to detect nonzero contributions, by method (line type), for each region of interest (panel),

across a range of contributions (x axis). Fifty mixtures of size 200 were simulated for each contribution level by region of interest combination; see text for details. Power estimates are the proportion of each scenario's simulated mixtures in which the method detected the contribution.
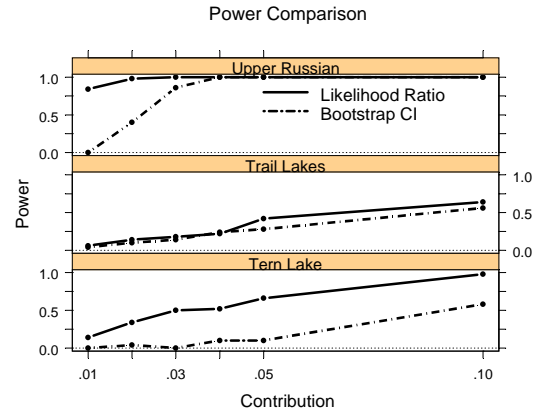
References

Aitchison, J. A. 1992. On criteria for measures of compositional difference. Math. Geol. 24: 365-379.

Banks, M.A. & W. Eichert. 2000. WHICHRUN (Version 3.2) a computer program for population assignment of individuals based on multilocus genotype data. Jrnl. of Heredity 91:87-89.

Begg, G. A., K. D. Friedland & J. B. Pearce. 1999. Stock identification – its role in stock assessment and fisheries management: a selection of papers presented at a symposium of the 128th annual meeting of the American Fisheries Society in Hartford, Connecticut, USA, 23-27 August 1998. Fish. Res. 43: 1-3.

Cornuet, J. M., S. Piry, G. Luikart, A. Estoup & M. Solignac. 1999. New Methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153: 1989 – 2000.

Davison, A. C. & D. V. Hinkley. 1997. Bootstrap Methods and their Application, Cambridge University Press, Cambridge, UK.

Debevec, E. M., R. B. Gates, M. Masuda, J. J. Pella, J. H. Reynolds & L. W. Seeb. 2000. SPAM (version 3.2): Statistics Program for Analyzing Mixtures. Jnl. Hered. 91: 509–511.

Fournier, D. A., T. D. Beacham, B. E. Ridell & C. A. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Can. Jrnl. Fisheries & Aquatic Science 41: 400-408.

Hoenig, J. M., & D. M. Heisey. 2001. The abuse of power: the pervasive fallacy of power calculations

---

[2] Reynolds, J. H. & P. A. Crane. In preparation. Power analysis for mixed stock analysis.

of data analysis. The American Statistician 55 (1): 19–25.

Ihssen, P. E., H. E. Booke, J. M. Casselman, J. M. McGlade, N. R. Payne & F. M. Utter. 1981. Stock identification: materials and methods. Can. Jrnl. Fisheries & Aquatic Science 38: 1838-1855.

Marshall, S., D. Bernard, R. Conrad, B. Cross, D. McBride, A. McGregor, S. McPherson, G. Oliver, S. Sharr & B. Van Allen. 1987. Application of scale patterns analysis to the management of Alaska's sockeye salmon (*Onchorhynchus nerka*) fisheries. pp. 207–326. *In*: H. D. Smith, L. Margolis, & C. C. Wood (ed.) Sockeye Salmon *(Oncorhynchus nerka)* Population Biology and Future Management, Canadian Special Publication on Fisheries and Aquatic Science 96.

Millar, R. B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Can. Jrnl. Fisheries & Aquatic Science 44: 583-590.

Milner, G. B., D. J. Teel, F. M. Utter & C. L. Burley. 1981. Columbia River stock identification study: validation of genetic method. Final report of research (FY80) financed by Bonneville Power Administration Contract DE-A179-80BP18488, National Marine Fisheries Service, Northwest and Alaska Fisheries Center, Seattle, WA.

Moles, A. & K. Jensen. 2000. Prevalence of the sockeye salmon brain parasite *Myxobolus arcticus* in selected Alaska streams. Alaska Fish. Res. Bull. 6: 85-93.

Pearce, J. M., B. J. Pierson, S. L. Talbot, D. V. Derksen, D. Kraege & K. T. Scribner. 2000. A genetic evaluation of morphology used to identify harvested Canada geese. Jrnl of Wildlife Management 64: 863-874.

Pella, J. J. & M. Masuda. 2001. Bayesian methods for stock-mixture analysis from genetic characters. Fish. Bull. 99: 151-167.

Pella, J. J. & G. B. Milner. 1987. Use of genetic marks in stock composition analysis. pp. 247 – 276. *In*: N. Ryman & F. Utter (ed.) Population Genetics and Fishery Management, Washington Sea Grant Program, Seattle, WA, USA.

Reynolds, J. H. 2001. SPAM Version 3.5: user's guide addendum. Addendum to special publication no. 15, Alaska Dept. of Fish and Game, Commercial Fisheries Division, Gene Conservation Lab, 333 Raspberry Road, Anchorage, Alaska, 99518. Available http://www.cf.adfg.state.ak.us/geninfo/research/genetics/Software/SpamPage.htm.

Redner, R. A. & H. F. Walker. 1984. Mixture densities, maximum likelihood and the EM algorithm. Soc. for Industrial and Applied Mathematics Review, 26: 195-239.

Ruzzante, D. E., C. T. Taggart, S. Lang & D. Cook. 2000. Mixed-stock analysis of Atlantic cod near the Gulf of St. Lawrence based on microsatellite DNA. Ecol. Appl. 10: 1090-1109.

Seeb, L. W. & P. A. Crane. 1999. Allozymes and mitochondrial NDA discriminate Asian and North American populations of chum salmon in mixed-stock fisheries along the south coast of the Alaska Peninsula. Trans. Am. Fish. Soc. 128: 88-103.

Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R. Z. Davis, L. K. Brannian & J. E. Seeb 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the *Exxon Valdez* Oil Spill. Trans. Am. Fish. Soc. 129: 1223-1249.

Smouse, P. E., R. S. Waples & J. A. Tworek. 1990. A genetic mixture analysis for use with incomplete source population data. Can. Jrnl. Fisheries & Aquatic Science 47: 620-634.

Stuart, A., J. K. Ord & S. Arnold. 1999. Kendall's Advanced Theory of Statistics Vol 2A: Classical Inference and the Linear Model (6th ed.), Oxford University Press, New York, NY, USA.

Weir, B. S. 1996. Genetic data analysis II, Sinauer Associates, Sunderland, MA., USA.

Appendix 1 Ideal Power to detect nonzero contributions with one-sided bootstrap confidence intervals.

Let Population A contribute a proportion $0 < \pi < 1$ of uniquely identifiable individuals to the independent random mixture sample of size n. The bootstrap mixture resample X then follows a Binomial (n, π), with Prob(X=0) = $(1 - \pi)^n$, so Prob(limit of the one-sided bootstrap lower confidence interval for $\pi = 0$| 95% percentile confidence interval, 1000 resamples, π, n) = Prob(The number of resamples with X=0 is $\geq 50 \mid X \sim \text{Binomial}(n, \pi)$) =

Prob($R \geq 50 \mid R \sim \text{Binomial}(1000, (1 - \pi)^n)$ =

$1 - \text{Prob}(R < 50 \mid R \sim \text{Binomial}(1000, (1 - \pi)^n) =$

$$1 - \sum_{R=0}^{49} \binom{1000}{R} p^R (1-p)^{1000-R} \qquad (\dagger).$$

The power to detect $\pi > 0$ using the one-sided bootstrap confidence interval method on this mixture sample is therefore

$$1 - (\dagger) = \sum_{R=0}^{49} \binom{1000}{R} (1-\pi)^{nR} (1-(1-\pi)^n)^{1000-R}.$$