Nuclear and mitochondrial SNPs provide high-throughput resolution for migratory studies of Chinook salmon

Christian T. Smith, William D. Templin, James E. Seeb and Lisa W. Seeb

Gene Conservation Laboratory
Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, Alaska 99518, USA

Abstract

Much is known of the freshwater life history and ecology of Chinook salmon; far less is known of the oceanic migration patterns and relative marine survival of individual stocks. Migration following stock-specific corridors may lead to differing marine survival and varying rates of return among stocks during periods of fluctuating marine conditions. We are developing markers based on single nucleotide polymorphisms (SNPs) to rapidly genotype large number of individual from high seas samples. Nuclear and mitochondrial SNPs provide useful and complementary information. We compare SNPs from representative populations originating from throughout the range of Chinook salmon to equivalent data from allozymes and microsatellite markers. Results to date show that SNP markers have the potential of becoming a rapid and cost effective approach to the analysis of large numbers of samples from complex mixtures.

Introduction

Studies of the ecological genetics of fish may require the analysis of thousands of individuals at many loci to characterize populations and estimate the composition of complex fisheries. Often results from fishery analyses are desirable under very tight time-constraints.

Several genetic markers described in the literature provide information potentially of great use to management. Many of these markers are never exploited for conservation or management, because assays are too slow and/or expensive.

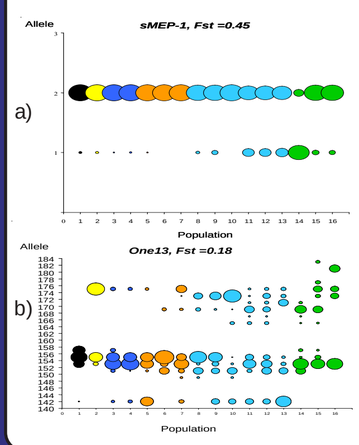
Methods

Ten 5'-nuclease reaction SNP genotyping assays (Fig 2) were developed in order to facilitate large-scale examination of several loci known *a priori* to be informative in Chinook salmon on smaller scales (Table 1). These 10 assays were tested on 1473 individuals from 16 sites around the Pacific basin (Fig 4). Data from the SNP loci were compared to data from other genetic markers commonly used for high-throughput fishery applications: allozymes and microsatellites. Microsatellite loci examined in this study were: *One9*, *One102*, *One13*, *Ots100*, *Ots107* and *µsat73*. A total of 29 allozyme loci were also analyzed in all samples. Genetic distances based on the three different marker types were compared graphically with multidimensional scaling plots and statistically by examining correlations between matrices of genetic distances between all populations. Statistical significance of each correlation was tested using a permutation procedure[1].

Table 1 Loci for which SNP assays were developed:

Assay Name	Locus description	Published assay type
ADFG|Ots_C3N3	Mitochondrial cytochrome oxidase III	PCR-RFLP[2]
ADFG|Ots_Prl2	Prolactin II	DNA sequence[3]
ADFG|Ots_P53	P53	PCR-RFLP[4]
ADFG|Ots_GH2	Growth hormone II intron D	PCR-RFLP[5]
ADFG|Ots_MHC1	Major histocompatibility complex class I α2 domain	DGGE[6]
ADFG|Ots_MHC2	Major histocompatibility complex class II β1 domain	DNA sequence[7]
ADFG|Ots_Ots2	Microsatellite *Ots2* flanking region	DNA sequence[8]
ADFG|Ots_Tnsf	transferrin	DNA sequence[9]
ADFG|Ots_SL	somatolactin	PCR-RFLP[10]
ADFG|Ots_P450	Cytochrome p450A	PCR-RFLP[10]

Fig 1. Allozyme and microsatellite alleles both distinguished known lineages of Chinook salmon



Allele frequencies for an allozyme locus (1a) and a microsatellite locus (1b). Populations (x-axis) correspond to those shown in Fig 4. The size of each bubble represents the frequency of the corresponding allele. Colors indicate larger geographic regions.

Fig 2. Genotyping without gels

The genotyping assays used for this study utilize the 5'-exonuclease activity of DNA polymerase to digest allele-specific probes in the course of the PCR[11]. The use of fluorogenic probes in a real-time PCR machine allowed amplification and genotyping to take place simultaneously[12] without any need for electrophoresis.

Components of the 5'-exonuclease reaction:

Template DNA: T/C

Allele-specific oligonucleotide probes:

Fluorescent tag — Quencher (prevents Fluorescence while in proximity of tag)

DNA polymerase featuring 5' ➡3' exonuclease activity:

Usual PCR reagents: buffer, dNTPs, etc...

1) probe matches template, appropriate fluorescence emitted

2) probe does not match template, no fluorescence emitted

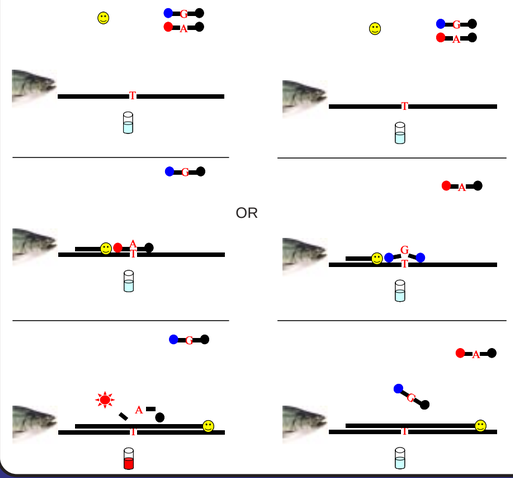


Fig 3. Rapid analysis of raw SNP data

Following amplification via the PCR, it takes less than 5 minutes to genotype a plate of 384 wells (fish). The axis on the resulting scatter plot indicate how much of each allele-specific probe was cleaved. Each dot represents one fish, the genotype of which is inferred by its position relative to dots from individuals of known genotypes

Fig 4. SNPs revealed geographic groupings similar to the ones indicated by the other marker types

Allele frequencies for an example SNP locus: ADFG|Ots_C3N3

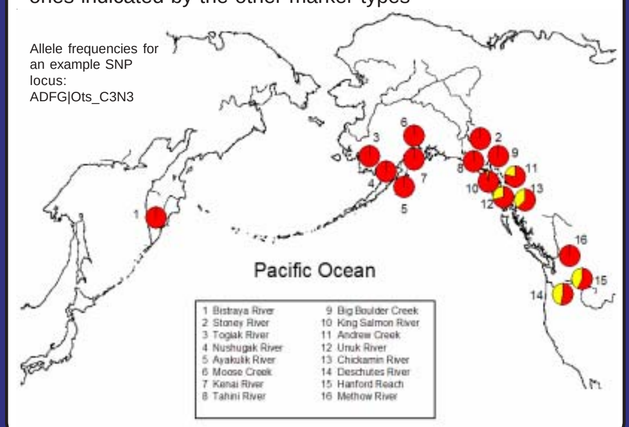


Fig 5. Multidimensional scaling plots of genetic distances[13] between populations

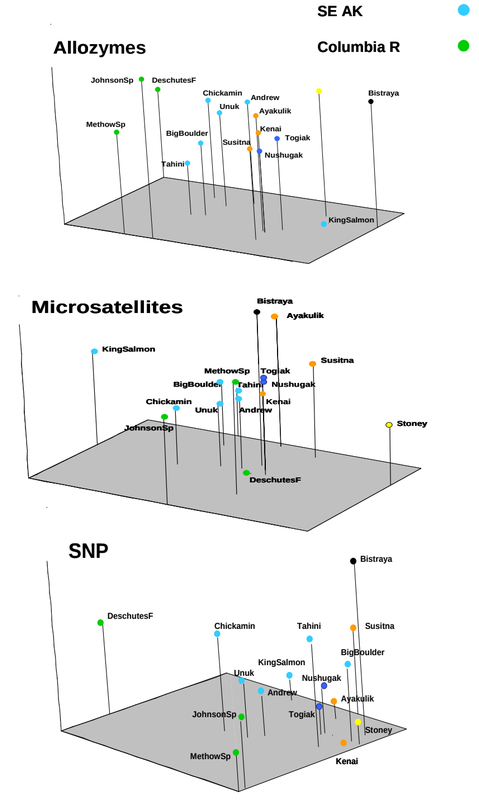


Fig 6. F_{ST} values observed for each locus

Average F_{ST} estimates[14] were similar for allozyme and microsatellite markers, but were higher for SNPs. Maximum F_{ST} estimates for microsatellite loci will be reduced as the number of alleles increase. Higher F_{ST} and genetic distance estimates for SNPs, however, are likely due to the fact that we intentionally developed assays for highly informative loci (Table 1).

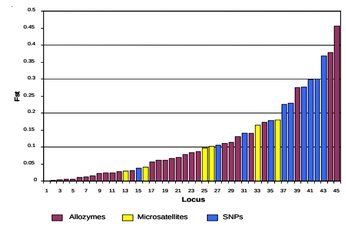
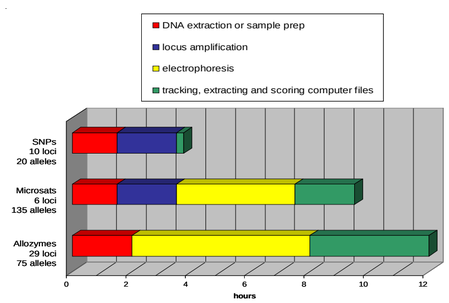


Fig 7. Comparison of the times required to analyze a single collection using each of the marker types

The SNP loci developed for this study required less time to run than either of the other two marker types. This is because the SNP assays require no electrophoresis (Fig 2), and much less interpretation of raw data (Fig 3) than do microsatellites or allozymes.



Results and Conclusions

Conversion of genetic markers from PCR-RFLP, DNA sequencing and DGGE to high-throughput SNP genotyping assays was successful. These SNP assays are cheaper and faster to run than genetic markers presently being used for stock structure analyses (Fig 7). Further, since the required number of individuals in baseline samples increases with increasing numbers of alleles per locus, SNP baselines should be relatively smaller and thus cheaper to produce.

A second advantage of SNP markers over microsatellites and allozymes is that SNP data are discrete (nucleotide bases) rather than continuous (relative mobilities). This allows immediate standardization of SNP allele definitions among laboratories, an exercise that has proven difficult and expensive (in some cases prohibitively so) in multi-agency studies of microsatellite and allozyme variation.

Mean genetic distance and F_{ST} estimates were higher for SNPs than for the other two loci (Fig 6), reflecting the fact that SNP loci were chosen based on *a priori* knowledge of their information content.

Patterns of variation revealed by the three marker types were largely concordant (Fig 5). The correlation was higher between allozymes and SNPs (r^2=0.57) than between either allozymes and microsatellites (r^2=0.35) or between SNPs and microsatellites (r^2=0.23). All correlations were highly significant (P<0.01)

Given that the time and monetary requirements for running SNP genotyping assays are low relative to other classes of genetic markers and that a wealth of previously described polymorphisms may be accessed using these new technologies, it is likely that SNPs will become increasingly important tools for population genetic studies in fisheries.

Acknowledgments: Funding for this study was received from the Southeast Alaska Sustainable Fisheries Fund, and the Chinook Technical Committee of the Pacific Salmon Commission. Technical Assistance was provided by Nick DeCovich, Zac Grauvogel, Judy Berger and Penny Crane. Tony Gharrett provided helpful insight regarding the analysis of these data. We are grateful to Shawn Narum for providing us with tissue samples.

Reference List

1. Manlel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209-20.
2. Schwenke, P., Rhyddereh, J., Ford, M., Marshall, A. and Park L. (2003) Identification of Upper Columbia Chinook spring-run (*Oncorhynchus tshawytscha*) for forensic and salmon predation research applications. *manuscript in prep.*
3. Xiong, F., Chin, R.A. and Hew, C.L. (1992) A gene encoding chinook salmon (*Oncorhynchus tschawytscha*) prolactin: gene structure and potential cis-acting regulatory elements. *Mol Mar Biol Biotechnol* **1**, 155-64.
4. Park, L.K., Moran, P. and Dightman, D.A. (1996) A chinook salmon PCR-RFLP marker in the p53 locus. *Anim Genet* **27**, 121-128.
5. Park, L.K., Moran, P. and Dightman, D.A. (1995) A polymorphism in intron D of the chinook salmon growth hormone 2 gene. *Anim Genet* **26**, 285
6. Miller, K.M. and Withler, R.E. (1998) The salmonid class I MHC: limited diversity in a primitive teleost. *Immunol Rev* **166**, 279-293.
7. Miller, K.M., Withler, R.E. and Beacham, T.D. (1997) Molecular evolution at Mhc genes in two populations of chinook salmon *Oncorhynchus tshawytscha*. *Mol Ecol* **6**, 937-54.
8. Blankenship, S.M., May, B. and Hedgecock, D. (2002) Evolution of a Perfect Simple Sequence Repeat Locus in the Context of Its Flanking Sequence. *Molec Biol Evol* **19**, 1943-1951.
9. Ford, M.J., Thornton, P.J. and Park, L.K. (1999) Natural selection promotes divergence of transferrin among salmonid species. *Mol Ecol* **8**, 1055-61.
10. Ford, M.J. (1998) Testing models of migration and isolation among populations of chinook salmon (*Oncorhynchus tschawytscha*). *Evolution* **52**, 539-557.
11. Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Nat Acad Sci* **88**, 7276-7280.
12. Lee LG, Connell CR, Bloch B (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nuc Acid Res* **21**, 3761-3766.
13. Cavalli-Sforza, L.L. and Edwards, A.W. (1967) Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* **19**, 233-257.
14. Nei, M. (1987) Molecular Evolutionary Genetics. New York: Columbia University Press. p164-165.