Detection of Variation in Highly Polymorphic Mhc Genes by Denaturing Gradient Gel Electrophoresis Using the DCode™ System

Kristi M. Miller, Tobi J. Ming and Ruth E. Withler, Pacific Biological Station, Department of Fisheries and Oceans, Nanaimo, B.C. V9R 5K6, Canada

Introduction

We describe a rapid and sensitive method for the detection of nucleotide sequence variation which can be used for large-scale screening of population markers. Denaturing gradient gel electrophoresis (DGGE) detects sequence variants of amplified fragments by the differences in their melting behavior (Fischer and Lerman 1983). By adding a GC-clamp to one of the primers, nearly all single base substitutions can be detected (Sheffield *et al.* 1989). Although DGGE has primarily been utilized for the detection of single base mutations in disease studies, it offers great potential for use in population analysis of genetic markers with higher levels of sequence variation. Herein we describe our DGGE analysis of a Mhc class I locus in salmon.

The methodology described was developed to identify the number and distribution of Mhc class I *A1* (exon 2) alleles among chinook salmon (*Oncorhynchus tshawytscha*) populations (Miller *et al.* 1997, Miller and Withler 1997, 1998). DGGE detects 25 of the 28 identified *A1* sequences, with a nucleotide sequence variation between them of 1–16 nt plus a 2 codon indel. By creating a unique network of control alleles, 22 of these alleles can be rapidly and accurately resolved by a single gel run condition. This technique has been used in surveys of over 20,000 salmon, representing two Mhc markers (class I *A1* and A2) in two salmon species (chinook and coho salmon). A single person in our laboratory now analyses 160 salmon per day with DGGE using two DCode systems.

Methods and Results

Single locus Mhc class I *A1* primers were used to amplify fragments of 301–307 bp (including primers) (Miller *et al*. 1996). Because a large proportion of the sequence variation and the lower melting domain reside on the 3' end of the fragment, a GC-clamp was added to the 5' primer.

DGGE was performed on the DCode system. We determined initial DGGE conditions by electrophoresing three the amplified chinook salmon *A1* chinook alleles (5, 9, and 10), differing from one another by 7–12 nt and a 2 codon indel on a perpendicular gradient gel (Figure 1). Dissociation of *A1* alleles occurred over a 36–54% gradient. The presence of three distinctive but overlapping curves indicated that all three *A1* alleles melted at different concentrations of denaturant and could be differentiated on parallel DGGE gels, but might shift relative positions if run into different portions of the denaturant gradient. Further analysis of these alleles by WinMelt™ software again showed three distinctive curves, with two melting domains still apparent (despite the GC-clamp), and elucidated the crossover point between alleles 9 and 10 at nt position 207 (Figure 2). This crossover point directly coincided with the location of the 2 codon insert, present in allele 10, but absent in allele 9. Thus, it appeared that the 2 codon insert and the presence of two melting domains may cause shifting of some alleles on parallel gels.

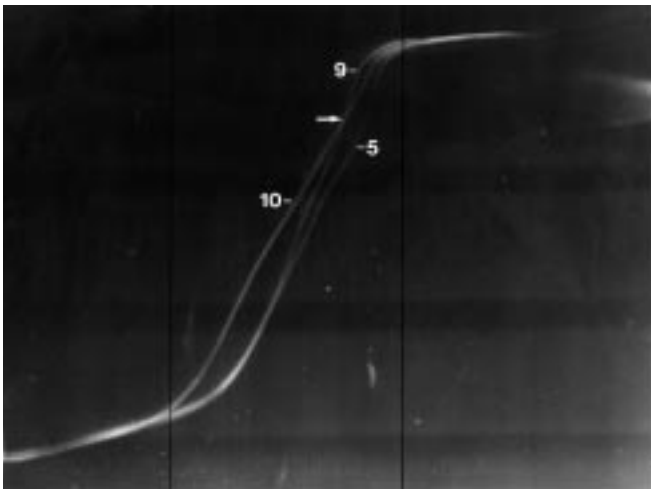

Fig. 1. Perpendicular DGGE of A1 *alleles 5, 9, and 10 run on a chemical gradient of 15–90%. Melting curves for* A1 *alleles ranged from 36–54%. Curves for alleles 9 and 10 cross at 51.5% denaturant.*

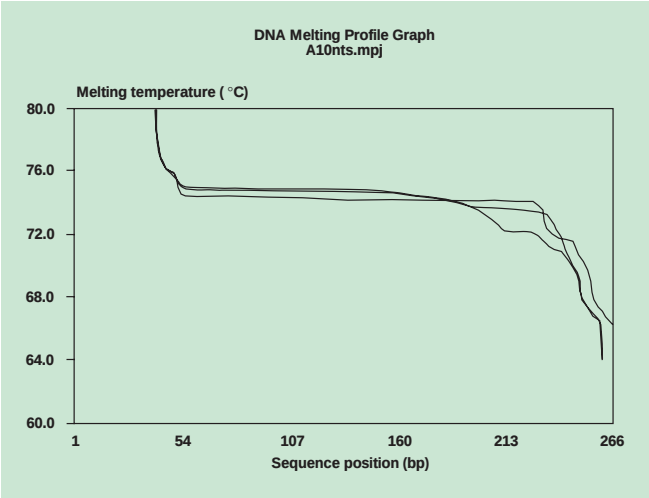

Fig. 2. Predicted melting curves for alleles 5, 9, and 10 calculated using WinMelt™ software. Alleles differ by 7 to 12 nt and a 2 codon indel.

Time-series analyses on 40–65% (7.5% acrylamide) parallel gradient gels was used to determine the run-time that separated most of the 28 *A1* alleles (homoduplexes only) and to assess changes in the relative positions of alleles at different denaturant concentrations. Virtually all of the alleles were differentiated by DGGE, with the exception of two sets of alleles with a single C-G transversion within 10 bp of the clamped primer and one set of alleles with a single A-T transversion at nt 237. The remaining 5 sets of alleles with single base differences at positions ranging from nt 85 to 282 (24 to 243 not including 5'-clamped primer) were differentiated, as well as all alleles with a greater number of differences. Thus, the range of denaturant gradient used and the addition of the GC-clamp facilitated sampling of sequence variation over virtually the entire *A1* fragment.

As expected from the presence of two melting domains within the *A1* sequence, time series analysis revealed a number of alleles which shifted relative positions on the gel when electrophoresed into different ranges of denaturant concentration. Moreover, some alleles that migrated differentially through the gel at low concentrations of denaturant became indistinguishable when electrophoresed in high denaturant concentrations (Figure 3). As a result, there was no single run-time that consistently separated all of the alleles. In the time series shown in Figure 3, the best separation of alleles was obtained at a run-time of 5.5 hr, which placed the DNA fragments between 48% and 53% denaturant, a range in which dissociation of the *A1* low melting domain was completed and the high melting domain was initiated. At this run-time, 18 alleles were distinctive, and 5 sets of alleles migrated to similar positions and were consequently binned into 5 allelic groups.

For the homoduplex *A1* allele fragments, increasing migration distance through the denaturing gradient was strongly correlated with increasing overall GC content. However, the presence/absence of the 2-codon insert, and the GC content of the last 70 nt (distal to the GC-clamp) also had apparent influence on allelic migration. Prior to any allele dissociation, alleles separated according to size, with the alleles containing the two-codon insert migrating slower than those without the insert. However, once alleles began to dissociate, the insert was apparently less influential on allele migration than was the GC content, and some of the alleles that contained the insert and a high overall GC content migrated past alleles that did not contain the insert. In addition, because much of the gradient sampled variation in the low melt domain, GC content distal to the GC-clamp was more influential on the overall migration distances of the homoduplexes than GC content close to the clamp.

Alleles that shifted (exchanged relative positions in a ladder of migration distance) often did not share the presence/absence of the 2-codon insert and/or differed in their relative GC content in the low and the high melt domains. The most notoriously shifting allele (9; fourth band from the bottom of lane A in Figure 3) shifts position with four other alleles so readily it can actually shift across a single gel if the denaturing gradient is not completely even. Careful monitoring of the standard allele sets is necessary to accurately score this allele, and it is often used as a marker for gel to gel run uniformity.

To facilitate rapid and accurate scoring of alleles on population gels, including those that shift relative positions during in different run times, subsets of the "known" alleles were combined into three standard sets to use as controls in the following way: The relative positions of alleles run under optimal conditions were determined from the time series analysis, and the fastest (latest melting) to slowest (earliest melting) alleles were alternately placed into each one of the 3 standard sets. Alleles that were known to shift relative positions at different denaturing concentrations (as identified in the time-series analysis) were placed in different standard sets. The A, B, and C lanes in the time-series are the actual standard sets used in the population gels.

On the population gels, allele standard sets were loaded in three adjacent lanes in the center and on each side of the gel, with 20 individuals from one or two populations run in the remaining lanes. Parallel DGGE gels were photographed with a CCD camera utilizing a 1024 x 1024 cm pixel array format and analyzed with standard gel documentation software. Standard allele sets were used to confirm that the electrophoretic conditions were sufficient to resolve all of the standard alleles, to monitor the relative positions of shifting alleles on the gel, and to form a network on which to score individual alleles. Alleles within the standard lanes were assigned "DGGE allele" numbers ranging from 50 to 400, approximately corresponding with their (logarithmic) positions on the gel. New alleles identified in the population survey with denaturing points that did not coincide with any of the known DGGE alleles were added to standard lanes as they were found.

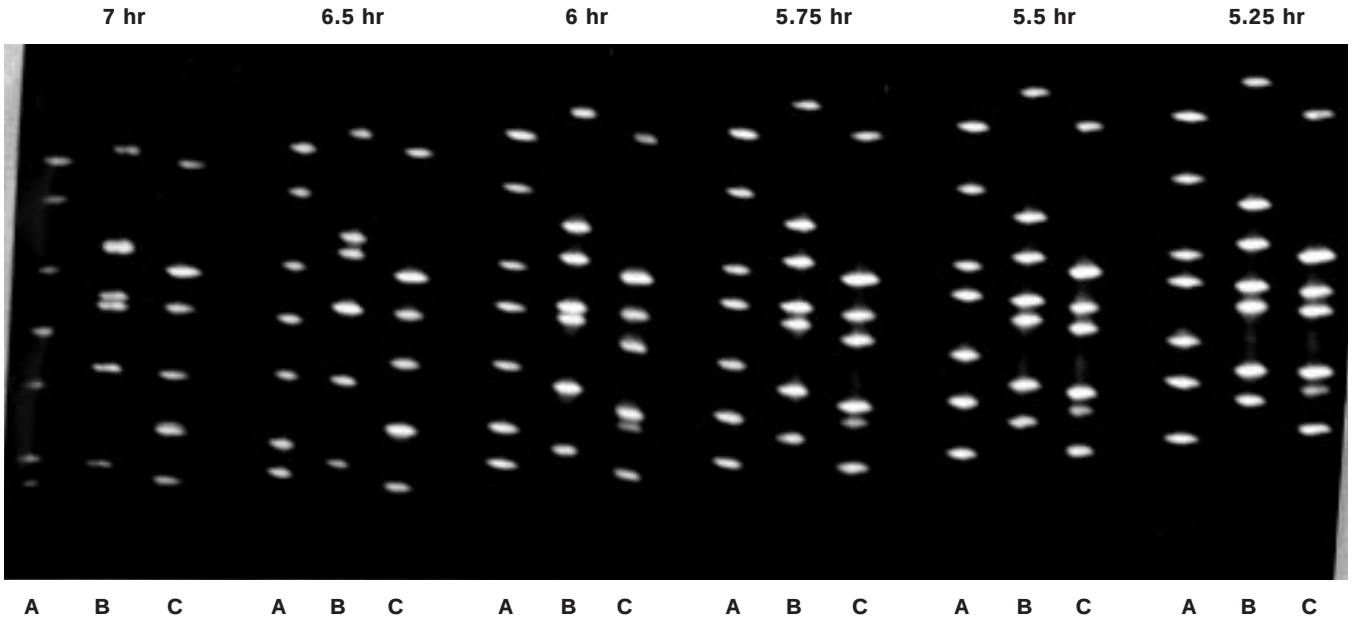

Fig. 3. Time-series analysis of chinook A1 *alleles run on a 45–60% denaturant gradient for 5 to 7 hours at 120 V. Lanes A, B, and C contain subsets of 21 of the 22 alleles distinguished by DGGE (allele 24 is not shown). Only a single allele from comigrating pairs of alleles is shown. The maximum resolution of alleles occurred at 5.5 hours, which was the run-time used for further population analyses.*

Troubleshooting

Several factors affect our ability to produce DGGE gels of sufficient quality to score reliably over time. The purity and age of the acrylamide and acrylamide denaturant solutions, the maximum voltage, and the overall run time all strongly affect the resolution and migration of fragments. Optimally, for our analysis DGGE was the most reliable when run slowly (60–80 V) overnight (<17 hours). However, to maximize the throughput of the DCode system, we conduct the *A1* analysis on daytime runs (at 120 V max) and analyze other loci on overnight runs. We use a single brand of acrylamide for analysis of each gene locus, and run a time series on new batches of acrylamide for slight run-time adjustments. Acrylamide over 6 months old and acrylamide denaturant solutions over 7 days old can be problematic and are avoided. In general, run-times for *A1* are decreased by 5 minutes per day after the acrylamide solution is made. Furthermore, in order to accurately reproduce the run-time conditions from the time-series to the population gels, population gels were pre-run for 1.5 hr prior to loading. A more refined time series (from 5 to 6 hr) would decrease the necessary pre-run time.

Discussion

DGGE has been utilized for population analysis in only a limited number of studies (Brown *et al.* 1997; Norman *et al*. 1994), most merely to identify haplotypes among a few individuals for sequencing. Our study offers the first example of the potential use of DGGE for large scale population surveys of highly polymorphic DNA loci. In our laboratory, each technician runs 160 fish per day on DGGE (8 gels) using two DCode systems (80 fish each for daytime (*A1* locus) and nighttime (*A2* locus) runs). Ambiguous scoring of shifting alleles necessitates the re-running of less than 2% of all fish analyzed. This rapid rate of analysis of nucleotide sequence variation makes it possible to quickly accumulate large databases, and to use DGGE for applied, real-time fisheries issues such as stock identification in mixed stock fisheries, escapement enumeration within river systems, and the identification of escaped farmed fish, to name a few.

The DGGE methodology described herein can be used on any region of DNA with nucleotide sequence variation. Best results are obtained with fragments of 300–500 bp, but larger fragments can be analyzed by first digesting them with restriction endonucleases (Sheffield *et al.* 1989). In DGGE, the detection of length variation is often superseded by nucleotide variation, but DGGE could be an invaluable tool for detecting homoplasy in microsatellites with similar migration on acrylamide gels. The use of DGGE virtually eliminates the need to sequence large numbers of clones to examine sequence variability. Direct sequencing of fragments cut out of gradient gels can further expedite the process. Finally, by careful use of standard networks of known alleles, DGGE is a highly sensitive and expeditious technique for detecting nucleotide sequence variation in large-scale population surveys.

References

1 *Brown, J. M., Leebens-Mack, J. H., Thompson, J. N., Pellmyr, O. and Harrison, R. G., Phylogeography and host association in a pollinating seed parasite Greya* politella (Lepidoptera: Prodoxidae), Mol. Ecol., ***6**, 215–224 (1997).*

2 *Fischer, S. G. and Lerman, L. S., DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: Correspondence with melting theory,* Proc. Natl. Acad. Sci. USA., ***80**, 1579–1583 (1983).*

3 *Miller, K. M., Withler, R. E. and Beacham, T. D., Molecular evolution at Mhc genes in two populations of chinook salmon* Oncorhynchus tshawytscha, Mol. Ecol., ***6**, 937–954 (1997).*

4 *Miller, K. M. and Withler, R. E., Mhc diversity in Pacific salmon: Population structure and trans-species allelism,* Hereditas, ***127**, 83–95 (1997).*

5 *Miller, K. M. and Withler, R. E.,* The salmonid class I MHC: limited diversity in a primitive teleost, Immunological Rev., ***166**, In Press, (1998).*

6 *Norman, J. A., Moritz, C. and Limpus, C. J., Mitochondrial DNA control region polymorphisms: genetic markers for ecological studies of marine turtles,* Mol. Ecol., ***3**, 363–373 (1994).*

7 *Sheffield, V. C., Cox, D. R., Lerman, L. S. and Myers, R. M., Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes,* Proc. Natl. Acad. Sci. USA, ***86**, 232–236 (1989).*

Bio-Rad Laboratories

Life Science Group

Website *www.bio-rad.com* ***U.S.*** *(800) 4BIORAD* ***Australia*** *02 9914 2800* ***Austria*** *(01)-877 89 01* ***Belgium*** *09-385 55 11* ***Canada*** *(905) 712-2771*
China *86-10-62051850/51* ***Denmark*** *45 39 17 99 47* ***Finland*** *358 (0)9 804 2200* ***France*** *01 43 90 46 90* ***Germany*** *089 318 84-0*
Hong Kong *852-2789-3300* ***India*** *(91-11) 461-0103* ***Israel*** *03 951 4127* ***Italy*** *39-02-216091* ***Japan*** *03-5811-6270* ***Korea*** *82-2-3473-4460*
Latin America *305-894-5950* ***Mexico*** *514-2210* ***The Netherlands*** *0318-540666* ***New Zealand*** *64-9-4152280* ***Russia*** *7-095-4585822*
Singapore *65-2729877* ***Spain*** *34-91-661-7085* ***Sweden*** *46 (0)8 627 50 00* ***Switzerland*** *01-809 55 55* ***United Kingdom*** *0800-181134*

Bulletin 2416 US/EG Rev A *98-541 1298 Sig 100798*