

**Final Report: Development of DNA Fingerprinting Markers for the American Dipper, *Cinclus mexicanus***

Cynthia Anderson, Shane Sarver  
Department of Biology  
Black Hills State University  
1200 University St. Unit 9067  
Spearfish, SD 57709  
605-642-6043  
[CynthiaAnderson@bhsu.edu](mailto:CynthiaAnderson@bhsu.edu); [shanesarver@bhsu.edu](mailto:shanesarver@bhsu.edu)

**Objective:** The primary objective of this project was to construct a small fragment genomic library for the American dipper using the PCR-based Identification of Microsatellite Arrays (PIMA) method as described by Lunt *et al.*, 1999. This method utilizes random primers to generate amplification products that are reported to contain a higher frequency of microsatellite sequences in many of the animals in which this method has been utilized. Once this library was constructed, DNA sequencing was to be used to identify and characterize the nature of the microsatellite region, and primers would be designed that would specifically amplify that locus. This would allow us to begin to identify polymorphic loci that could be used as microsatellite markers in future research involving the evolutionary population genetics of the American dipper in the Black Hills.

**Materials and Methods:** American dipper genomic DNA was isolated from a museum sample provided to us by the Museum of Vertebrate Zoology at the University of California, Berkeley. Genomic DNA was isolated using a DNEasy extraction kit (Qiagen, Valencia, CA) Using a method known as PCR-based Identification of Microsatellite Arrays (PIMA) described by Lunt *et al.*, 1999, we attempted to identify genomic regions containing microsatellite sequences. To construct the small fragment genomic library we used a set of 20 random primers to generate randomly amplified polymorphic DNA fragments. Those PCR reactions that yielded the most amplification products were then subcloned into plasmid pGemT (Promega, Madison, WI) and transformed into *E. coli* strain JM109 (Promega). Transformed bacteria were picked, grown in LB broth, and transferred into 96 well culture plates containing a glycerol solution for storage of the library at  $-70\text{ C}$ . Prior to freezing, a 96-pin replicator was used to inoculate agar plates with the bacteria for daily maintenance and for use in screening the library.

To screen the library, colony PCR was done using M13 forward and M13 reverse primers that bind to the plasmid on each side of the multiple cloning site, plus a primer that contains a repetitive sequence. The repetitive sequences we used were  $(AC)_{12}$ , and  $(AG)_{12}$  which are reported to be the most frequent dinucleotide repeats in many, though not all, vertebrates (Ender *et al.*, 1996; Zane *et al.*, 2002). PCR products were then run on 1.5% agarose gels. Those reactions that yielded both a full length amplicon generated from the forward and reverse primers, as well as a shorter fragment, generated by the repeat primer and either the forward or the reverse primer, were repeated, and the PCR product was used for sequencing to identify the region containing homology to the repetitive primer. Sequencing reactions were done using a BigDye Terminator Ready Reaction Kit v3.1 (Applied Biosystems) following the manufacturers protocol, and analyzed on a 3100 Avant Genetic Analyzer (Applied Biosystems).

We found it necessary to begin work to construct two more small fragment genomic libraries, one enriched for microsatellite containing sequences using a method described by Carleton *et al*, 2002, and one that was not enriched. Construction of both libraries is in progress, and the methods being used are as follows. Genomic DNA was extracted from dipper muscle tissue that had been given to us by D. Backlund of SD Game, Fish and Parks. In both cases, genomic DNA was digested overnight with *Sau3AI* then run on a 0.8% low melting point agarose gel to separate the fragments. The region between 700 and 1800 base was excised from the gel. For the non-enriched library, these fragments will be ligated directly into pBlueScriptKS vector (Stratagene, La Jolla, CA) that has been linearized with *BamHI* to yield compatible ends with the *Sau3A* digested genomic DNA. These recombinant plasmids will be transformed into *E. coli* strain JM109 (Promega). Colonies will be picked, grown overnight in liquid culture, aliquotted into cryo-tubes containing glycerol, and stored at  $-70^{\circ}\text{C}$ . Individual cultures from the library will be maintained on LB media containing Ampicillin during the process of screening for microsatellite sequences. This library will be screened by colony blotting using a set of biotin labeled repetitive probes using a DIG chemiluminescent detection kit (Roche, Indianapolis, IN). Positive clones will be sequenced. For the enriched library, the DNA was purified from the low melting point gel using a Wizard Gel Extraction Kit (Promega). These fragments were ligated to linkers that would provide a known sequence for PCR amplification, and contained a *BamHI* site. To enrich for repetitive sequences, single-stranded oligonucleotide probes were custom made for us by Integrated DNA Technologies that each contained a repetitive sequence and was biotinylated on the 5' end (table 2). These oligonucleotides will then be conjugated to magnetic beads coated with streptavidin. The isolated, linker-ligated fragments will then be incubated with the streptavidin coated beads, allowed to hybridize to the probes and the beads which will be washed to remove unhybridized fragments. The fragments that remained hybridized will then separated from the probes, and ligated into an appropriate cloning vector, transformed and maintained as above. These fragments will be screened by PCR to identify inserts greater than 300 base pairs in length. These will then be sequenced. In both cases, the sequences will be analyzed to identify those containing microsatellite sequences and primers will be designed and optimized for these markers.

### **Results/Discussion:**

Of the 20 RAPD primers that were used, only 3 yielded multiple PCR products, A3 5'-AGTCAGCCAC-3', A9 5'-GGGTAACGCC-3', and A13 5'-CAGCACCCAC-3'. These PCR products were subcloned into pGemT and transformed as stated above. The PIMA library consisted of 288 colonies. When PCR was done using the M13 forward, M13 reverse and repeat primers, more than half of the PCR reactions yielded more than two amplified fragments, but only 54 of those yielded a smaller fragment unique in size. These 54 clones were sequenced, and while many contained short dinucleotide repeats (3 to 4 pairs in length), none contained repetitive regions long enough to be useful as microsatellite markers. One confounding factor in the construction of this library was the quality and quantity of the DNA extracted from the museum tissue. Yields were less than 0.2 micrograms per 0.2 grams of tissue, and the DNA that was extracted was somewhat degraded. To make the most of the library we had to gel purify the non-degraded DNA. This lack of good quality genomic DNA most certainly reduced the number of markers we were able to obtain. Additionally, it is possible that long AC and AG dinucleotide repeats are not as common in the American dipper as in other vertebrates. It is also probable that by limiting our subset of fragments to those generated by only 20 RAPD primers,

and then further limiting our subcloning to only those that yielded multiple amplicons, as the Lunt method suggested, we excluded too much of the genome to efficiently identify microsatellites. While this method was touted as being able to yield several microsatellite markers by utilizing such a small subset of RAPD primers, we found that the number and diversity of amplicon sequences was just not adequate to obtain the markers we were looking for. One other difficulty with microsatellite development in birds is the fact that birds have reduced genome sizes, which means there is a lesser amount of non-coding DNA in the genome. This makes development of microsatellite markers more difficult.

Despite the negative results obtained from the PIMA method of marker development, we did feel this project was of significant importance to the Black Hills ecosystem, especially as it relates to water quality issues, that we pursued additional funding for this project through a collaborative grant with the Environmental Protection Agency which will allow some additional funding to continue to work on this project. Since there is some urgency to obtaining genetic markers for the Dipper, and since we had already spent several months only to get nothing valuable, we decided to put off any further screening utilizing the Lunt method in favor of older, though more tedious, methods that do not limit the fragments contained in a library to those of perhaps questionable utility. We were also able to obtain fresher dipper tissue from SD G, F and P. From this tissue we were able to isolate ample DNA of high quality to work with. The new libraries have been successfully constructed and we are in the process of screening those libraries.

## References

- Carleton KL, Streebman JT, Lee BY, Garnhart N, Kidd M, Kocher TD. 2002. Rapid isolation of CA microsatellites from the tilapia genome. *Animal Genetics* 33(2):140-144
- Ender A, Schwenk K, Stadler T, Streit B, Schierwater B. 1996. RAPD identification of microsatellites in *Daphnia*. *Molecular Ecology* 5:437-441
- Lunt DH, Hutchinson WF, and Carvalho GR. 1999. An efficient method for PCR-based isolation of microsatellite arrays (PIMA). *Molecular Ecology* 8:891-894
- Zane L, Bargelloni L, and Patarnello T. 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* 11:1-16