Selection of optimal spawning pairs to maintain genetic variation among captive populations of Acipenseridae based on the polymorphism of microsatellite loci

Dariusz Kaczmarczyk

Received – 16 December 2015/Accepted – 05 May 2016. Published online: 30 June 2016; ©Inland Fisheries Institute in Olsztyn, Poland

Citation: Kaczmarczyk D. 2016 – Selection of optimal spawning pairs to maintain genetic variation among captive populations of Acipenseridae based on the polymorphism of microsatellite loci – Arch. Pol. Fish. 24: 77-84.

Abstract. The American paddlefish, Polyodon spathula (Walbaum), is an endangered acipenserid fish. Its wild populations are supplemented with stocking material that is obtained by conducting artificial spawning in aquaculture conditions. When fish are bred in captivity, it is important to select breeding pairs that will produce the most genetically diverse progeny, since this permits maintaining the fitness of wild populations. Breeding pairs of land animals are selected successfully based on the polymorphism of their microsatellite loci. This theoretical paper asks how to adapt this technique to fish so that American paddlefish spawners can be paired with the aim of producing restocking material in aquaculture that maintains genetic variation. To test our calculating techniques, we used actual data on the polymorphism of the microsatellites from paddlefish broodstock at the Pogorze fish farm (Poland). The data enabled us to do calculations that showed which spawner pairs would create the most genetically diverse offspring and how to assemble sets of spawning pairs that would be best for maintaining genetic variation. The method presented in this paper can be used for breeding fish in aquaculture to help conserve species. It could also be used in a computer program which would automate calculations and present them in easy-to-read tables and graphs.

Keywords: genetic variation, microsatellite loci, assemblage of spawning pairs, conservation of populations, paddlefish

Introduction

The American paddlefish, Polyodon spathula (Walbaum), is an acipenserid fish inhabiting the large rivers of central North America (Nelson 1984). In the twentieth century, its populations were severely reduced, sometimes to extirpation, as a result of over-harvesting, pollution, and damming rivers (Birstein 1993). The paddlefish is an endangered species in need of conservation; it is listed in the “red book” as an endangered or threatened species in many states of the USA (Birstein 1993). Conservation strategies are often based on artificial spawning and fish stocking performed in aquaculture. The aquaculture of this species also generates valuable products for the US and foreign markets (Billard and Lecointre 2001, Mims 2001, Vedrasco et al. 2001, FAO 2005, Pikitch et al. 2005), including Europe and Poland. The creation of hatchery stocks that will be provide progeny for the supplementation of natural populations is essential for conservation work. Thus, it is essential to select and use appropriate...
spawning pairs to ensure appropriate genetic variation in subsequent generations of spawners.

Broodstocks held in hatcheries are usually limited in size, which means they are vulnerable to progressive domestication (Koljonen et al. 1999, 2002, Verspoor et al. 2005). In hatcheries, natural selection favors individuals that are better adapted to life in captivity than in the natural environment (Kim et al. 2004). In addition to the founder effect and the crossbreeding of closely related individuals, allelic diversity, and heterozygosity can be reduced within a conserved population, which, consequently, reduce its viability (Bryant et al. 1986, Shikano and Taniguchi 2002). The consequences of decreasing genetic variation in populations of paddlefish or any other long-lived, late maturing animal species would likely be difficult to reverse (Olech 2003, Kaczmarczyk et al. 2007).

To effectively manage paddlefish and other populations that are dependent on humans, up-to-date knowledge of the structure and mating patterns of the broodstock is necessary. This permits maintaining genetic variation in populations, and this can be achieved by first rearing individuals that differ most genetically, and then arranging them in appropriate spawning pairs. However, if spawning pairs are assembled based on phenotype, closely related individuals can be crossed resulting in decreased genetic variation (Bryant et al. 1986) and viability in the progeny cohort, which render conservation programs unsuccessful (Hallerman 2003).

One technique for selecting optimal pairs of spawners is based on the polymorphism of their microsatellite loci (Bruford and Wayne 1993, McConnel et al. 1994, Hansen et al. 2000, Henderson et al. 2004). This method can be used as a tool for managing genetic variation in fish species and other vertebrates that are dependent on humans. Variation of microsatellite markers permits matching spawning pairs according to the genetic differences identified among spawners and reducing the likelihood of inbreeding events. Here, we used three measures of genetic variation that are known to be important for species conservation (Stow and Briscoe 2005): heterozygosity; the share of weak heterozygotes; the number of different alleles potentially inherited by the progeny of each spawning pair (Kaczmarczyk and Fopp-Bayat 2013). Weak heterozygotes are individuals that are tetraploid at a given locus, and which, at that locus, have three alleles that are the same and only one allele that is different, e.g., AAAB. Crossing weak heterozygotes should be avoided because it can lead to a marked decrease in the heterozygosity of the progeny cohort.

This study proposes a method to conserve genetic variation by selecting appropriate pairs of paddlefish for spawning. The selection should be based on: heterozygosity, the share of weak heterozygotes, and the allelic diversity of progeny that are known to be important for population viability. Based on the values of these indicators of genetic variation, we attempted to construct an optimal set of three spawning pairs that would best maintain the genetic variation of the broodstock while also producing highly heterozygous progeny which would be best for the conservation of the human-dependent species such the paddlefish.

Materials and methods

Fish samples

Before sampling, all individuals were tagged by using the TROVAN FDX-B system. The transponders were preprogramed with 15 digit ID codes, and they were implanted in the dorsal muscles near the head of the fish. The tag numbers can be read with portable readers designed for this system. This tagging method enabled identifying the fish at a distance of up to 0.5 m without having to remove it from the tank. Pelvic or pectoral fin clips (approximately 100 mm²) were collected from all individuals (24 fish aged about 11 years) from the broodstock at Pogorze fish farm located in southern Poland near the Czech border. Next, the sex of each fish was determined. The fish investigated in this study were a part of samples taken from 100 individuals of the Polish stock, from Wasosze, Pogorze as well as from a stock from Gorny Tykich in Ukraine (Kaczmarczyk et al. 2013).
The fin clips were placed in Eppendorf tubes and kept in 97% ethanol at a temperature of 4°C until examination.

**DNA extraction**

Genomic DNA was extracted and purified from fin tissues using a DNA Genomic Wizard Kit (Promega) following the procedure described by Fopp-Bayat and Luczynski (2006). DNA samples were stored at a temperature of –20°C. The integrity of the DNA samples was visually inspected after their electrophoresis in 1.5% agarose gel stained with ethidium bromide. All agarose gels were photographed using a gel imaging system and the pictures were computer recorded. Samples of the DNA yields were quantified by spectrophotometric analysis. Samples containing more than 30 pg μl⁻¹ of double stranded DNA were qualified for the PCR stage.

**PCR amplification**

The microsatellites examined were Psp12, Psp18, Psp20, Psp21, Psp26, Psp28, Psp29, and Psp32 (May et al. 1997, Heist et al. 2002). All forward primers were 5’-labeled using phosphoramidite dyes (D2-PA, D3-PA and D4-PA). In the case of base dilutions, the number of freezing–thawing cycles was reduced as much as possible. Details of the PCR reaction are described in Kaczmarczyk et al. (2007).

**Genotyping**

The lengths of amplified DNA fragments were determined using a Beckman Coulter CEQ 8000 sequencer against 60-400 bp size standard. Amplified fragments were arranged in three sets. Set No. 1 consisted of amplicons Psp12, Psp18, and Psp26; set No. 2 consisted of amplicons Psp20, Psp21, and Psp28; and set No. 3 of Psp29, Psp32, and Ls57. Within these sets, each individual microsatellite was amplified using primers with different phosphoramidite labels attached, thus enabling their separation in a multiplex mode. Fragment size and allele determination was performed using CEQ 8000 Fragment Analysis software following the manufacturer’s recommendations.

**Mathematical analysis**

Heterozygosity, the percentage of weak heterozygotes, and the number of alleles in parental group, and that expected in the progeny of each spawning pair were calculated. We calculated the heterozygosity for each locus in the parental group. These values were then used to calculate the average observed heterozygosity ($H_o$) across all di- and tetrasomic loci. With the genotyping results of the parental group, we identified weak heterozygous genotypes, and calculated their share for each given locus, for tetrasomic loci, and for the number of alleles and allelic diversity in the broodstock. The estimation of the indicators of genetic variation in the progeny was based on genetic differences among spawners. The best spawning pair was designated based on the values of the potential heterozygosity of their progeny expressed as the average value of heterozygosity calculated across all of the loci investigated. The calculation also included the percentage of weak heterozygotes (calculated only as the average for tetrasomic loci) and the number of alleles expected in the progeny of each potential spawning pair across all investigated loci according to the method (Kaczmarczyk and Fopp-Bayat 2012, Kaczmarczyk 2015). All values of genetic variation indicators expected in the progeny of each possible spawning pair were computed with Genassemblage 1.0 software (Kaczmarczyk 2015).

The suitability of spawning pairs for breeding was evaluated by simultaneously comparing the heterozygosity, the weak heterozygote percentage expected in their progeny, and the number of different alleles the next generation could potentially inherit. The relative importance of these three indicators to the final evaluation of the spawning pair was expressed in v index values. The weight of each indicator in the calculation value of the v index was: heterozygosity ($i_H = 0.40$), share of weak heterozygotes ($i_{wh} = 0.20$) and number of different
alleles \((i_{ar} = 0.40)\). The value of the \(v\) index was calculated as described in the Genassemblage manual (http://pracownicy.uwm.edu.pl/d.kaczmarczyk/genassemblage/Genassemblage%20manual.pdf) using Algorithm 1:

\[
v = i_h \left[ \frac{H_n}{H_{\text{max}}} \right] + i_{wh} \left[ \frac{wh_{\text{max}} - wh_n}{wh_{\text{max}}} \right] + i_{ar} \left[ \frac{ar_n}{ar_{\text{max}}} \right]
\]

\(i_h + i_{wh} + i_{ar} = 1\)

\(wh_{\text{max}} > 0\)

In this algorithm, \(H_n\), \(wh_n\), and \(ar_n\) are the values of the indicators of genetic variation expected in the progeny of each potential pairing, and \(H_{\text{max}}\), \(wh_{\text{max}}\), and \(ar_{\text{max}}\) are the maximal values of those indicators detected in all analyzed pairings. A higher \(v\) index indicates the best parental combinations that will produce more genetically diverse, heterozygous progeny, and, therefore, it is the optimal choice for the purposes of supplementing the species in the wild.

Based on the \(v\) index values, we have chosen a set of three spawning pairs that best maintain the genetic variation of the broodstock and also produce highly heterozygous progeny. For this set we calculated the average heterozygosity of their progeny and the share of weak heterozygotes (as average values calculated for each pair included in this set). Moreover, we calculated how many different alleles the progeny of those three pairs will inherit. This was calculated by counting the number of different alleles in genotypes of fish descended from those three pairs.

A set of three spawning pairs that have the lowest \(v\) index value was identified, and these should not be bred. The indicators of genetic variation in this set of spawning pairs was calculated in the same way as is described above with the best set of spawning pairs.

### Results

Eighteen out of the 24 fish from the Pogorze broodstock were sexed successfully—seven males and 11 females. The amplification of all the examined (eight) microsatellites extracted from these 18 fish was successful, but only seven microsatellites were genotyped successfully (Table 1).

Seven samples containing locus \(Psp12\) were not genotyped because of numerous stutter bands and

<table>
<thead>
<tr>
<th>Fish No.</th>
<th>Sex</th>
<th>Locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pog 1</td>
<td>F</td>
<td>170 170 250 250 138 152 150 152 208 210 179 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 2</td>
<td>F</td>
<td>170 170 244 246 136 138 146 150 208 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 3</td>
<td>F</td>
<td>170 170 244 246 138 142 146 150 208 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 4</td>
<td>F</td>
<td>170 170 244 246 136 142 150 152 210 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 5</td>
<td>F</td>
<td>170 170 230 254 136 142 146 152 208 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 6</td>
<td>F</td>
<td>170 170 230 244 138 152 150 150 210 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 7</td>
<td>F</td>
<td>170 170 244 246 138 142 150 150 210 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 8</td>
<td>F</td>
<td>170 170 244 246 136 142 150 152 210 210 181 181 195 195 211 215</td>
</tr>
<tr>
<td>Pog 9</td>
<td>M</td>
<td>168 170 244 246 138 152 150 150 210 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 10</td>
<td>F</td>
<td>170 170 246 254 136 142 146 150 208 210 179 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 11</td>
<td>F</td>
<td>170 170 246 254 138 138 150 154 208 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 12</td>
<td>F</td>
<td>170 170 246 254 138 138 150 154 208 210 181 181 195 203 211 215</td>
</tr>
<tr>
<td>Pog 13</td>
<td>F</td>
<td>170 174 246 254 138 138 150 154 210 210 181 181 195 211 215 211</td>
</tr>
<tr>
<td>Pog 14</td>
<td>F</td>
<td>170 170 230 244 152 160 150 154 208 208 179 181 195 211 211 211</td>
</tr>
<tr>
<td>Pog 15</td>
<td>F</td>
<td>170 170 230 244 138 152 150 150 208 208 179 181 195 203 211 215</td>
</tr>
</tbody>
</table>
unspecific products; therefore, this locus was excluded from further calculations. Genotyping of samples included in this study, as well as all samples investigated by Kaczmarczyk et al. (2012), has shown the tetrasomic inheritance of locus $Psp29$. We found fish that were homozygous as well as heterozygotes with two, three, or four different alleles. No individual was noted among the investigated group of 18 fish that was a weak heterozygote at tetrasomic locus $Psp29$.

The heterozygosity observed in the broodstock was 0.60, which was higher than the average heterozygosity calculated across all possible 77 groups of progeny (0.56). The average potential heterozygosity ($h$) within the progeny of each pair varied from 0.27 to 0.68 when calculated across loci (Table 2). Only 27 spawning pairs (35.1%) would create progeny with heterozygosity equal to or higher than that observed in the parental generation.

No weak heterozygous genotype was detected in the parental group, but such genotypes would be possible in the progeny cohorts of some spawning pairs. The average share of weak heterozygous genotypes for all progeny groups would be 0.08, ranging from 0.00 to 0.44 in the potential progeny cohorts of all the possible spawning pairs. Of the spawning pairs studied,
only 35 (45.5%) could create progeny with a share of weak heterozygous genotypes of 0.00 (Table 2).

Twenty five alleles of all microsatellites were detected in the broodstock, for an average of 3.51 alleles/locus. The progeny of each potential spawning pair would inherit between 14 and 20 (average 17) different alleles at the investigated microsatellite loci. The lowest values of allelic diversity (14-15 alleles) would be found in the progeny of twelve spawning pairs, of which the progeny of four spawning pairs would inherit fourteen alleles. The highest allelic diversity (19-20 alleles) would be found in the progeny of seven combinations of spawners, with the progeny of only two spawning pairs inheriting 20 alleles (Table 2).

When the three indicators were combined (heterozygosity, share of weak heterozygotes, and allelic diversity expected in progeny) in the \( v \) index, the best combination was male pog 9 × female pog 6 (\( v \) index 1.00). This pair would create many progeny that are highly heterozygous (0.71), with no weak heterozygotes, and it would transfer the highest (20) number of different alleles to the next spawning pairs (0.81). These combinations would be expected to produce highly heterozygous progeny with diverse alleles and a low share of weak heterozygous genotypes. Therefore, they could be used to replace or supplement the best spawning pair (Table 2).

The optimal set of three spawning pairs would be male pog 9 and female pog 6, male pog 24 and female pog 5, and male pog 13 and female pog 1 (Table 2). The \( v \) index for this set is 0.96, and they would produce a progeny cohort with very high average heterozygosity (0.70), a very low share of weak heterozygous genotypes (0.06), and allelic diversity that is very close to that observed in the broodstock (24 alleles, 3.43 alleles/locus). Certain spawning pairs should be avoided when constructing a set of spawning pairs: male pog 9 and female pog 6, male pog 19 and female pog 6, and male pog 23 and female pog 6 (Table 2). The \( v \) index calculated for this set is 0.42, and they would produce a progeny cohort with an average heterozygosity of 0.42, a high share of weak heterozygotes (0.21), and an allelic diversity of 18 alleles (2.57 alleles/locus). Moreover, all of the spawning pairs in this set include female pog 6, thus increasing the risk of inbreeding in future generations.

**Discussion**

This paper presents a calculation technique that can be a useful tool for maintaining genetic diversity in fish species by designating the best possible spawning pairs. The usefulness of this method in pairing fish for spawning is based on the fact that detection of genetic differences by using sets of microsatellite markers (Nilsisson 1997) is higher than using coding regions, for example those of code proteins (allozymes) (Stahl 1987) or mtDNA (O’Connell and Wright 1997). Therefore, there is the assumption that if fish are similar to each other at microsatellite loci they are even more similar at coding regions. This method is simple and based on the class of markers that are used in studies to evaluate kinship and genetic differences among animals (Heath et al. 2001, Villanueva et al. 2002, Radko 2010). In this method, the precision of detection of genetic differences can be further increased when more loci are used and their polymorphism is higher (O’Connell and Wright 1997). To our knowledge, the limitation of this method is that it is necessary to include the genetic analysis of all fish comprising the broodstock. Moreover, their sex must be clearly identified, and information regarding the spawner pairs that produced given progeny has to be recorded. For the best results, it is a good idea to use spawning sets that include numerous pairs.

By identifying genetic differences among spawners and then assembling them into ideal pairings, highly diverse heterozygous progeny can be produced. Although it may be desirable to use more microsatellite markers to increase the accuracy of the detection of genetic differences among spawners (Ruzzante 1998), here it was possible to identify the genetic differences between the spawners and select the most favorable pairing using only a few loci. The patterns of inheritance of these loci were used to calculate the heterozygosity, the share of weak
heterozygotes, and the allelic diversity of the progeny cohorts of each potential spawning pair. These three measures were combined in the v index, which was used to construct the most favorable sets of spawning pairs. These pairs would enable the transfer to the next generation of allelic diversity that is very close to that in the broodstock, and the production of highly heterozygous progeny. Moreover, there would be a low share of weak heterozygotes in the progeny cohort, which would prevent a decrease in heterozygosity in future generations.

Although it would have been possible to create sets of pairs with a higher average heterozygosity or v index (with pairings like male pog 23 and female pog 6 or male pog 24 and female pog 6), it is better to use different individuals in each pairing to increase genetic variation and reduce the risk of inbreeding.

A similar method has been used with diploid land animals to select breeding pairs based on genetic differences revealed in their microsatellite DNA, for example, in the restoration of the Mexican wolf (Hedrick et al. 1997, Parsons 1998). For fish, a similar approach has also been proposed for managing genetic variation in sturgeons, a fully tetraploid organism (Kaczmarczyk and Fopp-Bayat 2013). To the best of our knowledge, this manuscript is the first attempt to use this method with paddlefish.

The method presented in this paper can also be used for breeding paddlefish and other sturgeons that depend on humans to produce juveniles in aquaculture to produce material for restocking their populations or to prevent their extinction. By maintaining genetic diversity, the fitness of wild populations can be increased. If this method is to be used for conservation purposes, the unique genetic characteristics of supplemented populations should be taken into account. Therefore, the assemblage of spawning pairs using fish taken from different populations should be avoided. To conserve a whole range of genetic diversity, spawning numerous pairs that will produce highly heterozygous and allelic diverse progeny should also be included. The use of numerous pairs that will produce genetic diverse progeny is also important for successful fish culture.

Acknowledgements. The author is grateful to Dr. Klaus Kohlmann and Petra Kersten from the Institute of Freshwater Ecology and Inland Fisheries in Berlin, Germany, for their help in paddlefish genotyping, and to Prof. Dr. Hab. Ryszard Kolman, Institute of Inland Fisheries in Olsztyn, Poland, for providing material for this study.

References

FAO 2005 – Region Review of Aquatic Development in Central and Eastern European Region.– Fisheries and Aquaculture Department, 97 p.
Henderson A.P., Spidle A.P., King T.L. 2004 – Genetic diversity, kinship analysis, and broodstock management of
Shikano T., Taniguchi N. 2002 – Relationships between genetic variation measured by microsatellite DNA markers and a fitness-related trait in the guppy (Poecilia reticulata) – Aquaculture 209: 77-90.