# The application of microsatellite DNA markers for forensic analysis of koi carp (*Cyprinus carpio*)

**J.P. Grobler[a*], A. Kotze[b], H. Swart[b] and E.M. Hallerman[c]**

**We describe the use of microsatellite DNA markers, previously developed for common carp, to investigate a forensic case involving ornamental koi carp. Two South African breeders offered koi strains for sale, which a third breeder claimed were taken illegally from his ponds. Screening of four microsatellite markers provided polymorphic results for koi, demonstrating their applicability for forensic studies. Amplification product sizes were comparable to those published for common carp at three of the four loci. We observed higher allelic diversity among koi carp than had been reported for a panel of inbred common carp. Coefficients of population differentiation showed no significant differences between the populations involved. Consequently, our analyses could not convincingly prove or disprove foul play. The latter ambiguity was most likely the result of sampling constraints rather than the suitability of the markers. Our results provide a foundation for future application of microsatellite markers for forensic investigation, marker-assisted breeding, and population diversity analysis in koi.**

Forensic science is commonly applied in fisheries science in cases involving misappropriation of resources, misrepresentation of fisheries products, and illegal commerce in fish or fisheries products.[1] Animal forensic science services are well established in South Africa, with accredited routine screening services for domestic animals. Protocols for screening terrestrial agricultural species are standardized internationally, using microsatellite DNA markers (tandem repeats of motifs of 1–8 nucleotides at particular loci in the genome). In recent years, South African forensic services have been expanded to include game species and marine fish stock identification.[2] In this article, we describe the use of microsatellite markers to resolve a forensic case involving koi carp (that is, ornamental strains of common carp, *Cyprinus carpio*) allegedly stolen from a South African breeder, the claimant, and offered for sale by two other breeders. An extensive description of koi varieties is provided by Axelrod.[3] RAPD markers were previously tested for molecular verification of koi colour variations,[4] but we preferred the higher repeatability and evidentiary acceptability of the microsatellite technique.

It is critical that the breeding system used by the claimant is understood. Males and females showing desirable traits are crossed, with two fish of each sex used and with free mixing of eggs and sperm cells. Offspring in the pooled $F_1$ generation then are assigned to 'strains' after a few months. There is thus no classic line breeding for specific colour variants; a single cross can result in offspring assigned to a number of colour variants, given the likely multi-locus system of inheritance of koi colour patterns.

We sampled 27 fish from the claimant (C1) and 84/29, respectively, from the two breeders accused of selling the allegedly misappropriated koi (A1/A2). In addition, we sampled 19 fish originating from the claimant, but which had been supplied to a retailer lawfully (C2). Each koi sampled was already assigned to a nominal colour variant. The colour variants involved were nominally Kohaku Muji, Kumonryo, Ogon Gin Matsuba, Sanke, speckled Sanke, Showa, Shusui, Soragou Ginrin, Utsuri (Hi), Utsuri (Shiro) and Yambuki. All fish sampled from populations C1, A1, A2 and C2 potentially represented the same pooled $F_1$ generation from two males and two females. The $P_1$ fish were either sold or dead, which severely limited the options for analysis.

A fin clip (~1 $cm^2$) was taken from each individual screened and DNA was extracted using a Chelex-based technique.[5] Selection of four suitable microsatellite loci from among the 32 developed for common carp by Crooimans *et al.*[6] was based on a common annealing temperature and sufficient (~75 base-pair) differences between sizes of amplification products among loci, contributing to the ability to multiplex primers for PCR reactions. All microsatellites were of the poly-(CA) type.

Reaction mixtures consisted of 50–200 ng genomic DNA, 0.75 U *Taq* polymerase (Amplitaq Gold©, Applied Biosystems), 1× buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP and 4 pmol of each primer, with double-distilled $H_2O$ added to a total volume of 12 $\mu$l. Forward primers were labelled with 6-FAM, JOE or 6-TAMRA; and ROX 350 was used as the size standard. Reaction conditions were 10 min at 95°C, followed by 35 cycles each of: 45 s at 95°C, 80 s at 58°C and 80 s at 72°C, and with a final extension step of 10 min at 72°C. Fragments were separated on an ABI377 automated sequencer (Applied Biosystems) and scored using GENESCAN and GENOTYPER software.

All four primer pairs used provided a range of polymorphic amplification fragments within and among koi strains (Tables 1 and 2). An average number of 7.75 alleles was observed per locus (Table 1). This is higher than the average of 4.7 alleles per locus resolved by Crooimans *et al.*;[6] the difference is most likely attributable to larger sample sizes and to a wider variety of strains screened, since the latter authors used a panel of eight common carp, compared to the 159 that we screened. Further, the carp used by Crooimans *et al.* represented one inbred line,[6] whereas the koi strains were derived from multiple founding events and intense selection from a highly heterogeneous common carp gene pool. Fragment size ranges we observed for the *MFW09, MFW13* and *MFW20* loci (Table 1) were comparable to those published by Crooimans *et al.*[6] The fragment lengths resolved for the *MFW16* locus are shorter than those reported by Crooimans *et al.*, which may again reflect the small sample size of the latter authors, or differentiation from common carp during the selec-

**Table 1**. Polymorphism and amplification fragment sizes of four microsatellite markers screened in koi carp.

| Locus | No. of alleles | Fragment length (bases) |
|---|---|---|
| *MFW09* | 7 [6]* | 84–116 [117] |
| *MFW13* | 11 [5] | 161–205 [184] |
| *MFW16* | 4 [7] | 130–132 [171] |
| *MFW20* | 9 [6] | 213–231 [249] |
| Mean | 7.75 ± 2.99 [4.7] | |

*Data for common carp from Crooimans *et al.*[6] are presented in brackets for comparative purposes.

[a]Department of Biodiversity, School of Molecular and Life Sciences, University of Limpopo, Private Bag X1106, Sovenga 0727, South Africa.
[b]Animal Improvement Institute, Agricultural Research Council, Private Bag X2, Irene 1675, South Africa.
[c]Department of Fisheries and Wildlife Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.
*Author for correspondence. E-mail: paulg@ul.ac.za

tion history of the koi.

An analysis of molecular variance (AMOVA[7]) was used to measure the relative distribution of diversity in the sampled fish, using ARLEQUIN[8] software. Results showed that most diversity (86.2%) was present within populations ($P < 0.01$). Differences between populations contributed 13.6% of diversity, whereas diversity between colour variants was very low at 0.22% ($P < 0.01$). These results conform to the breeding system used by the claimant, and the precise reflection also validates the inheritance of the four microsatellite loci used. Based on these AMOVA results and the breeding system used, colour pattern was not considered a factor during subsequent analysis.

To test for the pairwise distribution of population differentiation, we used $R_{ST}$[9] (as a coefficient representative of the stepwise mutation mode) using $R_{ST}$Calc software.[10] Results (Table 2) show no significant differentiation for pairwise combinations of populations ($P = 0.08$–$0.83$). Least differentiation ($R_{ST} = -0.00625$) was observed between populations C1 and A1, followed by A2/C2 ($R_{ST} = -0.019$). Goodman[10] suggested that negative $R_{ST}$ values mean that populations are essentially panmictic. Note that differentiation between C1 and C2 ($R_{ST} = 0.031$) shows less differentiation than between the fish of either of the accused and those of the claimant.

We then attempted to gauge the status of individual fish within the putatively stolen populations. Since no $P_1$ fish were available, parentage analysis such as that available in CERVUS-based analyses[11] was not possible. An assignment test (as implemented in GENECLASS software[12]) was therefore used to determine how frequently individual genotypes could be assigned correctly to their source populations. A distance-based method was used to assign the individual to the closest population, and a probability that the individual belongs to each population was then computed as a measure of confidence. We note that the lack of between-population differentiation and the small number of loci used was not optimal for assignment tests. An assignment test could thus provide only approximate results.

We created a reference data file containing genotypes for koi from the claimant (C1 and C2) and both accused. The assignments of fish [with % probability] were as follows: the majority of fish (51.2% [46.8]) in A1 classified to claimant (24.4% C1; 26.8% C2), 31.7% [66.0] self-classified to A1 and 17.1% [52.5] to A2. For the fish in population A2, results were: 39.6% [36.4] to claimant (24.1% C1; 13.8% C2), 44.8% [63.0] self-classified to A2 and 17.2% [40%] to A1. By comparison, for population C2, 42.11% [29.6] self-classified correctly, with 21.1% each assigned to C1 and A2 [49.5; 61.8] and 10.5% [2.5] to A1. Note that equal numbers from population C2 were assigned to C1 and A2, suggesting no difference between the relationship of C2 with the known sibling population C1, compared to that with accused population A2. In conjunction with the high number of A1 which classified to the claimant, these results suggest close overall similarity among all populations screened.

The final statistical approach used was a basic Mendelian analysis, screening for the presence/absence of specific alleles in the putative source and recipient populations (Table 3). Five alleles not scored in the claimant's population were scored in population A1, and four in A2. Note, however, that even the most prolific of these unique alleles occurred at a

**Table 2**. Pairwise differentiation ($R_{ST}$), gene flow (Nm) and significance of differentiation among four populations of koi carp. See text for abbreviations used.

| | Claimant — C1 | Accused 1 — A1 | Accused 2 — A2 |
|---|---|---|---|
| Accused 1 — A1 | $R_{ST} = -0.006$<br>(Nm = infinite)<br>$P = 0.78$ | – | – |
| Accused 2 A2 | $R_{ST} = 0.051$<br>(Nm = 4.63)<br>$P = 0.10$ | $R_{ST} = 0.054$<br>(Nm = 4.33)<br>$P = 0.08$ | –<br>$R_{ST} = -0.019$ |
| Control C2 | $R_{ST} = 0.031$<br>(Nm = 7.85)<br>$P = 0.29$ | $R_{ST} = 0.040$<br>(Nm = 6.03)<br>$P = 0.16$ | (Nm = infinite)<br>$P = 0.83$ |

**Table 3**. Allele frequencies observed in four populations of koi.

| Locus | Allele | Population | | | |
|---|---|---|---|---|---|
| | | Claimant (C1) | Accused 1 (A1) | Accused 2 (A2) | Control (C2) |
| *MFW09* | 84 | 0.559 | 0.318 | 0.259 | 0.556 |
| | 86 | 0.059 | – | – | – |
| | 88 | 0.206 | 0.240 | 0.519 | 0.333 |
| | 90 | 0.059 | 0.351 | 0.222 | 0.111 |
| | 92 | 0.059 | 0.026 | – | – |
| | 94 | – | **0.032*** | – | – |
| | 98 | 0.029 | 0.013 | – | – |
| | 116 | 0.029 | – | – | – |
| | 120 | – | **0.019** | – | – |
| *MFW13* | 161 | 0.020 | – | – | – |
| | 163 | – | – | – | **0.071** |
| | 165 | – | **0.022** | – | **0.143** |
| | 167 | 0.260 | 0.268 | 0.448 | 0.214 |
| | 169 | – | – | **0.017** | – |
| | 171 | – | – | **0.017** | – |
| | 177 | 0.020 | – | – | – |
| | 185 | 0.060 | 0.065 | 0.121 | 0.036 |
| | 187 | 0.260 | 0.130 | 0.086 | 0.143 |
| | 189 | 0.200 | 0.413 | 0.224 | 0.321 |
| | 199 | 0.160 | 0.072 | 0.086 | 0.071 |
| | 201 | 0.020 | 0.014 | – | – |
| | 205 | – | **0.014** | – | – |
| *MFW16* | 130 | 0.981 | 0.976 | 1.000 | 0.947 |
| | 132 | – | **0.012** | – | – |
| | 136 | – | **0.012** | – | **0.053** |
| | 138 | 0.019 | – | – | – |
| *MFW20* | 213 | – | **0.015** | **0.021** | **0.167** |
| | 215 | – | – | – | **0.083** |
| | 217 | 0.280 | 0.462 | 0.438 | 0.458 |
| | 219 | 0.620 | 0.477 | 0.458 | 0.208 |
| | 221 | – | **0.008** | – | – |
| | 225 | – | – | **0.083** | – |
| | 227 | 0.040 | 0.015 | – | – |
| | 229 | 0.020 | – | – | – |
| | 231 | 0.040 | 0.023 | – | 0.083 |

*Frequencies in bold denote alleles in given populations not present in the presumed parental population.

frequency of only 0.032 (for *MFW09* allele 94 in A1). To determine the possibility of exclusion through sampling error, we pooled genotypes from all populations to simulate a hypothetic pre-subdivision $F_1$ population. We then calculated Hardy-Weinberg expectations for allele 94 of *MFW09* in a sample of 84 A1 fish taken from the claimant. The expected occurrence is only 3 fish, a number that could indeed potentially not be present in a population of 84 due to sampling error. The presence of apparently unique alleles in populations A1 and A2 thus does not necessarily provide proof that the fish were not stolen. Also note that three apparently unique alleles are present in population C2 as well (and at greater frequencies than in A1 and A2), yet the known history of group C2 proves that these alleles can be present only because of sampling error.

In summary, there was very little differentiation between fish from the claimant and accused, but we note that any similarities revealed by these results also could reflect a low overall level of differentiation within koi strains, resulting from the shared selection history of these ornamental fish. Foul play cannot be proven unconditionally, but also there is no way in which the current results can be used to acquit the accused. The ambiguous outcome is most probably the result of the sampling strategy enforced by circumstances, since no $P_1$ fish of allegedly stolen stock were available. Including a larger number of microsatellite loci may have assisted, but it should be borne in mind that forensic analysis should remain affordable. Screening of maternally inherited mitochondrial DNA markers may have yielded additional insights. However, given that effective population size for mitochondrial markers is one quarter that for nuclear markers,[13] we believe that variation in highly selected and bottlenecked stocks may have proved limiting, providing few tests of innocence or guilt.

Although we could not definitively prove or disprove the allegation of theft, our data revealed microsatellite variation in koi carp. This knowledge could prove useful for marker-assisted breeding of koi or for testing in other forensic cases where potential parents of progeny of interest are available.

1. Hallerman E. (2003). Forensics. In *Population Genetics: Principles and Applications for Fisheries Scientists,* ed. E. Hallerman, pp. 395–402. American Fisheries Society, Bethesda, MD.
2. Flint N.S., Van der Bank F.H., Theron P.J. and Staby A. (1998/1999). Genetic variation in two populations of the orange roughy (*Hoplostethus atlanticus*) from Namibia. *S. Afr. J. Aquat. Sci.* **24**(1/2), 71–83.
3. Axelrod R. (1988). *Koi Varieties: Japanese colored carp — nishikigoi.* TFH Publications, Neptune City, NJ.
4. Jackson K., Goldberg D., Yehuda Y. and Degani G. (2000). Molecular DNA variation in koi (*Cyprinus carpio*) of various color patterns. *Israeli J. Aquaculture — Bamidgeh* **52**(4), 151–158.
5. Walsh P.S., Metzger D.A. and Higuchi R. (1991). Chelex 100 as medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**(4), 506–513.
6. Crooimans R.P.M.A., Bierbooms V.A.F., Komen J., Van der Poel J.J. and Groenen M.A.M. (1997). Microsatellite markers in common carp (*Cyprinus carpio* L.). *Animal Genetics* **28**, 129–134.
7. Michalakis Y. and Excoffier L. (1996). A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 1061–1064.
8. Schneider, S., Roessli, D. and Excoffier L. (2000). Arlequin, ver. 2.000: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
9. Slatkin M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–462.
10. Goodman S.J. (1997). RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.* **6**, 881–885.
11. Marshall T. (2001). CERVUS. http://helios.bto.ed.ac.uk/evolgen
12. Cornuet J.M., Piry S., Luikart G., Estoup A. and Solignac M. (1999). New methods employing mutinous genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989–2000.
13. Billington N. (2003). Mitochondrial DNA. In *Population Genetics: Principles and Applications for Fisheries Scientists,* ed. E. Hallerman, pp. 59–100. American Fisheries Society, Bethesda, MD.