Mol Genet Genomics (2001) 266: 353–362
DOI 10.1007/s004380100569

ORIGINAL PAPER

L. David · P. Rajasekaran · J. Fang · J. Hillel · U. Lavi

Polymorphism in ornamental and common carp strains (*Cyprinus carpio* L.) as revealed by AFLP analysis and a new set of microsatellite markers

Received: 2 April 2001 / Accepted: 20 July 2001 / Published online: 4 September 2001

Abstract Forty-seven new microsatellite markers were generated and applied, together with the AFLP (Amplified Fragment Length Polymorphism) technique using two different enzyme combinations, to the genetic analysis of two carp species, *Cyprinus carpio* L. and *Ctenopharyngodon idella*. The extent of polymorphism and the genetic relationships between nine carp populations were studied. The incidence of microsatellites containing CA and CT motifs was estimated to be one every 17.4 and one every 126.3 kb, respectively, and their average allele numbers were four and five, respectively. Across populations, the average proportion of individuals that were heterozygous for microsatellite markers was 44.2% and the average allele number was 4.02. The *Eco*RI/*Taq*I combination generated more analyzable AFLP bands than the *Eco*RI/*Mse*I pair, making the former preferable for the analysis of carp populations. The proportion of polymorphic AFLP bands within populations ranged from 6.7% in grass carp to 59.9% in Kohaku strain (Koi) of the ornamental carp. The fixation index (F_{ST}) for microsatellites in these populations was estimated to be 0.37, and for AFLP markers the value was 0.39. Genetic distance matrices derived from microsatellites and from two AFLP analyses were positively correlated. Grass carp showed fewer AFLP bands than other populations and was genotyped by only half of the microsatellite markers. These findings agree with genetic distance estimates in suggesting that the grass carp is phylogenetically quite remote from all the other populations examined.

Keywords Koi carp · Grass carp · Microsatellite · AFLP · Genetic distance

Communicated by G. P. Georgiev

L. David · J. Hillel
Department of Field Crops and Genetics,
Faculty of Agriculture, Food and Environmental Sciences,
The Hebrew University of Jerusalem,
P.O.Box 12, Rehovot 76100, Israel

P. Rajasekaran
Unit of Applied Genetics,
Scottish Crop Research Institute,
Invergowrie DD2 5DA, Dundee, Scotland, UK

J. Fang
College of Horticulture,
Nanjing Agricultural University,
Nanjing 210095, People's Republic of China

U. Lavi (✉)
Institute of Horticulture, ARO-Volcani Center,
P.O.Box 6, Bet-Dagan 50250, Israel
E-mail: ulavi@agri.gov.il
Fax: +972-3-9669583

Electronic supplementary material to this paper can be obtained by using the Springer Link server located at http://dx.doi.org/10.1007/s004380100569.

Introduction

The Japanese ornamental carp (or Koi) is a colored variant of the common carp (*Cyprinus carpio* L.) that has been developed in Japan since the 17th century. Traditional breeding using positive assortative mating is the most commonly practiced method for production of excellent Koi (S. Rothbard, personal communication).

The common carp is an important freshwater aquaculture species, widely distributed from Southeast Asia to European and Mediterranean countries. The European races originated by introduction from the Far East. The carp has been intensively bred since domestication, resulting in a large number of strains and races showing morphological and physiological variability. Kohlmann and Kersten (1999) used isozyme markers to show that Koi is genetically more distant from edible carps than any of the edible carps are from one another. They also showed that Koi is genetically closer to the Far Eastern strains of carps than to the European strains.

A small population of the albino mutant of grass carp (*Ctenopharyngodon idella*) was introduced into Israel from Japan in 1988 and since then has been produced

commercially as an ornamental fish (S. Rothbard, personal communication).

Koi and common carp belong to the same species but, having been selected for different purposes, they differ in many morphological and other traits. Both the grass carp and the common carp are classified in the Subfamily Cyprininae within the Family Cyprinidae.

Microsatellite markers, also named SSRs (Simple Sequence Repeats), have been developed for large number of organisms for various purposes. These markers are codominantly inherited, highly polymorphic, abundant throughout genomes and can be detected with high reproducibility. These markers are therefore considered as markers of choice for many linkage and biodiversity studies (Bowcock et al. 1994; MacHugh et al. 1998). Microsatellite markers have also been developed and used in studies involving a number of fish species, including zebrafish (*Brachydanio rerio*; Knapik et al. 1998), Japanese pufferfish (*Fugu rubripes*; Edwards et al. 1998), tilapia (*Oreochromis niloticus*; Kocher et al. 1998), salmon (*Oncorhynchus* spp.; Sletan et al. 1997) and catfish (*Ictalurus* spp.; Liu et al. 1999). Some 32 microsatellite markers were developed for the common carp by Crooijmans et al. (1997) and three microsatellites were developed from a genomic library constructed from Koi DNA by Aliah et al. (1999).

AFLP markers (Vos et al. 1995) provide an efficient method for DNA fingerprinting in many species of plants and animals (Ajmone-Marsan et al. 1997; Besse et al. 1998). Two main factors affect the total number of AFLP bands generated per individual: the specific endonucleases used and the number of selective bases added to the primers in the selective amplification step. For all complex genome organisms a two-step selective amplification is recommended (Vos et al. 1995). AFLP analysis has also been applied to several fish species – to tilapia (*Oreochromis* spp.) by Agresti et al. (2000), to rainbow trout (*Onchorynchus mykiss*) by Young et al. (1998) and to catfish (*Ictalurus* spp.) by Liu et al. (1998). The AFLP markers in these studies were generated using *Eco*RI and *Mse*I for restriction digestion.

In the present study we have generated 47 new microsatellite markers and used them, together with two methods of AFLP, to analyze polymorphism and genetic relationships between and among various carp populations.

Materials and methods

Fish populations

Nine fish populations were analyzed: (1) the albino mutant of grass carp (*Ctenophryngodon idella*); three strains of the common carp (*Cyprinus carpio*), (2) a Yugoslavian strain (Moav et al. 1975), (3) the Dor 70 strain (Moav et al. 1975) and (4) an Israeli commercial hybrid (Moav et al. 1975); and five variants of Koi (*Cyprinus carpio* var. Koi), (5) the white variant (Platinum Ohgon), (6) the black variant (Karasu), (7) the Kohaku variant, (8) the Sanke variant, and (9) the Showa variant.

AFLP analysis using the enzyme combination *Eco*RI/*Mse*I method was performed on five individual fish randomly selected from each of the above-mentioned nine populations. For AFLP analysis with *Eco*RI and *Taq*I, and for microsatellite analyses, one individual was chosen from each population.

Samples of all populations were kindly provided by the Gan-Shmuel Fish Breeding Center, Gan-Shmuel, Israel.

Generation of microsatellite markers

Genomic DNA was extracted from Koi (Showa) blood using the Purgene kit (Gentra Systems, Minneapolis, Minn.). The genomic DNA library was prepared by DNA Technologies (Gaithersburg, Md.). The library was constructed in pBluescript from DNA digested with *Bam*HI + *Eco*RI and size selected for fragments of 0.5–0.7 kb. About 5200 clones were screened using a mixture of synthetic ^{32}P end-labeled $(GT)_{12}$ and $(GA)_{12}$ oligonucleotides. Plasmids were isolated from positive clones using the High Pure Plasmid Isolation kit (Boehringer Mannheim, Mannheim, Germany). Sequencing was carried out using the BigDye Terminator kit (PE Applied Biosystems, Foster City, Calif.) with the KS (Bluescript) or the P79 primer (5′-GGTGGCGGCCGCTCTAG-3′) on an ABI Prism 377 DNA Sequencer (Perkin Elmer, Foster City, Calif.). Primers were designed, from sequences flanking repeats of at least 16 bp long, using Oligo 4.04 Primer Design software (National Biosciences, Plymouth, Minn.). The microsatellites and their sequences are listed in Table 1 in Electronic Supplementary Material (http://dx.doi.org/10.1007/s004380100569). Four more microsatellite markers were retrieved from the literature, MFW 4 and MFW 7 (Crooijmans et al. 1997) and CCA 17 and CCA 30 (Aliah et al. 1999).

Genotyping of microsatellite marker loci

PCRs contained 70 ng of DNA template, 0.5 U of AmpliTaq Gold polymerase (Perkin Elmer), 1 µl of 10× AmpliTaq Gold polymerase buffer, 0.25 mM of each of the four dNTPs, 2.5 mM $MgCl_2$ and 5 pmol of primers in a total volume of 10 µl. For some of the microsatellite markers one primer was labeled at the 5′ end with fluorescent HEX, FAM or TET. The other markers were used for genotyping with unlabeled primers and the reaction included 1 µM fluorescently labeled dUTP (PE Applied Biosystems). The PCR program consisted of 12 min at 96°C, followed by 35 cycles of 96°C for 30 s, 1 min at the primer-specific annealing temperature (see Table 1 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569) and 72°C for 1 min, and a final extension step at 72°C for 15 min. Reaction products were mixed 1:1 with an internal size standard (*gen*SIZE R350 or T350, Genpak, Brighton, UK) diluted 1:4 with loading buffer. The mixture was fractionated in a 6% polyacrylamide gel (Bio Lab, Jerusalem, Israel) on an ABI Prism 377 DNA Sequencer (Perkin Elmer). Sizing of fragments was carried out using the Genscan 2.0 software (Perkin Elmer). Genotypes were determined using the software Genotyper 2.0 (Perkin Elmer).

AFLP analysis

Two combinations of endonucleases were used. For the combination *Eco*RI/*Mse*I, genomic DNA (300–500 ng) was incubated for 2 h at 37°C with 2 U of *Mse*I, 5 U of *Eco*RI, 1.2 U of T4 DNA ligase (New England Biolabs, Beverly, Mass.), 50 pmol of *Mse*I adapters and 5 pmol of *Eco*RI adapters (see Table 2 in Electronic Supplementary material for sequences, http://dx.doi.org/10.1007/s004380100569). This reaction was done in a volume of 50 µl of restriction-ligation buffer containing 10 mM TRIS-acetate pH 7.5, 10 mM $MgCl_2$, 50 mM potassium acetate, 5 mM dithiothreitol (DTT), 1 mM ATP and 50 ng/µl BSA.

For digestion with *Eco*RI/*Taq*I, genomic DNA (300–500 ng) was digested at 65°C for 1.5 h with 5 U of *Taq*I in a volume of

Table 1 Prevalence of microsatellite motifs

Motif[a]	Number of positive clones[b]	Percentage of total in library[c]	Average distance between arrays (kb)	Average size of array (bp)[d]	Proportion of the genome (%)[e]
CA/GT	80	2.9	17.4	39	0.222
CT/GA	11	0.4	126.3	32	0.026
CA + CT	13	0.5	106.9	77	0.072
CA + Other	28	1.0	49.6	80	0.162
CT + Other	5	0.2	277.9	96	0.034
CA + CT + Other	12	0.4	115.8	93	0.080
All CA types	133	4.9	10.5	51	0.488
All CT types	41	1.5	33.9	70	0.208

[a]The values for each type of motif are based on the analysis of positive clones identified by screening the genomic library with the various microsatellites and on sequencing data
[b]Number of clones containing the corresponding motif based on sequencing data
[c]Proportion of clones in the library that contain the corresponding motif
[d]Average array size of microsatellite type in the library
[e]Ratio of the average array size to the average distance between two such motifs

Table 2 Polymorphism in nine populations based on the analysis of 47 microsatellite markers

Population[a]	Average number of alleles per locus	Range of allele number per locus	Percentage of heterozygous loci	Number of markers yielding more than two PCR products
Grass carp	1.4	1–2	43.5	0 (0%)
Yugoslavian carp	1.8	1–4	50.0	10 (22.7%)
Dor-70 carp	1.6	1–4	46.7	6 (13.3%)
Commercial carp	1.4	1–3	36.4	3 (6.8%)
White Koi	1.7	1–4	45.2	6 (14.3%)
Black Koi	1.6	1–4	40.5	5 (11.9%)
Kohaku Koi	1.6	1–4	45.2	5 (11.9%)
Sanke Koi	1.7	1–4	52.4	6 (14.3%)
Showa Koi	1.6	1–4	47.7	5 (11.4%)

[a]The data are based on the use of a single individual from each of the nine populations

25 µl of the restriction-ligation buffer described above. The reaction was cooled to 37°C, supplemented with 15 µl of the restriction-ligation buffer containing 5 U of *Eco*RI and incubated at 37°C for an additional 2 h. For adapter ligation, 10 µl of the restriction-ligation buffer, containing 50 pmol of *Taq*I adapters, 5 pmol of *Eco*RI adapters (see Table 2 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569), 0.5 mM ATP and 1.2 U of T4 DNA ligase, was added, and the reaction was incubated in 37°C for 3 h.

In both cases, a 30-µl aliquot of the adapter-ligated DNA was diluted 1:10 with distilled water to serve as template in the preselective PCR. The remaining 20-µl portion was used to verify that digestion was complete.

The preselective PCR contained 5 µl of template, 1 U of AmpliTaq polymerase, 2 µl of 10× AmpliTaq polymerase buffer, 0.25 mM of each of the four dNTPs, 2.5 mM $MgCl_2$ and 25 ng of *Eco*RI and either *Taq*I or *Mse*I primers with one selective nucleotide (A) (see Table 2 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569), in a total volume of 20 µl. The PCR program consisted of thirty cycles of 30 s at 94°C, 1 min at 56°C and 1 min at 72°C, followed by 10 min at 72°C.

The selective PCR contained 5 µl of the diluted (1:20) product of the preselective PCR, 0.5 U of AmpliTaq polymerase, 2 µl of 10× AmpliTaq polymerase buffer, dNTPs and $MgCl_2$ as mentioned above, 50 ng of *Taq*I or *Mse*I selective primer and 5 ng of ^{33}P end-labeled *Eco*RI selective primer (see Table 2 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569), in a total volume of 20 µl. The first amplification cycle was carried out for 30 s at 94°C, 30 s at 65°C and 1 min at 72°C. In each of the following 10 cycles, the annealing temperature was reduced by 1°C. The last 25 cycles were carried out at an annealing temperature of 56°C, and the final extension step was carried out at 72°C for 10 min.

Each sample was diluted 1:1 with loading buffer, denatured and fractionated on a 5% polyacrylamide sequencing gel in TRIS-borate-EDTA buffer. Gels were run at constant power (110 W), and then dried for 1 h in 80°C in a vacuum gel dryer. The gel was exposed to Kodak Biomax MR film for 2–5 days and the band pattern was read manually.

Data analysis

The AFLP band patterns were encoded in a matrix (using 1 or 0 for the presence or absence of a band, respectively) for each individual and for each band. A subset of clearly delineated bands was chosen in each gel and analyzed across all lanes, resulting in a scoring matrix. Estimation of genetic distance and bootstrapping for AFLP data were done with Phyltools software (Phylogenetic Computer Tools v. 1.2; available from J. B. Buntjer, Wageningen University, Wageningen, The Netherlands) or with PHYLIP (Phylogeny Inference Package version 3.57c; available from J. Felsenstein, University of Washington, Seattle, Wash.). Genetic distances were calculated according to Nei's index (Nei and Li 1979) for data obtained from AFLP analysis and according to Reynolds's index (Reynolds et al. 1983) for microsatellite data. Analysis of Molecular Variance (AMOVA) and comparison of genetic distances matrices (Mantel test) were done using Arlequin software (Arlequin: a software for population genetics data analysis, Ver 2.000; developed by and available from S. Schneider, D. Roessli and L. Excoffier, Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva), for both AFLP and SSR data. Other statistical analyses were done using JMP software (SAS Institute, Cary, N.C.).

Results

Generation of microsatellite markers and estimation of their prevalence

The genomic library of Showa Koi (5184 clones) was screened with a mixture of $(GT)_{12}$ and $(GA)_{12}$ olgonucleotides. Two hundred and ninety one clones (5.61%) were positive. On this basis, the average frequency of these microsatellites in the genome was calculated to be one per 8.9 kb. A total of 156 clones were sequenced, of which two (1.28%) were found not to include SSRs. Of the remaining 154 clones, 149 contained a CA or a CT repeat and in five clones another type of repeat was found (see Table 1 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569). Fifty-eight of the 149 clones (38.9%) were found to have a compound motif. The frequencies of the different motifs, based on an average insert size of 500 bp, are given in Table 1. The frequency of CA motifs in the Koi genome is about seven times higher than that of CT motifs. Compound motifs containing CA are also more abundant than the simple CT motif. The compound motifs, containing CA or CT, are less prevalent than the respective simple ones. Within each of the two motif types examined (CA or CT), compound repeats form longer arrays than simple ones. These microsatellite sequences are estimated to account for about 0.6% of the sequences in the Koi genomic library (Table 1).

About 40 sequences from common carp deposited in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) were also screened for the presence of microsatellites. The screen was carried out with all possible di-, tri- and tetranucleotide repeat sequences and yielded 18 positive hits. Primers were designed for 12 sequences allowing the generation of seven SSR markers (see Table 1 in Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569).

Assessment of polymorphism using microsatellite markers

Forty-three of the microsatellite markers were used to assess the degree of polymorphism in a panel made up of one fish from each of the nine populations. The numbers of alleles detected for each of these microsatellites at each locus and the numbers of heterozygous fish are listed in Table 1 of the Electronic Supplementary Material http://dx.doi.org/10.1007/s004380100569. Five markers were homozygous in all nine individuals and five markers were heterozygous in all nine individuals. The remaining markers identified various numbers of heterozygous fish in the sample. The average proportion of heterozygous individuals was 44.2% and the average number of alleles across the nine individuals and across all markers was 4.02. No significant difference was found either in number of alleles or in number of heterozygous individuals between CA- and CT-type loci.

Out of 47 SSRs used (including four markers retrieved from the literature), only 23 (49%) could be used for genotyping the grass carp.

At twelve of the 47 SSR loci (25.5%), more than two PCR products per individual were observed in at least one of the nine individuals. Excluding these markers, the average proportion of heterozygous individuals is 40.6% and the average number of alleles across nine individuals and across all markers is 3.95.

Data on the average number of alleles per locus and the percentage of heterozygous loci, for each of the nine fish, are presented in Table 2. Both these estimates show low variability among the different populations. The number of alleles per locus ranges between 1.43 and 1.77, and the percentage of heterozygous loci ranges between 36.4 and 52.4. In all populations (except for the grass carp), markers showing more than two PCR products were found; such markers made up between 6.8 and 22.7% of the total, depending on the population.

Analysis of Molecular Variance (AMOVA) was carried out on the data for 41 SSRs. The variance among populations was estimated to be 576.6, equivalent to 37% of the total variance [i.e., the estimated fixation index (F_{ST}) is 0.37].

Comparison of AFLP levels detected with different enzyme and selective-primer combinations

AFLP band patterns generated using seven *Eco*RI/*Mse*I and six *Eco*RI/*Taq*I selective combinations were analyzed. An example of an AFLP pattern produced using the *Eco*RI-ACA/*Mse*I-CTG primer combination is presented in Fig. 1. Only clearly delineated bands within the range of 70–450 bp were scored. The average band number per lane was 79 and 65 for the *Eco*RI/*Mse*I and *Eco*RI/*Taq*I combinations, respectively, of which an average of 27 (34%) and 44 (67.7%) bands were analyzed. For each enzyme pair, band number varied depending on the primer combination employed, but no pattern had more than 96 bands, making all the primer combinations used analyzable.

The AFLP characteristics of the different fish populations, as determined with both enzyme combinations are summarized in Table 3. All fish populations, except the grass carp, showed similar average numbers of bands per lane. The grass carp showed on average 17% and 26% fewer bands per lane in the *Mse*I and *Taq*I experiments, respectively. On each gel the area with the clearest bands was chosen for analysis. These bands were scored across all individuals and recorded as presence/absence scoring matrices. Scoring matrices for the *Mse*I experiments contained 75 bands on average and these for *Taq*I 106 bands on average. An average of 74 out of the 75 bands scored on *Mse*I gels (99%) were polymorphic per scoring matrix across all 45 individuals, while an average of 32 bands (43%) were polymorphic within fish populations. For the *Taq*I data, an average of 88 out of 106 (83%) polymorphic bands per scoring matrix was

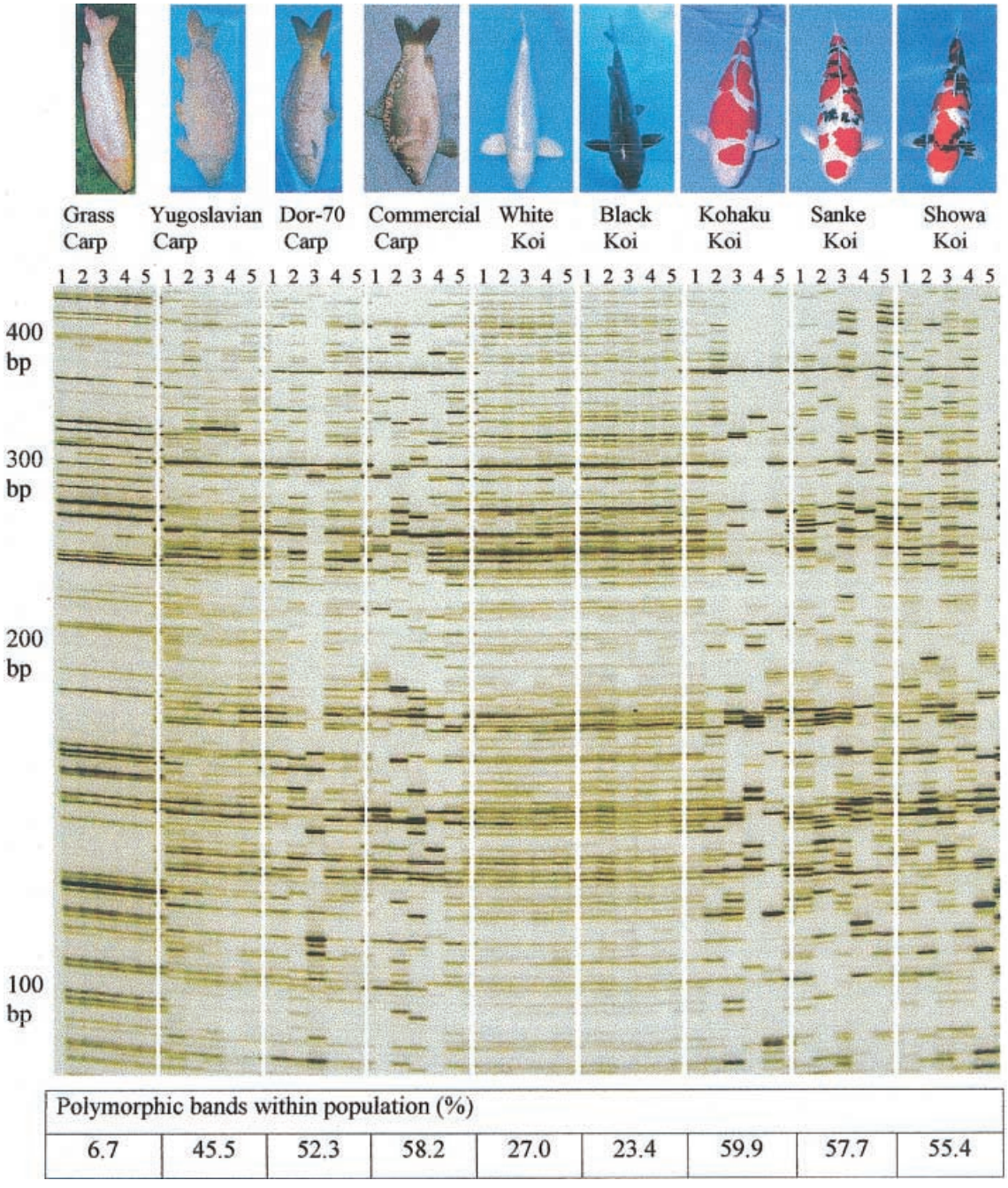

Fig. 1 AFLP patterns (*Eco*RI-ACA/*Mse*I-CTG primer combination) and polymorphism in nine fish populations, each represented by five individuals

calculated across nine individuals. The average proportion of unique bands (present in one population and absent in all other populations) was also determined. In the grass carp 10.3% and 39.7% of scored bands were unique, based on *Mse*I and *Taq*I data, respectively. All the other eight fish populations had less than 1.1% unique bands per *Mse*I combination and less than 1.7% per *Taq*I combination (Table 3). The grass carp accounted for an average of 68.5% (*Mse*I) and 82.4% (*Taq*I) of the total number of unique bands scored.

Assessment of polymorphism using AFLP markers

The AFLP method based on *Eco*RI/*Mse*I digestion was used to genotype 45 individuals (five individuals from each of the nine populations). Seven selective primer combinations were used, resulting in a total of 505 bands. All analyzed bands were polymorphic when scored across the 45 individuals. AFLP patterns visually reflect the polymorphism level of these nine populations (Fig. 1). The percentage of polymorphic bands within populations gives a direct estimate of the level of genetic diversity. The proportion of polymorphic bands within populations varies among the different populations, roughly separating them into four groups: grass carp

Table 3 Characterization of the nine carp populations by two AFLP methods

Population	Average number of bands per lane[a]		Average number of bands per scoring matrix[b]		Polymorphic bands within population (%)[c]	Unique bands (%)[d]	
	*Eco*RI/*Mse*I	*Eco*RI/*Taq*I	*Eco*RI/*Mse*I	*Eco*RI/*Taq*I	*Eco*RI/*Mse*I	*Eco*RI/*Mse*I	*Eco*RI/*Taq*I
Grass carp	66.4	49.1	29.2	28.9	6.7	10.3	39.7
Yugoslavian carp	82.6	61.1	24.2	40.9	45.5	0.3	0.1
Dor-70 carp	77.2	65.3	25.8	45.6	52.3	0.9	0.8
Commercial carp	73.0	68.1	22.2	45.9	58.2	1.1	0.4
White Koi	83.9	66.3	35.8	47.4	27.0	0	1.7
Black Koi	80.9	71.6	34.6	46.6	23.4	0	1.6
Kohaku Koi	81.8	67.1	24.4	45.1	59.9	1.0	0.7
Sanke Koi	80.8	66.9	21.6	47.7	57.7	0.6	1.7
Showa Koi	79.5	68.0	24.4	46.7	55.4	0.8	1.3

[a]Averaged over the seven (*Mse* I) or six (*Taq*I) selective combinations used in each case
[b]A set of clearly delineated bands was chosen in each gel and analyzed across all lanes to generate a scoring matrix
[c]The*Eco*RI/*Taq*I combination was used on only one individual per population, so within-population comparisons cannot be made
[d]Bands present in one population but absent in all others

had 6.7% polymorphic bands, white Koi and black Koi had 27.0% and 23.4% polymorphic bands, respectively and Yugoslavian carp had 45.5% polymorphic bands. The remaining populations (Dor-70 and commercial carp, Kohaku, Sanke and Showa) had between 52.3% and 59.9% polymorphic bands (Table 3 and Fig. 1).

An AMOVA analysis was carried out on each of the seven data sets resulted in an average variance component of 5.17 for the among-populations variance and an estimated F_{ST} of 39.4%. This index ranged between 30.6% (for the primer combination where the among-populations variance was the lowest) and 44.9% (for the most variable one). Each of the estimated F_{ST} indices differed significantly from zero ($P < 0.01$, 1000-permutation test).

Genetic relationships among the fish populations

Assessment of genetic relationships among populations was based on the use of seven selective primer combinations for AFLP with *Eco*RI/*Mse*I. Band frequency was calculated for each of the bands in each of the fish populations separately. These band frequencies were used to estimate Nei's genetic distance statistic among pairs of populations, using the PHYLIP package. Each one of the seven primer combinations was used to estimate pairwise distances. The seven genetic distance matrices were compared pairwise using the Mantel test. All the correlation coefficients were significant ($P < 0.02$, 1000 permutations) and ranged between 0.60 and 0.94. Consequently an average matrix of genetic distances was calculated, bootstrapped (1000 times) and a consensus phylogenetic tree was constructed using the UPGMA method implemented in the PHYLIP package (Fig. 2b). The matrix of average genetic distances is presented in Table 4. The grass carp is seen to be quite remote from all other populations. The white and the black Koi are equidistant from the edible carps and from the rest of the Koi populations. The Kohaku, Sanke and Showa Kois are close to each other and equidistant from all the rest. Bootstrapping values for all nodes of the phylogenetic tree were not significant, ranging between 330 (out of 1000) and 448, except for the branch separating the grass carp, which had a bootstrapping score of 1000.

Nei's genetic distances were also estimated among pairs of populations, based on the use of one individual from each of the nine populations, using six selective primer combinations and the enzyme pair *Eco*RI/*Taq*I. An average matrix of these distances was calculated (lower diagonal, Table 5). A consensus dendogram based on 1000 bootstraps was constructed using the UPGMA method and is presented in Fig. 2c. This method resolves the populations into three groups: grass carp, edible carps and ornamental carps. The matrix of average distances derived from *Eco*RI/*Taq*I data was similar ($R = 0.83$, $P = 0.017$) to the average matrix obtained with *Eco*RI/*Mse*I.

A matrix of genetic distances (Reynolds distance) based on SSRs and on data for one fish from each of the nine populations was calculated (upper diagonal, Table 5). A consensus dendogram based on 1000 bootstraps was constructed using UPGMA method and is presented in Fig. 2a. This dendogram places the Yugoslavian carp together with the ornamental carps. Comparison of this matrix with the matrices based on AFLP data showed a correlation coefficient of 0.631 ($P = 0.037$) with the matrix derived from the *Eco*RI/*Mse*I results and of 0.786 ($P = 0.002$) with the matrix constructed with the *Eco*RI/*Taq*I data.

Estimated Mean Genetic Distance (MGD) for each population (Table 4) correlated negatively ($r = -0.94$, $P = 0.0002$) with the percentage of within-population polymorphic bands (Table 3), both obtained by the same AFLP analysis. Estimates of MGD calculated from *Eco*RI/*Taq*I AFLP data and from SSRs (based on one fish from each population) were higher for grass carp and uniform for the rest of the populations. A positive correlation coefficient of 0.83 ($P = 0.0062$) was found between the latter two MGD sets.

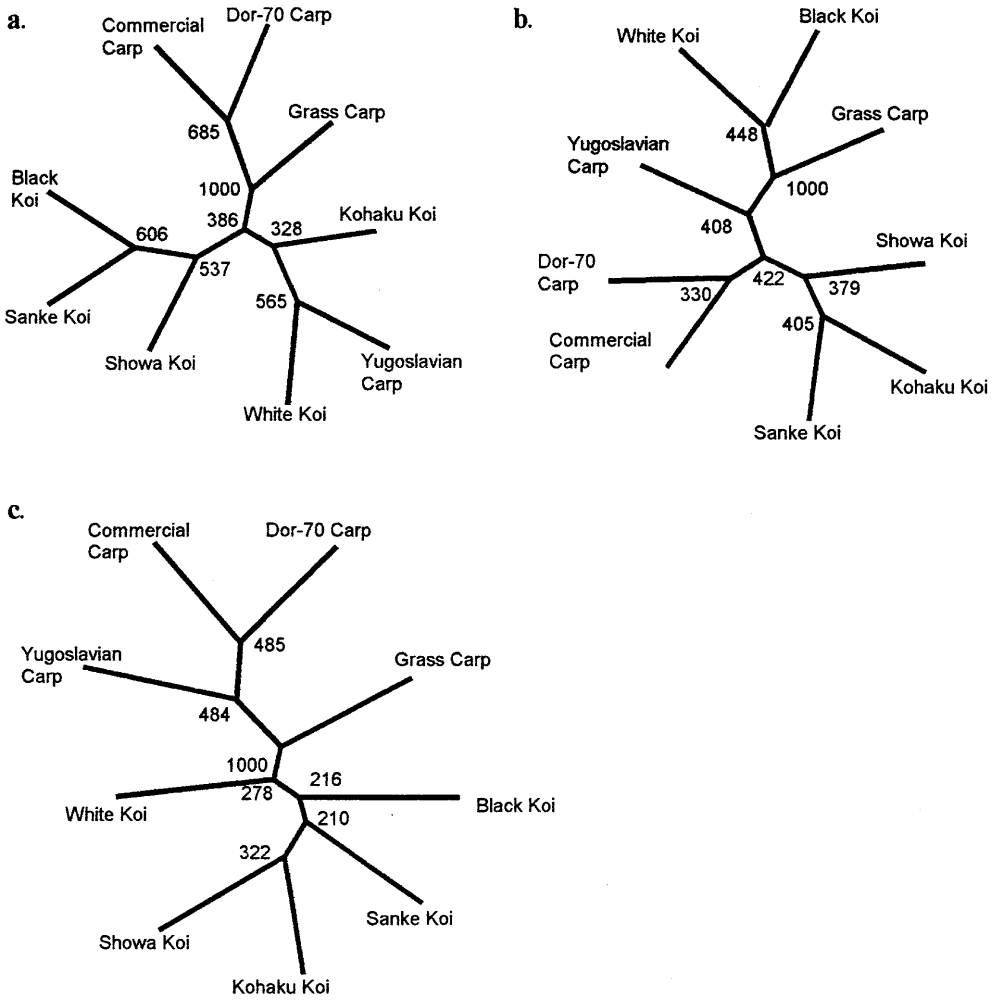

Fig. 2a–c Genetic relationships among the nine fish populations. Dendograms were constructed using the UPGMA method. **a** Tree based on data from one individual per population, genotyped with microsatellite markers. **b** Five individuals representing each population were genotyped by AFLP (*Eco*RI/*Mse*I). **c** One individual representing each population was analyzed using AFLP (*Eco*RI/*Taq*I). The numbers at nodes are bootstrapping values (out of 1000 repetitions)

Discussion

Frequency of microsatellite markers

In the current study, 47 new markers for the common carp were generated. The CA-containing repeats were estimated to occur once every 10.5 kb on average. This estimate is about three times higher than the previous estimate in carp and is closer to the values found in the Atlantic cod (one every 7 kb; Crooijmans et al. 1997) and in the pufferfish (one every 6.6 kb; Edwards et al. 1998). The proportion of compound motifs (40%) was about two times higher than the previously reported estimate for common carp and higher than the values for the Atlantic cod, rainbow trout, or the Atlantic salmon (Crooijmans et al. 1997). These discrepancies can probably be attributed to differences in library construction and/or in screening methodologies.

In Koi, CA motifs occur about seven times more often than CT motifs, as was also found in pufferfish (*Fugu*) by Edwards et al. (1998). The CA motif was found to be the most abundant dinucleotide microsatellite in the *Fugu* genome (and also in some vertebrate genomes) and the CT motif was found to be the fourth most prevalent (Edwards et al. 1998).

No difference was found in the frequency of heterozygotes or in the number of alleles between CA and CT microsatellites. Their prevalence, together with their level of polymorphism, justifies further searches for more CA repeats in common carp and probably for other repeats that have not been examined in this study.

The proportion of the genome occupied by CA repeats in common carp (0.22%) is only half of that estimated in *Fugu* (0.408%). The same is true of CT repeats. This may simply reflect the fact that carp has a larger genome, as the *Fugu* genome is known to be quite compact (400 Mb; Edwards et al. 1998).

AFLP Analysis

AFLP analyses of fish species have been reported by Liu et al. (1998), Spruell et al. (1999) and Agresti et al. (2000). The average number of bands produced in the pink salmon was about 30 (Spruell et al. 1999) and in the catfish it varied from 50 to 120 (Liu et al. 1998). Both these studies used the enzyme combination *Eco*RI/*Mse*I.

Table 4 Average genetic distances (GD) among nine fish populations based on AFLP analysis

Population[a]	Grass carp	Yugoslavian carp	Dor-70 carp	Commercial carp	White Koi	Black Koi	Kohaku Koi	Sanke Koi	Showa Koi
Yugoslavian carp	0.57 (0.19)								
Dor-70	0.48 (0.15)	0.13 (0.16)							
Commercial carp	0.43 (0.11)	0.15 (0.05)	0.08 (0.06)						
White Koi	0.71 (0.21)	0.20 (0.12)	0.19 (0.08)	0.27 (0.13)					
Black Koi	0.73 (0.21)	0.25 (0.15)	0.25 (0.09)	0.32 (0.10)	0.12 (0.19)				
Kohaku Koi	0.44 (0.13)	0.17 (0.06)	0.10 (0.04)	0.08 (0.04)	0.25 (0.10)	0.23 (0.09)			
Sanke Koi	0.43 (0.12)	0.18 (0.07)	0.11 (0.08)	0.09 (0.07)	0.29 (0.11)	0.26 (0.11)	0.05 (0.02)		
Showa Koi	0.42 (0.09)	0.18 (0.05)	0.10 (0.07)	0.09 (0.07)	0.28 (0.11)	0.25 (0.08)	0.06 (0.02)	0.05 (0.02)	
Mean GD	0.52	0.23	0.18	0.19	0.29	0.30	0.17	0.18	0.18

[a]Values of Nei's genetic distance estimate are listed, based on band frequency for seven primer combinations (*Eco*RI/*Mse*I). The standard error of each estimate is given in *parentheses*. Data were obtained from five individuals per population

Table 5 Average genetic distances (GD) among nine fish based on AFLP and SSR markers

Population[a]	Grass carp	Yugoslavian carp	Dor-70 carp	Commercial carp	White Koi	Black Koi	Kohaku Koi	Sanke Koi	Showa Koi	Mean GD[c]
Grass Carp	**0.00**	0.86	0.80	1.29	1.06	0.93	1.06	0.94	1.04	1.00
Yugoslavian carp	0.56 (0.07)	**0.00**	0.14	0.32	0.56	0.49	0.60	0.37	0.50	0.48
Dor-70	0.58 (0.08)	0.08 (0.04)	**0.00**	0.30	0.37	0.51	0.60	0.37	0.41	0.44
Commercial carp	0.58 (0.08)	0.08 (0.03)	0.05 (0.03)	**0.00**	0.69	0.86	0.98	0.69	0.80	0.74
White Koi	0.57 (0.07)	0.25 (0.06)	0.27 (0.04)	0.27 (0.04)	**0.00**	0.68	0.69	0.44	0.59	0.64
Black Koi	0.59 (0.07)	0.25 (0.07)	0.26 (0.07)	0.25 (0.06)	0.19 (0.05)	**0.00**	0.00	0.08	0.18	0.47
Kohaku Koi	0.57 (0.07)	0.26 (0.07)	0.28 (0.05)	0.26 (0.05)	0.19 (0.01)	0.18 (0.03)	**0.00**	0.06	0.22	0.53
Sanke Koi	0.58 (0.08)	0.26 (0.05)	0.28 (0.06)	0.26 (0.06)	0.18 (0.02)	0.19 (0.04)	0.17 (0.02)	**0.00**	0.03	0.37
Showa Koi	0.58 (0.05)	0.26 (0.07)	0.29 (0.05)	0.28 (0.06)	0.18 (0.04)	0.16 (0.02)	0.16 (0.02)	0.17 (0.04)	**0.00**	0.47
Mean GD[b]	0.58	0.25	0.26	0.26	0.26	0.26	0.26	0.26	0.26	

[a]The values *below* the diagonal give the average Nei's genetic distance (upper value) together with the standard error of the estimate (in *parentheses*), based on six AFLP primer combinations (*Eco*RI/*Taq*I). The values *above* the diagonal give Reynolds' genetic distance estimate based on microsatellite markers

[b]Mean Genetic Distance from all others for each fish based on AFLP analysis

[c]Mean Genetic Distance for each fish based on SSR analysis

The carp, which shows an average of 80 bands under these conditions (Table 3) probably has a more complex genome than the salmon. The efficiency of AFLP marker generation, their numbers and their reproducibility are important for genetic analyses (Knorr et al. 1999). The choice of enzyme combination for restriction digestion is thus a major consideration. The difference in average band number per lane indicates that *Taq*I cleaves the carp genome less frequently than *Mse*I. The lower average band number and the higher proportion of analyzable bands produced by the *Eco*RI/*Taq*I pair make it the preferred combination for AFLP analysis in the carp. This is particularly true for polymorphic populations where many bands have to be scored accurately (in the current study, only 18% of the bands generated with *Taq*I were shared by all fish).

Assessment of polymorphism using microsatellite markers

Average heterozygosity of all markers was 44.2%, compared to 60.4% estimated in outbred lines of carp by Crooijmans et al. (1997). The average of four alleles per marker found in this study is similar to the average of 4.7 alleles found by Crooijmans et al. (1997), although these two studies used different test panels of carps. Numbers of alleles and percentage of heterozygous loci across markers were similar in all populations (Table 2). This could result from the use of one fish per population.

Twelve SSR loci out of 47 showed more than two PCR products in one individual. These were found in all strains but the grass carp. This may lead to an overestimation of the allele number and the heterozygosity in common carps. Allele numbers higher than two may result either from PCR artifacts or a higher level of ploidy in the common carp (Crooijmans et al. 1997).

All microsatellite markers could be used for genotyping both edible and ornamental carps. Though they are phenotypically different, the genome is conserved. Conservation of SSRs among different species and families was previously reported (Schlötterer et al. 1991). Here, the level of cross-species amplification (49%) was much lower than in whales (100%) and similar to that in sheep and cow (40%), even though the latter three are all mammals. This estimate may reflect for the relative time of divergence of grass carp and common carp.

Polymorphism estimated using AFLP markers

The proportion of polymorphic bands within populations generated with *Eco*RI/*Mse*I provides an estimate of the level of polymorphism in the populations (Fig. 1 and Table 3). On the basis of these values, the populations could be separated into four groups in ascending order of polymorphism: (1) grass carp; (2) Black and white Kois; (3) Yugoslavian carp; and (4) Dor-70 and commercial carp, Kohaku, Sanke and Showa Kois. These polymorphism levels probably reflect the history of these populations.

The degree of polymorphism within populations revealed by AFLP (*Eco*RI/*Mse*I method) differs from the estimates obtained by SSRs. The number of heterozygous loci in each population (each represented by one fish) estimated by SSRs showed little variance. Variability of the AFLP estimates is probably due to the use of five fish from each population. The lower polymorphism values estimated by AFLP may result from the dominant mode of inheritance of AFLP markers and/or from the different sources of polymorphism between these two marker types.

Distinguishing carp populations using AFLP markers

Unique bands and differences in band frequencies can be used to distinguish between populations. Grass carp can be distinguished from other populations using a single primer combination (Table 3). In order to distinguish between other populations more than one combination must be used.

The average number of unique bands per combination was higher with *Taq*I than with *Mse*I method, probably because the first was used on a single fish per population.

Genetic relationships among fish populations

High F_{ST} values result from similarity within populations and diversity among populations. The estimate of F_{ST} (0.37) using 41 SSR loci demonstrates that high proportion of the molecular variance is due to the differences among-populations. This estimate is similar to the estimate obtained from the AFLP analysis (0.39) corroborating the genetic structure of the selected populations. However, the value of variance-component among groups based on AFLP was much smaller than the one based on the SSRs. This smaller value probably stems from several factors such as: the use of five fish per population for AFLP analysis, dominance of AFLP loci and a higher polymorphism level of SSR loci.

High correlation coefficients found between matrix pairs for different primer combinations (*Eco*RI/*Mse*I method) reflect the ability of even single AFLP combinations to reveal the genetic relationships among these populations. This should probably be attributed to the large number of markers revealed by each combination and to the level of polymorphism within and among these populations.

It is noteworthy that the grass carp could be genotyped by only half of the SSR markers. Exclusion of these markers reduces the number of alleles per locus mainly in polymorphic markers. This may imply on the contribution of different alleles by this species. These results indicate genotypic difference of the grass carp from the common carp species, as indicated also by the lower number of bands revealed by AFLP.

The genetic distance matrix based on all seven combinations of AFLP (*Eco*RI/*Mse*I method) (Table 4) shows the genetic relationships between populations. These results show that grass carp is phylogenetically quite remote from all other populations, as also indicated by the bootstrapping values for the phylogeny. A second (non-significant) distinction is that between polymorphic populations (e.g., edible carps, Kohaku, Sanke and Showa Kois) and less polymorphic ones (e.g., black and white Kois). This may suggest the need to use larger samples from each population to better characterize these polymorphic populations. Grass carp was separated by all three analyses (two AFLP and SSRs) as shown by the genetic distances and the dendograms (Table 5 and in Fig. 2). AFLP analysis (with *Eco*RI/*Taq*I) also separates the edible carps from the ornamental ones. This reflects the ability of the latter AFLP method to generate reliable patterns even when based on one individual fish from each population.

Genetic distance estimates, based on Nei's index in this study, are similar to isozyme-based values between Dor-70 and Koi (Kohlmann and Kersten 1999).

Positive and significant correlation coefficients of the SSR matrix with the matrices derived by the two AFLP methods suggest that the genetic structure estimated by the three methods is consistent. This is true when a single individual represented each population, as in SSR and *Eco*RI/*Taq*I AFLP analyses, or five fish, as in the *Eco*RI/*Mse*I analysis. The higher correlation coefficient between SSR and *Eco*RI/*Taq*I methods is probably due to the fact that the same fish samples were used for both analyses.

Mean genetic distances calculated for all populations (Table 4) show a high negative correlation with the level of polymorphism as estimated by the proportion of polymorphic bands within each population. This reflects the situation where the less polymorphic populations present different AFLP patterns or gene pools. Mean genetic distances shown in Table 5 could represent only the difference in the gene pool between grass carp and the rest of the carps.

In summary, the current study describes the development and application of 47 new SSR markers and of AFLP to the common carp. It demonstrates the potential of both SSRs and AFLP for future biodiversity and other genetic studies in this species.

Acknowledgments We would like to thank Dr. S. Rothbard and the staff of the Gan-Shmuel Fish Breeding Center for the provision of fish populations for this study.

References

Agresti JJ, Seki S, Cnaani A, Poompuang S, Hallerman EM, Umiel N, Hulata G, Gall GAE, May B (2000) Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. Aquaculture 185:43–56

Ajmone-Marsan P, Valentini A, Cassandro M, Vecchiotti-Antaldi G, Bertoni G, Kuiper M (1997) AFLP markers for DNA fingerprinting in cattle. Anim Genet 28:418–426

Aliah RS, Takagi M, Dong S, Teoh CT, Taniguchi N (1999) Isolation and inheritance of microsatellite markers in the common carp *Cyprinus carpio*. Fisheries Sci 65:235–239

Besse P, Taylor G, Carroll B, Berding N, Burner D, McIntyre CL (1998) Assessing genetic diversity in a sugarcane germplasm collection using an automated AFLP analysis. Genetica 104:143–153

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR Cavalli-Sforza, LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Crooijmans RPMA, Bierbooms VAF, Komen J, Van Der Poel JJ, Groenen MAM (1997) Microsatellite markers in common carp (*Cyprinus carpio* L.). Anim Genet 28:129–134

Edwards YJK, Elgar G, Clark MS, Bishop MJ (1998) The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: perspectives in functional and comparative genomic analyses. J Mol Biol 278:843–854

Knapik EW, Goodman A, Ekker M, Chevrette M, Delgado J, Neuhauss S, Shimoda N, Driever W, Fishman MC, Jacob HJ (1998) A microsatellite genetic linkage map for zebrafish. Nat Genet 18:338–343

Knorr C, Cheng HH, Dodgson JB (1999) Application of AFLP markers to genome mapping in poultry. Anim Genet 30:28–35

Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B (1998) A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics 148:1225–1232

Kohlmann K, Kersten P (1999) Genetic variability of German and foreign common carp (*Cyprinus carpio* L.) populations. Aquaculture 173:433–443

Liu Z, Nichols A, Li P, Dunham RA (1998) Inheritance and usefulness of AFLP markers in channel catfish (*Ictalurus punctatus*), blue catfish (*I. furcatus*), and their F1, F2, and backcross hybrids. Mol Gen Genet 258:260–268

Liu Z, Tan G, Kucuktas H, Li P, Karsi A, Yant DR, Dunham RA (1999) High levels of conservation at microsatellite loci among Ictalurid catfishes. J Hered 90:307–311

MacHugh DE, Loftus RT, Cunningham P, Bradley DG (1998) Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers. Anim Genet 29:333–340

Moav R, Hulata G, Wohlfarth GW (1975) Genetic differences between the Chinese and European races of the common carp. I. Analysis of genotype-environment interactions for growth rate. Heredity 34:323–340

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Reynolds J, Weir BS, Cockerham CC (1983) Estimation for the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779

Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. Nature 354:63–65

Sletan A, Olsaker I, Lie O (1997) Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. Heredity 78:620–627

Spruell P, Pilgrim KL, Greene BA, Habicht C, Knudsen KL, Lindner KR, Olsen JB, Sage GK, Seeb JE, Allendorf FW (1999) Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon. J Hered 90:289–296

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH (1998) A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148:839–850