A linkage map of common carp (Cyprinus carpio) based on AFLP and microsatellite markers

L. Cheng*†, L. Liu*†, X. Yu*, D. Wang*† and J. Tong*
*State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, the Chinese Academy of Sciences, Wuhan 430072, P.R.China. †Graduate School of the Chinese Academy of Sciences, Beijing 100039, P.R.China

Summary
Common carp (Cyprinus carpio) is an important fish for aquaculture, but genomics of this species is still in its infancy. In this study, a linkage map of common carp based on Amplified Fragment Length Polymorphism (AFLP) and microsatellite (SSR) markers has been generated using gynogenetic haploids. Of 926 markers genotyped, 151 (149 AFLPs, two SSRs) were distorted and eliminated from the linkage analyses. A total of 699 AFLP and 20 microsatellite (SSR) markers were assigned to the map, which comprised 64 linkage groups and covered 5506.9 cM Kosambi, with an average interval distance of 7.66 cM Kosambi. The normality tests on interval map distances showed a non-normal marker distribution. Visual inspection of the map distance distribution histogram showed a cluster of interval map distances on the left side of the chart, which suggested the occurrence of AFLP marker clusters. On the other hand, the lack of an obvious cluster on the right side showed that there were a few big gaps which need more markers to bridge. The correlation analysis showed a highly significant relatedness between the length of linkage group and the number of markers, indicating that the AFLP markers in this map were randomly distributed among different linkage groups. This study is helpful for research into the common carp genome and for further studies of genetics and marker-assisted breeding in this species.

Keywords AFLP, common carp (Cyprinus carpio), haploids, linkage map, microsatellite (SSR).

Introduction
Common carp (Cyprinus carpio) is an important freshwater species for aquaculture, with an annual production currently near 3.2 million tons (FAO 2006). It was domesticated about 4000 years ago and is widely distributed from Asia to Europe. In addition, numerous strains and breeds including ornamental colour varieties known as ‘Koi’ carp have been bred from their wild ancestor (Balon 1995). Most economic traits of farmed fishes are complex traits that are governed by multiple genes and can only be located using linked polymorphic marker loci ( Tanksley 1993). Thus, a genetic map where the relative chromosomal locations of different genetic markers are ordered is necessary for genetic analysis of complex traits (Ferguson & Danzman 1998). A preliminary genetic map of this taxon (genus: Cyprinus) was reported (Sun & Liang 2004) and a locus associated with the cold tolerance trait of common carp was mapped on linkage group 5. However, compared with other animals and some model fishes, common carp genomics is still in its infancy. For the studies of genetics and genomics towards initiating selective breeding programmes in common carp, more genetic maps with different molecular markers and higher density are needed.

Microsatellites (also known as simple sequence repeats, SSRs) are inherited in a Mendelian fashion as co-dominant markers. This is a strength of microsatellite markers in addition to their abundance, even genomic distribution, high polymorphism and ease of typing via PCR. Microsatellites have become one of the preferred marker systems for animal genetic mapping. However, the use of microsatellite markers involves a large amount of up-front investment and effort (Chistiakov et al. 2006). In contrast with microsatellites, Amplified Fragment Length Polymorphism (AFLP) is an approach that offers rapid marker development and genotyping. The power of AFLP analysis is tremendously high for revealing genomic polymorphisms and it does not require any prior molecular information. Like Random Amplification of Polymorphic DNA, the major disadvantage of AFLP...
markers is their dominant inheritance (Bensch & Akesson 2005).

Although the meiosis and reproduction of the common carp are the same as for diploid fish, evidence in the form of the DNA content, chromosome counts (Yu et al. 1989) and gene duplication (Ferris & Whitt 1977; Larhammar & Risinger 1994; Zhang et al. 1995) in this species has supported the hypothesis that the common carp is a tetraploid species and that polyploidy may have occurred by hybridization (allopolyploidy) (David et al. 2003). Construction of linkage maps in polyploid organisms has lagged behind that in diploid organisms, because the interpretation of genotypes is more difficult in polyploids. Wu et al. (1992) proposed a general strategy for mapping in polyploid species based on the segregation of single-dose markers (SDM) with 1:1 ratio (presence: absence) in the gametes of heterozygous parents. A single-dose marker in P1 or in the progeny of P1 × P2 can be identified and analyzed as pseudo-test crosses according to two criteria: (i) the fragment should be present in P1 yet absent in P2, and (ii) the fragment should segregate at a 1:1 ratio (presence: absence, determined by a chi-square test) in the progeny of P1 × P2. These criteria are valid regardless of whether the species is allopolyploid, autopolyploid or diploid. Until recently, genetic maps based on SDM were obtained in polyploid species such as zebrafish (Danio rerio) (Postlethwait et al. 1994), tilapia (Oreochromis niloticus) (Kocher et al. 1998), walking catfish (Clarias macrocephalus) (Poompuang & Na-Nakorn 2004) and turbot (Scophthalmus maximus) (Fortes et al. 2008). AFLP markers have not been used for the genome mapping of common carp. In this study, a linkage map of common carp based on AFLP and microsatellite markers was generated by using gynogenetic haploids, with the aim of elucidating its genome for further studies of genetics and marker-assisted breeding in the species.

Materials and methods

Haploid mapping panel

Gynogenetic haploids were prepared following the methods of an earlier study (Cherfas 1981). Milt was collected from a male goldfish (Carassius auratus) and was diluted to about 1:3 in Hank’s solution. Diluted milt was placed into Petri dishes and irradiated by using a 15 W UV germicidal tube from a distance of 16–17 cm for 10–15 min. Approximately 3000 eggs were obtained from a female common carp, and activated (‘fertilized’) with irradiated sperm of goldfish. Then they were divided into batches of 100–150 eggs in each Petri dish and incubated at about 20 °C in water until hatching. Because haploid embryos of common carp die shortly after hatching, haploid populations for mapping in this study were collected 50–55 h after fertilization.

Genomic DNA extraction

Fin clips of the female parent and the haploid embryos were soaked in 95% ethanol, and then stored at 4 °C until DNA extraction. Genomic DNA was purified using a high salt method (Aljanabi & Martinez 1997) with slight modifications. Briefly, tissues or individual embryos were homogenized in 400 μl of sterile TEN buffer (0.4 M NaCl, 10 mM Tris–HCl pH 8.0, and 2 mM EDTA pH 8.0). Then 40 μl of 20% SDS and 8 μl of 20 mg/ml proteinase K were added and mixed well. The mixture was incubated at 55 °C for about 3 h, and then 300 μl of 6 mol/l NaCl (NaCl saturated H2O) was added to each tube. After addition of each reagent, the samples were mixed thoroughly by inversion and then spun down for 30 min at 10 000 g. The supernatant was transferred to a fresh tube, and an equal volume of isopropanol was added. DNA was washed with 70% ethanol, air-dried and dissolved in 1 × TE.

AFLP assay

The AFLP assay was carried out according to the protocol as described by David et al. (2001) with the following modifications: (i) 30 ng of EcoRI selective primer, rather than 5 ng of 32P end-labelled EcoRI selective primer was used in the selective amplification. (ii) PCR products were size fractionated on 6% polyacrylamide gels (38 × 40 cm) on a Sequi-Gen GT Sequencing Cell (Bio-Rad Inc.) and visualized by silver staining. All scored AFLP bands were derived from the EcoRI/TaqI endonuclease system, and the primer combinations were abbreviated in a matrix manner (Table 1). Sequences of adapters and primers used in AFLP assay are provided in Table S1.

Microsatellite genotyping

PCR amplifications for microsatellite markers were carried out in a 10-μl volume. Primer sequences and PCR conditions are detailed as follows: (i) for markers with the prefix MFW, see Crooijmans et al. (1997); (ii) for markers with the prefix HLJ, see Wei et al. (2001); (iii) for markers with the prefix Koi, see David et al. (2001); and (iv) for markers with the prefix CCE, see Wang et al. (2007). In the pilot studies,
the female parent of gynogenesis and its seven haploid progenies were used to screen heterozygous loci that segregated in the progenies. PCR products were size fractionated on 8% native polyacrylamide gels and visualized by Ethidium Bromide staining. Primer sequences, annealing temperatures and expected sizes of the SSR loci are provided in Table S2.

Marker scoring and nomenclature
Alleles of the heterozygous loci in the female parent segregating in 82 haploid embryos were scored. The microsatellite markers were designated by the notations of primers, and the AFLP markers were named according to primer combination and the size order of the AFLP bands. Specifically, the first three letters of an AFLP marker represent the selective bases of EcoRI primer, and the following three letters behind the slash denote the selective bases of TaqI primer. The number following the letters indicates the size order (from large to small) of the AFLP bands scored. For instance, an AFLP marker AGC/ATG-7 indicated the 7th AFLP band which was scored from the primer combination of E-AGC/T-ATG.

Map construction
Chi-square tests were performed to determine whether the segregation ratios of the markers were significantly different from the expected ratio of 1:1. Markers deviating from the expected ratio at the level of significance $P < 0.05$ were eliminated for the linkage analyses. The dataset was designated as ‘backcross’ for the purpose of computer analysis. As a result of the lack of phase information in pseudo-test cross populations, each locus was paired with a ‘dummy locus’ as described by Lodhi et al. (1995). For example, the genotypes of the first 17 progenies at locus ACA/ATG-14 were coded as follows:

ACA/ATG-14 AAAHHAAHHHAAHHAA
ACA/ATG-14d HHAHHAAHHAAAAHAAH

Here ‘H’ represents presence and ‘A’ represents absence of a specific AFLP band or microsatellite allele. A dummy locus was then re-coded in mirror image. This arrangement of the data results in two mirrored linkage maps. Only one map was selected randomly for presentation and further analysis.

The software Map Manager QTX (Manly et al. 2001) was employed to carry out linkage analyses. Linkage groups were formed with an initial $P$-value of 1E-5 by using the ‘Make Linkage Groups’ command. $P$-values in Map Manager indicated the probability of a Type I error, which is the probability of a false positive linkage. Once the initial linkage groups were established, those previously unlinked markers were allocated to these linkage groups again using the ‘Distribute’ command with the $P$-value of 1E-4. The map distances between markers were calculated by the Map Manager in centiMorgans (cM) using the Kosambi function. Charts of linkage groups were drawn using an EXCEL (Microsoft Inc.) macro ‘MAPDRAW’ (Liu & Meng 2003).

Map statistics
The Shapiro–Wilk test was used to determine whether distribution of standardized map distances between adjacent markers differed from normal distribution. The Pearson’s correlation coefficient between the number of markers per linkage group and the total length of the linkage group was calculated to evaluate the random distribution of markers throughout the maps. Histogram drawing, normality test and correlation analysis were performed using the software SPSS 13.0 (SPSS Inc.).

Table 1

<table>
<thead>
<tr>
<th>Primer combination</th>
<th>T-AAT</th>
<th>T-ACA</th>
<th>T-ACC</th>
<th>T-ACG</th>
<th>T-ACT</th>
<th>T-AGA</th>
<th>T-AGT</th>
<th>T-ATC</th>
<th>T-ATG</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-AAA</td>
<td>10/1</td>
<td>10/0</td>
<td>3/0</td>
<td>21/4</td>
<td>17/5</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AAC</td>
<td>34/3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AAG</td>
<td>20/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AAT</td>
<td>NS</td>
<td>NS</td>
<td>17/4</td>
<td>16/5</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-ACA</td>
<td>7/0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-ACC</td>
<td>NS</td>
<td>9/2</td>
<td>8/0</td>
<td>11/1</td>
<td>11/2</td>
<td>7/0</td>
<td>9/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AGC</td>
<td>15/7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-ACT</td>
<td>13/3</td>
<td>22/3</td>
<td>15/1</td>
<td>14/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AGA</td>
<td>21/2</td>
<td>19/7</td>
<td>NS</td>
<td>15/4</td>
<td>25/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AGC</td>
<td>21/2</td>
<td>19/3</td>
<td>11/1</td>
<td>13/2</td>
<td>12/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AGG</td>
<td>18/6</td>
<td></td>
<td>13/3</td>
<td>24/6</td>
<td>5/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-AGT</td>
<td>17/1</td>
<td>6/0</td>
<td>16/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-ATA</td>
<td>5/1</td>
<td>25/4</td>
<td>NS</td>
<td>22/3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-ATG</td>
<td>NS</td>
<td></td>
<td></td>
<td>10/3</td>
<td>9/2</td>
<td>11/1</td>
<td>8/4</td>
<td>7/2</td>
<td></td>
</tr>
</tbody>
</table>

NS, not scored.
Results

Marker genotyping

In the pilot studies, the female parent of gynogenesis and its seven haploid progenies were screened for a total of 61 microsatellite markers, and 24 (39.3%) loci were heterozygous in the parent and segregated in the progenies. An additional 75 haploid progenies were scored for those 24 polymorphic SSR loci. In total, 73 AFLP primer combinations were assayed in the haploid mapping panel, and 63 (~86.3%) of the primer combinations produced clear and easy-to-score bands. A total of 902 segregated AFLP loci were identified from these primer combinations, and the number of markers per combination ranged from 2 to 34 with a mean of 14.3 (Table 1).

Segregation distortion

Segregation distortion was observed in Chi-square tests. At the level of significance $P < 0.01$, 78 AFLP and 1 SSR (CCE48) markers showed segregation distortion. When the level of significance was raised to the level of $P < 0.05$, an additional 71 AFLP and 1 SSR (Koi45) markers were distorted. When accurate counts were made in correspondence with the primer combinations, the results showed that the majority of the distorted markers occurred from some specific primer combinations (Table 1), and these distorted markers were eliminated for the linkage analyses. Therefore, the final dataset consisted of the results of genotyping for 22 SSR and 753 AFLP markers in a panel of a female parent and its 82 haploids (Table 2).

Map construction

Initially, a total of 714 markers (696 AFLPs, 18 SSRs) were assigned to 67 linkage groups by using the ‘make linkage group’ command at a $P$-value of 1E-5. Three small linkage groups, which contained only two closely linked AFLP markers, were excluded from the original map. Thus the original map of common carp contained 708 markers spanning 64 linkage groups. After the ‘Distribute’ command was applied at a $P$-value of 1E-4, 11 accessory markers (two SSRs, nine AFLPs; indicated with ‘+’ in Fig. 1) from previously unlinked markers, which included six markers in those eliminated small linkage groups, were also allocated to the original map. Consequently, 719 (including 20 SSRs) of the total 775 usable markers were assigned on the final map (Table 2).

The number of markers per linkage group ranged from 2 to 33, with a mean of 11.2. Fifty-three major linkage groups were established with 5–33 markers, and 11 smaller linkage groups with 2–4 markers (Fig. 1). The final linkage map had 64 linkage groups covering 5506.9 cM of the common carp genome, with the average length of linkage group being 86.0 cM. The largest linkage group spanned 205.9 cM with 20 markers, and the shortest one only 2.4 cM with three markers (details are provided in Table S3).

Map characteristics

Visual inspection of the distance distribution histogram (Fig. 2) showed a cluster of interval map distances on the left side of the chart. The results of normality tests on interval map distances for the dataset of this study appear to show a non-normal distribution. In addition, the Pearson’s correlation coefficient between the length of linkage groups and the number of markers per linkage group was 0.87, and correlation was significant at the 0.01 level (two-tailed).

Discussions

The karyotype of common carp consists of 50 pairs of chromosomes (Yu et al. 1989), but our linkage map comprised 64 linkage groups. The presence of extra linkage groups could be explained by the lack of intermediate markers between linkage groups belonging to the same chromosome. This phenomenon could also account for the non-linked markers observed in our study. In a previous study, 50 linkage groups were identified, which were equal to the number of chromosomes in the haploid genome of common carp (Sun & Liang 2004). However, because both studies have not assigned linkage groups to specific chromosomes, the gaps and/or undetected chromosomes probably still exist in both cases. In addition, both studies showed that the genome size of common carp is quite large. Except for higher coverage and resolution, a notable difference between this linkage map and map constructed by Sun & Liang (2004) is that the mapping panel was generated from a female common carp rather than an interspecific hybrid female produced by crossing a female common carp (C. carpio) and a male Boshi carp (C. pellegrini). The hybrid $F_1$ female was more likely to be heterozygous, which means the mapping population will be more informative. However, in that case, each $F_2$ haploid offspring only represents one single meiosis in the hybrid $F_1$ dam rather than common carp (C. carpio) or Boshi carp (C. pellegrini). The interspecific hybrid $F_1$ generation is neither a valid species nor a cultured

<table>
<thead>
<tr>
<th>Class</th>
<th>Total</th>
<th>AFLP</th>
<th>SSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markers genotyped</td>
<td>926</td>
<td>902</td>
<td>24</td>
</tr>
<tr>
<td>Markers distorted</td>
<td>151</td>
<td>149</td>
<td>2</td>
</tr>
<tr>
<td>Markers finally used</td>
<td>775</td>
<td>753</td>
<td>22</td>
</tr>
<tr>
<td>Markers grouped ($P &gt; 1E-5$)</td>
<td>708</td>
<td>690</td>
<td>18</td>
</tr>
<tr>
<td>Markers assigned ($P &gt; 1E-4$)</td>
<td>11</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Markers mapped</td>
<td>719</td>
<td>699</td>
<td>20</td>
</tr>
<tr>
<td>Markers unlinked</td>
<td>56</td>
<td>54</td>
<td>2</td>
</tr>
</tbody>
</table>

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, Animal Genetics, 41, 191–198
A linkage map of common carp using AFLPs and SSRs

Figure 1 Linkage map of *Cyprinus carpio* based on AFLP and microsatellite markers. The microsatellite markers were designated by the notations of primers, and the AFLP markers were named by the information of primer combination and the size order of the AFLP bands. Map distances are in cM estimated by the Kosambi mapping function. Markers with ‘+’ are accessory markers allocated to map at the *P*-value of 1E-4.
fish for marketing purposes, and this kind of hybrid panel is normally analysed as ‘two-way pseudo-testcross’ for construction of linkage maps of each parental species separately (Grattapaglia & Sederoff 1994).

Deviations from the expected Mendelian ratios have been observed in previous efforts to construct linkage maps using molecular markers. Segregation distortion was observed for 151 of 926 markers in this study. A similar distortion rate of ~16% was reported for catfish using AFLP markers (Liu et al. 2003). Except for chance deviation, amplification of a single sized fragment derived from several genomic regions may account for the segregation distortion (Faris et al. 1998). In addition, the presence or absence of specific genes or haplotypes that are not compatible with the survival and/or fertilization of eggs may influence the ratio (Li et al. 1994).

Clusters of AFLP markers were also observed in many previous studies of several fish species, including medaka (Oryzias latipes) (Naruse et al. 2000), tilapia (O. niloticus) (Kocher et al. 1998; Agresti et al. 2000), rainbow trout (Oncorhynchus mykiss) (Young et al. 1998; Nichols et al. 2003) and channel catfish (Ictalurus punctatus) (Liu et al. 2003). In parallel with the results of Alonso-Blanco et al. (1998), we also found that a small proportion of the clustering markers were the results of allelism between some AFLP bands. For example, ACT/ACA14 and ACT/ACA15, which were completely linked markers with different lengths in opposite phases, suggested that we have identified codominant alleles of a length variant. However, a large proportion of the marker clusters could not be explained by allelism. Other potential causes, as suggested by Liu et al. (2003), may include: (i) a reduced recombination rate around centromeric regions and/or telomeric regions; (ii) a bona fide enrichment of AFLP markers in these regions as a result of an uneven distribution of restriction sites; (iii) presence of highly repetitive elements within these genomic regions with great variation in both the lengths and the sequences among the repetitive elements. Nonetheless, such high levels of marker clustering hinder the effectiveness of AFLP markers, and therefore it is warranted to elucidate the nature of the genomic sequences surrounding the regions of clustering markers.

In a saturated genetic map, the frequency of genetic distances with markers randomly located on the chromosomes should show a normal distribution (Krutovskii et al. 1998). In this study, a shift of the interval map distances on the left side of the distance distribution chart suggested the occurrence of AFLP marker clusters. On the other hand, the lack of an obvious cluster on the right side showed that there were a few big gaps which need more markers to bridge. When markers are randomly distributed throughout the genome, the correlation between the number of markers per linkage group and the length of the linkage groups must be linear, because the bigger the chromosome, the higher the number of markers that can be grouped (Cervera et al. 2001). As the correlation analyses in our study are highly significant, we conclude that the AFLP markers in this map of common carp are randomly distributed among different linkage groups.

Construction of a framework genetic map for future saturation and localization of genes or valuable trait loci is the major aim of many genomic studies. The map obtained here would be a useful tool for the study of the genomics of the common carp. Our map is based on the segregation of DNA markers using gynogenetic haploids representing meiosis in the female parent, and so is a female linkage map of common carp. It has been generally proposed that the difference in recombination rate between sexes is found in most animals, including fish species (e.g Agresti et al. 2000; Sakamoto et al. 2000). It is, therefore, necessary to develop a male linkage map for common carp using conventional mapping strategies. For the mapping of qualitative or quantitative trait loci in the genome of the common carp, diploid panels whose phenotypic traits can be measured should also be used.

Acknowledgements

We thank Dr. Y. Li for assistance in data analyses and Dr. C. You and Mr. W. Li for laboratory work. This study was partially supported by ’973’ (2004CB117405, 2010CB126305), ’948’ (2006-G55) and NSF (30771643) project.

References

A linkage map of common carp using AFLPs and SSRs


Sakamoto T., Danzmann R.G., Gharbi K. et al. (2000) A microsatellite linkage map of rainbow trout (Oncorhynchus mykiss)


Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Sequences of AFLP adapters and primers used in Cyprinus carpio mapping.

Table S2 Microsatellite markers segregating in the gynogenetic haploid panel of Cyprinus carpio (Ta, annealing temperature; Es, expected size).

Table S3 Tabular linkage map of Cyprinus carpio based on AFLP and microsatellite markers. The numbers on the left of markers are genetic distance in Kosambi centiMorgans (cM) between markers. The numbers on the right of markers are cumulative genetic distance from top end in Kosambi centiMorgans (cM). Markers with ‘+’ are accessory markers allocated to map at the P-value of 1E-4.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting (other than missing files) should be addressed to the authors.