Recent Duplication of the Common Carp (*Cyprinus carpio* L.) Genome as Revealed by Analyses of Microsatellite Loci

Lior David,*† Shula Blum,† Marcus W. Feldman,* Uri Lavi,‡ and Jossi Hillel†

*Department of Biological Sciences, Stanford University; †Department of Field Crops and Genetics, Faculty of Agriculture, Food and Environmental Quality Sciences, The Hebrew University of Jerusalem, Rehovot, Israel; and ‡Institute of Horticulture, ARO-Volcani Center, Bet-Dagan, Israel

Genome duplications may have played a role in the early stages of vertebrate evolution, near the time of divergence of the lamprey lineage. Additional genome duplication, specifically in ray-finned fish, may have occurred before the divergence of the teleosts. The common carp (*Cyprinus carpio*) has been considered tetraploid because of its chromosome number (2n = 100) and its high DNA content. We studied variation using 59 microsatellite primer pairs to better understand the ploidy level of the common carp. Based on the number of PCR amplicons per individual, about 60% of these primer pairs are estimated to amplify duplicates. Segregation patterns in families suggested a partially duplicated genome structure and disomic inheritance. This could suggest that the common carp is tetraploid and that polyploidy occurred by hybridization (allotetraploidy). From sequences of microsatellite flanking regions, we estimated the difference per base between pairs of alleles and between pairs of paralogs. The distribution of differences between paralogs had two distinct modes suggesting one whole-genome duplication and a more recent wave of segmental duplications. The genome duplication was estimated to have occurred about 12 MYA, with the segmental duplications occurring between 2.3 and 6.8 MYA. At 12 MYA, this would be one of the most recent genome duplications among vertebrates. Phylogenetic analysis of several cyprinid species suggests an evolutionary model for this tetraploidization, with a role for polyploidization in speciation and diversification.

Introduction

Genome duplications may have occurred in the early stages of vertebrate evolution, enabling organisms to evolve through modification of duplicated genes and acquisition of new functions (Ohno 1970). Whole-genome duplication may explain the variation in chromosome numbers as well as the multiple gene copies and chromosome segments in species of vertebrates (Postlethwait et al. 1998; Wolfe 2001). One or two rounds of genome duplication in vertebrate evolution have been suggested to occur before the divergence of the lamprey lineage and after this divergence, about 450 MYA (Holland et al. 1994; Sidow 1996; Skrabanek and Wolfe 1998). Additional genome duplication, specific to ray-finned fish, possibly occurred about 360 MYA, preceding the divergence of the teleosts. This duplication could have enabled the major diversification of the teleosts, the most species-rich group of vertebrates (Amores et al. 1998; Meyer and Málaga-Trillo 1999; Taylor et al. 2001). In general, polyploidy is much more prevalent in plants than among animals, where it is found mostly in insects, amphibians, and fish. Creative roles in evolution such as speciation, adaptation, diversification, and promotion of new functions have been attributed to polyploidy (Otto and Whitton 2000).

Assuming a whole-genome duplication to explain the genome structure of intensively studied species such as yeast, *Arabidopsis thaliana*, and *Homo sapiens* raised several difficulties, mostly because of the time elapsed since the duplication event (Wolfe 2001). The low proportion of duplicated segments in these species and the shuffling in the structure of their ancestral chromosomes (Postlethwait et al. 2000; Friedman and Hughes 2001) leave the theory of genome duplication in debate. Although a substantial amount of evidence suggests a whole-genome duplication (e.g., Gu, Wang, and Gu 2002; McLysaght, Hokamp, and Wolfe 2002), many segmental duplications were also found (Bailey et al. 2002). The partially duplicated structure of the yeast and human genomes raised the possibility that this structure could also be a result of multiple independent segmental duplications (Llorente et al. 2000; Hughes, da Silva, and Friedman 2001).

A few fish species are supposed to have had an additional recent round of genome duplication late in the evolution of the teleosts that might have led to their speciation. Among these are the catastomid fishes (suckers), with an estimated duplication time of 50 MYA (Uyeno and Smith 1972); the salmonids, with an estimated time of duplication of 25 to 100 MYA (Allendorf and Thorgaard 1984); and the common carp and the goldfish (Ohno et al. 1967; Larhammar and Risinger 1994). Polyploidization has also been documented in some loaches (*cobitidea*) (Ferris and Whitt 1977*b*) and in sturgeons (Ludwig et al. 2001). The rarity of polyploidy in higher vertebrates is probably a result of genetic sex determination, and few fish serve as unique examples of polyploids with genetic sex determination (Otto and Whitton 2000).

Genome duplication in the evolution of the common carp (*Cyprinus carpio*) is supported by the following observations. Its chromosome number (n = 50) is twice that of other *Cyprinidae*, and its DNA content is higher (Ohno et al. 1967). In addition, about 52% of this carp enzymes show a pattern consistent with duplication (Ferris and Whitt 1977*a*). Tetraploidization of carp was suggested to take place about 50 MYA, similarly to catastomids since both express a similar proportion of enzymes in duplicates. The c-*myc* genes in carp gave an estimate of 58 MYA for the event of tetraploidization (Zhang, Okamoto, and Ikeda 1995). Other duplicated genes of the carp suggest a more

recent divergence time of less than 16 MYA (Larhammar and Risinger 1994).

Carps of the family *Cyprininae* are the most cultivated species in aquaculture. Common carp is the third most cultivated species worldwide and is important in the European freshwater aquaculture, where its production has increased substantially over the last decade (FAO: http://www.fao.org/fi/default.asp). Though agriculturally important, genomic information on the carp is limited, and only 1,227 nucleotide and 606 protein sequences are currently available in GenBank (http://www.ncbi.nlm.nih. gov/). The common carp lacks a genetic linkage map, and about 100 microsatellite markers have been developed for this species (Crooijmans et al. 1997; Aliah et al. 1999; David et al. 2001).

Microsatellite markers provide a codominantly inherited tool for studying mostly noncoding regions of the genome. In a few studies, microsatellites have been useful for the study of duplications (Ludwig et al. 2001; Pyatskowit et al. 2001; Angers, Gharbi, and Estoup 2002). In this paper we use microsatellite markers to investigate the existence and extent of duplications in the common carp genome.

Materials and Methods
Fish Families

Several families were generated. Family I is a cross between a transparent koi female and a Yugoslavian carp male (Moav, Hulata, and Wohlfarth 1975). Family II is a cross between a yellow (Ohgon) koi female and a Yugoslavian carp male. From family I and family II, 59 and 54 offspring were genotyped, respectively. Four F_2 families were produced by crossing fish with known genotypes from family II. Each F_2 family contained about 50 offspring. Families were produced and maintained at the Gan-Shmuel Fish Breeding Center in Israel.

Microsatellite Loci

Genotyping at microsatellite loci was done as in David et al. (2001). When fluorescent primers were used (family II), PCR products were separated on a 6% polyacrylamide gel (Bio Lab Ltd. Jerusalem, Israel) on an ABI Prism™ 377 DNA Sequencer (Perkin Elmer, Foster City, Calif.). Otherwise, labeling was done by adding 2.5 µCi of ^{32}P in a total volume of 10 µl reaction. Radioactive products were denatured and separated on 5% polyacrylamide gels in TBE buffer (as in David et al. [2001]). Primer sequences and PCR conditions are detailed as follows: (1) for markers with the prefix MFW, see Crooijmans et al. (1997); (2) for markers with the prefix CCA, see Aliah et al. (1999); and (3) for markers with the prefix Koi, see David et al. (2001). Markers were first tested for polymorphism among parents of each family and for the number of fragments per individual. We used 53 microsatellite markers to genotype the four parents of family I and family II, whereas six additional markers were genotyped in only two parents of either family. A subset of 39 polymorphic markers that resulted in two or more

fragments per individual was then used for genotyping the progeny. Genotypes were scored visually.

Cloning and Sequencing of Alleles

We chose seven primer pairs that represented the various segregation patterns and sequenced 36 different fragments from these loci. For each of these primer pairs, we amplified DNA of individuals that represented the whole set of observed fragments. When fragments of the same size were cloned twice from different families, their sequences were found to be identical. Six sequences of stuttered fragments (that are typical of microsatellites) were excluded, and our further analyses are based on 30 sequences. Each PCR product was cleaned using a purification kit (High Pure PCR product, Roche Applied Sciences GmbH, Mannheim, Germany). Cloning was done using pGEM®-T vector system kit (Promega, Madison, Wis.). About 50 positive colonies were chosen for each PCR reaction and cultured overnight in 100 µl of liquid LB substrate in a 96-well plate. Identification of isolated fragments was done by a radioactive PCR reaction using 0.5 µl from the cultured clone as a template for the corresponding primer pair. Products were separated as detailed above. Clones for sequencing were grown overnight in 7.5 ml of LB medium, and plasmids were isolated using the Wizard® Plus SV Miniprep DNA purification system (Promega, Madison, Wis.). Sequencing was done by ABI 3100 automated sequencer (Perkin Elmer, Foster City, Calif.) using T7 or SP6 universal primers.

Analyses of Sequences

Alignment of sequences was done using the BioEdit software (Hall 1999) and then visually refined. For each sequence, the number of microsatellite repeats was counted and their sequence was deleted. The flanking regions were used to estimate the genetic relationship between fragments using Tamura-Nei distance (Tamura and Nei 1993) in the MEGA2 software (Kumar et al. 2001). Alignment data were bootstrapped, and gene trees were constructed using the Neighbor-Joining method (Saitou and Nei 1987), as implemented in MEGA2. Scales of genetic distances and divergence times for the gene trees were calculated by this software.

For analysis of microsatellites' flanking regions, we used a nucleotide substitution rate of 3.71×10^{-9} per site per year. Since these sequences are noncoding regions, we used the highest rate found for fourfold degenerate sites based on 47 mammalian genes, assuming a divergence time of 80 Myr between human and rodent lineages. For analysis of genes, we used the rate of 3.51×10^{-9} as estimated by Li (1997, p. 90) from synonymous sites in coding regions. Time of divergence was calculated using the formula $T = K/2r$, where K is the number of substitutions per base between homologous sequences and r is the rate of substitution. Estimation of means, distributions, and variances, as well as statistical analyses of substitution rates, divergence times, and repeat numbers were carried out using JMP4 statistical software (SAS Institute, Cary, N.C.).

Table 1
Summary of Segregation Patterns of Microsatellite Loci in Families

Number	Type[a]	Primer Pair	Family I ($n = 59$ fish)	Family II ($n = 54$ fish)	F_2 Families ($n = 50$ fish)
1	1	*Koi53-54*	1 or 2 fragments in parents[b]	Duplicated[c]	Duplicated[d]
2	1	*Koi35-36*	>2 fragments in 59 offspring[d]	Duplicated[c]	Duplicated[c]
3	1	*CCA 17*	Duplicated[c] (differential fixation[e])	Duplicated[c]	No data
4	1	*Koi17-18*	1 or 2 fragments in parents[b]	>2 fragments in parents[b,f]	Duplicated[c]
5	1	*Koi111-112*	>2 fragments 58 offspring[d]	>2 fragments in parents[b,f]	Duplicated
6	1	*Koi89-90*	Duplicated[c]	Duplicated[c]	Duplicated[c]
7	1	*Koi49-50*	Duplicated[c]	>2 fragments in parents[b,f]	Diploid
8	1	*Koi3-4*	>2 fragments in 5 offspring[d]	Duplicated[c]	Diploid
9	1	*Koi115-116*	Duplicated[c]	>2 fragments in parents[b,f]	No data
10	1	*Koi19-20*	Duplicated[d]	>2 fragments in parents[b,f]	No data
11	1	*MFW 14*	No data	Duplicated[c]	No data
12	1	*MFW 7*	>2 fragments in parents[b,f]	Duplicated[c] (differential fixation[e])	No data
13	1	*MFW 4*	>2 fragments in parents[b,f]	Duplicated[c] (differential fixation[e])	No data
14	1	*MFW 22*	No data	Duplicated[c] (differential fixation[e])	No data
15	1	*Koi55-56*	1 or 2 fragments in parents[b]	Duplicated[c] (differential fixation[e])	No data
16	1	*MFW 26*	>2 fragments in parents[b,f]	Duplicated[c] (differential fixation[e])	No data
17	1	*Koi79-80*	>2 fragments in parents[b,f]	Duplicated[c] (differential fixation[e])	No data
18	1	*Koi93-94*	1 or 2 fragments in parents[b]	Duplicated[c]	Diploid
19	2	*Koi5-6*	1 or 2 fragments in parents[b]	>2 fragments in parents[b,f]	1 or 2 fragments/offspring[d]
20	2	*Koi107-108*	>2 fragments in 5 offspring[d]	>2 fragments in parents[b,f]	No data
21	2	*Koi43-44*	>2 fragments in 33 offspring[d]	No data	No data
22	2	*Koi105-106*	>2 fragments in 59 offspring[d]	No data	No data
23	2	*Koi81-82*	>2 fragments in parents[f]	>2 fragments in 35 offspring[d]	No data
24	2	*Koi29-30*	>2 fragments in 20 offspring[d]	1 or 2 fragments in parents[b]	Diploid
25	2	*MFW 31*	No data	>2 fragments in 30 offspring[d]	No data
26	3	*MFW 23*	>2 fragments in parents[b,f]	>2 fragments in 16 offspring[d]	Duplication (tandem)
27	4	*MFW 20*	No data	Diploid	No data
28	4	*CCA 30*	No data	Diploid	No data
29	4	*MFW 18*	No data	Diploid	No data
30	4	*Koi67-68*	1 or 2 fragments in parents[b]	Diploid	No data
31	4	*MFW 10*	No data	Diploid	No data
32	4	*Koi85-86*	Diploid	No data	Diploid (with null allele)
33	4	*MFW 2*	No data	Diploid	No data
34	5	*MFW 30*	No data	Diploid (with null allele)[d]	No data
35	5	*MFW 15*	No data	Diploid (with null allele)[d]	No data
36	5	*MFW 29*	No data	Diploid (with null allele)[d]	No data
37	5	*Koi75-76*	1 or 2 fragments in parents[b]	Diploid[d]	No data
38	5	*MFW 11*	No data	Diploid[d]	No data
39	5	*Koi27-28*	1 or 2 fragments in parents[b]	Diploid[d]	No data

[a] Types of segregation pattern: (1) duplicated and disomic inheritance; (2) suggested as duplicated based on fragment number, but disomic inheritance is distorted; (3) tandem duplication; (4) diploid and disomic inheritance; (5) suggested as diploid but disomic inheritance is distorted.

[b] Information from parents is presented where progeny were not genotyped.

[c] Duplicated loci that fitted a disomic model and can be explained either as tandem or separate loci.

[d] The observed proportion of genotypes in progeny is significantly different from the expected ratio based on either disomic or tetrasomic inheritance (χ^2 test, $P < 0.05$).

[e] Duplicated loci each fixed on a different allele.

[f] At least one of the parents has more than two fragments.

Sequences of genes were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) using the Blast tool and taxonomic information was found in the taxonomy browser (http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/).

Results

Screening Parents at Microsatellite Loci

The 59 studied primer pairs were classified into three categories according to the number of PCR fragments found in the four parents of families I and II: (1) six primer pairs (10%) with one fragment per individual, (2) 21 primer pairs (36%) with two fragments in at least one parent, and (3) 31 primer pairs (54%) that exhibited more than two fragments in at least one parent. Of the 59 primer pairs, 53 (90%) were considered as polymorphic because they had two or more fragments for at least one individual, and 60% of those had more than two fragments per individual.

Patterns of Allele Segregation in Progeny

In total, 54 segregation patterns were analyzed among 39 polymorphic primer pairs (table 1). Some primer pairs were genotyped in more than one family. For example, primer pair *Koi89-90* (no. 6 in table 1) was studied in an F_2 family where both parents had three fragments each, and in total four fragments were found. The segregation pattern in progeny revealed that 83% of the offspring had three fragments (table 2 and fig. 1a). All offspring had a 164-bp fragment, whereas the other three fragments segregated (fragments 194 and 184 are present in approximately half of the fish, and fragment 182 is present in 76% of them). The proportions of the genotypes in progeny approach

Table 2
Examples of Duplicated Loci with Disomic Inheritance

Primer Pair	Family	Genotypes in Parents		Fragments in Progeny			Genotypes in Progeny				
		Parent	Locus1 // Locus2	Size (bp)	n/Total	Frequency	Locus1 // Locus2	n	Observed[a]	Expected[b]	χ^2
Koi89-90	F₂; n = 41	Grand dam	164/164 // 184/194[c]	194	20/41	0.49	164/164 // 182/182[c]	7	0.17	0.25	2.41;
		Grand sire	164/164 // 182/194	184	24/41	0.59	164/164 // 182/184	14	0.34	0.25	P = 0.49
		Dam	164/164 // 182/194	182	31/41	0.76	164/164 // 182/194	10	0.24	0.25	
		Sire	164/164 // 182/184	164	41/41	1.00	164/164 // 184/194	10	0.24	0.25	
Koi89-90	I; n = 55	Dam	164/164 // 194/194	194	54/54	1.00	164/164 // 184/194	54	1.00	1.00	0; P = 1
		Sire	164/164 // 184/184	184	54/54	1.00					
				164	54/54	1.00					
Koi53-54	II; n = 53	Dam	158/183 // 183/183	183	53/53	1.00	183/183 // 183/183	11	0.21	0.25	1.11;
		Sire	175/183 // 183/183	175	26/53	0.49	158/183 // 183/183	16	0.30	0.25	P = 0.77
				158	28/53	0.53	175/183 // 183/183	14	0.26	0.25	
							158/175 // 183/183	12	0.23	0.25	
CCA 17	II; n = 53	Dam	150/150 // 181/184	187	53/53	1.00	150/150 // 181/187	26	0.49	0.5	0.02;
		Sire	150/150 // 187/187	184	27/53	0.51	150/150 // 184/187	27	0.51	0.5	P = 0.89
				181	26/53	0.49					
				150	53/53	1.00					
Koi35-36	F₂; n = 47	Grand dam	172/172 // 184/204	204	18/47	0.38	172/172 // 196/196	18	0.38	0.25	4.8;
		Grand sire	172/172 // 196/196	196	39/47	0.83	172/172 // 196/204	10	0.21	0.25	P = 0.18
		Dam	172/172 // 184/196	184	19/47	0.40	172/172 // 184/204	8	0.17	0.25	
		Sire	172/172 // 196/204	172	47/47	1.00	172/172 // 184/196	11	0.23	0.25	
MFW 23	F₂; n = 39	Grand dam	101/126 // 117/126	126	29/39	0.74	126/126 // 126/126	10	0.26	0.25	0.03;
		Grand sire	126/126 // 126/126	117	29/39	0.74	101/101 // 117/117	10	0.26	0.25	P = 0.99
		Dam	101/126 // 117/126	101	29/39	0.74	101/126 // 117/126	19	0.48	0.5	
		Sire	101/126 // 117/126								

[a] Observed fraction of the genotype.

[b] Expected fraction based on parents' genotypes and disomic inheritance.

[c] The genotypes of the two loci are separated by double slash. Alleles of one duplicate are italicized in parents and in progeny.

a 1:1:1:1 ratio. Primer pair *Koi53-54* amplified three fragments in parents of family II (table 2). Fragments 158 and 175 were found in about 50% of the offspring, whereas fragment 183 was detected in all fish. Four different genotypes were found in proportions approaching a 1:1:1:1 ratio; about one-quarter of the offspring had all three parental fragments. From such segregation patterns (i.e., frequencies of alleles and genotypes in progeny) and genotypes of the parents, we inferred which fragments have allelic relationships and which fragments are not alleles but duplicates (paralogs). The expected ratio of genotypes in the progeny was based on genotypes of the parents and assuming either a disomic or tetrasomic mode of inheritance. For example, four fragments were amplified in parents of family II using primer pair *CCA 17* (table 2). The two paternal fragments were found in all offspring, and two of the three maternal fragments segregated in a 1:1

ratio. The segregating fragments were defined as alleles since none of the offspring had both. Fragment 187 was assumed to be in a homozygous state with fragment 150 as its duplicate since all offspring had both. These relationships predict two genotypes in progeny, with an expected ratio of 1:1 under disomic inheritance. The observed proportions of genotypes fits this postulated ratio ($\chi^2 = 0.02$, $P = 0.89$). Fourteen patterns out of 54 (25.9%) deviated significantly from expected ratios (χ^2 test, $P < 0.05$ [table 1]). All primer pairs that fitted a tetrasomic mode of inheritance also fitted a disomic model. On the other hand, some of the patterns that fitted a disomic mode of inheritance could not be explained by tetrasomic inheritance. Loci that were defined as diploid and mainly those whose alleles segregated in a 1:1 ratio could fit a tetrasomic model as well. However, disomic inheritance was used to explain segregation patterns, since it is more conservative and explains more segregation ratios at both single and duplicated loci.

The duplicated nature of loci was inferred based on the number of fragments per individual. Segregation patterns in families supported this assessment in 86% of the cases. The segregation patterns of the 39 primer pairs were categorized into five types (table 1): (1) 18 primer pairs that were found to be duplicated and to fit a disomic mode of inheritance; (2) Seven primer pairs with more than two fragments per individual but with segregation patterns that deviated from any expected ratio; (3) one primer pair that fits a tandem duplication pattern of inheritance; (4) seven primer pairs that fit a diploid mode of inheritance; and (5) six primer pairs that showed no indication of being duplicated but had distorted diploid segregation.
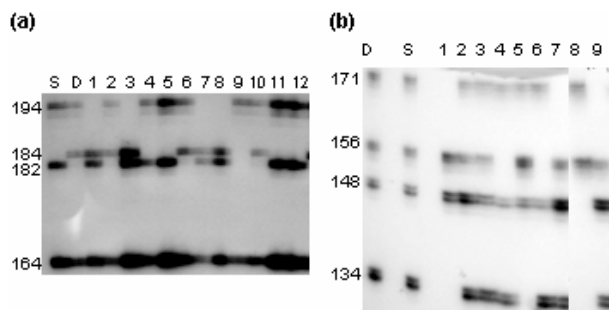

Fɪɢ. 1.—Segregation patterns of duplicated loci. Dam (D), sire (S), and progeny (1 to 12). (*a*) Primer pair *Koi89-90*; (*b*) Primer pair *Koi111-112*. Numbers on the left side are size of fragments (bp).
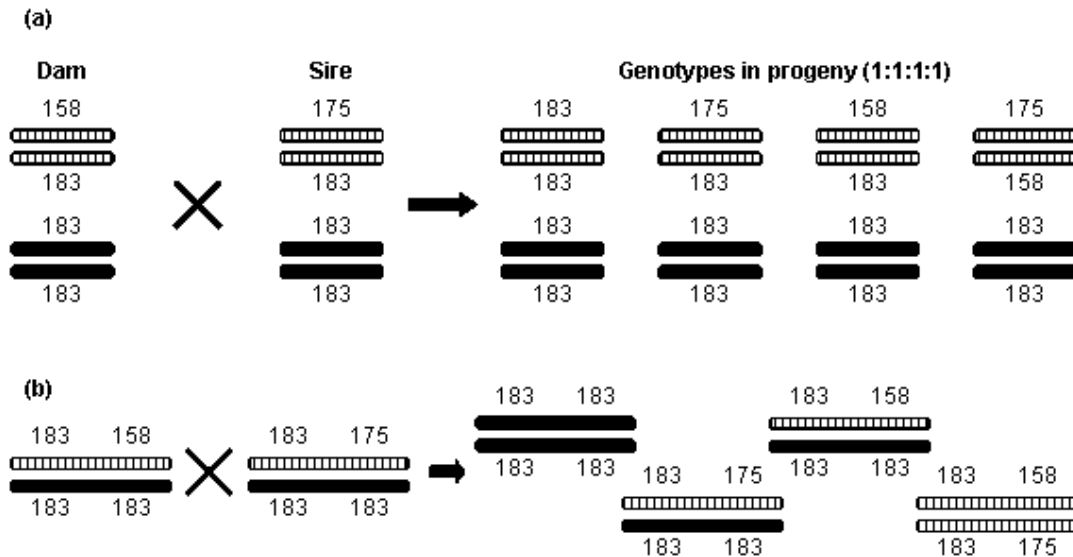
Fig. 2.—Options of duplication with disomic segregation at locus *Koi53-54*. (*a*) Duplicates are unlinked and labeled in different colors. (*b*) Duplicates are in tandem, and homologs are labeled in different colors. Note that both configurations result in the same four genotypes with a 1:1:1:1 ratio in progeny.

For genotyping of progeny, we chose primer pairs where segregation was expected based on the genotypes of the parents. Primer pair *Koi111-112* had four fragments in both parents, and all these alleles segregated in the progeny (fig. 1*b*). Such loci were defined as duplicated, based on both fragment number and segregation pattern. However, some primer pairs that had no more than two fragments per parent gave rise to a segregation pattern that revealed a duplicate locus (e.g., primer pair *Koi53-54* in family II [table 2 and fig. 2]). Furthermore, primer pairs where both parents appeared to be heterozygous were found to amplify paralogous loci; thus, all progeny had all the fragments from both parents (e.g., primer pair *Koi89-90* in family I [table 2]). Such primer pairs, where the fragment number indicated a diploid locus while the segregation revealed paralogs, demonstrate the complex relationship between duplicates and point towards the importance of the segregation analysis.

Differential fixation defines a situation in which the two paralogs are homozygous for different alleles. This state was found in seven of the 25 (28%) segregation patterns that did not deviate from their expected ratios and in which duplicates were suggested (table 1, in types 1 to 3). For example, at least one parent of family I at primer pair *Koi89-90,* as well as the grand sire of the F_2 families at primer pair *Koi35-36,* are apparent heterozygotes but actually possess a genotype with differential fixation (table 2).

Of the 21 segregation patterns that fitted a disomic and duplicated model (type 1, [table 1]) only primer pair *Koi111-112* in the F_2 family supported two separate duplicates. All 20 of the other patterns under disomic inheritance can be equally explained as either tandem duplication or unlinked duplicates. All of these loci are either differentially fixed or have only one duplicate segregating. This monomorphism, in at least one duplicate, does not allow segregation between duplicates, which is necessary to determine whether loci are or are not linked. For example, primer pair *Koi53-54* in family II can be

explained as a tandem duplication if the sire had two haplotypes (183, 183; 175, 183) and the dam had two haplotypes (183, 183; 158, 183). We could obtain the same genotypic proportions in progeny if the two duplicates were unlinked (fig. 2). A primer pair where parents had four alleles (*Koi111-112* [fig. 1*b*]) allows the two separate loci to segregate.

Primer pair *MFW 23* is the only suggested tandem duplication (tables 1 and 2) since it has two segregating haplotypes (126, 126); (101, 117). The two duplicates that resulted in three fragments for some of the individuals appear to be completely linked. This explanation can be made because there is polymorphism both within and between duplicates.

Sequences of Cloned Alleles

For each primer pair, the range of the allelic fragment size and the repeat number is presented in table 3. Variation was found in the number of tandem repeats and in the flanking regions of the repeat, both contributing to the observed variation in fragment size. For each primer pair, we found fragments differing in the number of repeats but identical in the sequence of the flanking regions (zero genetic distance). On the other hand, fragments differing in the sequence of the flanking region differed also in the number of repeats. The maximal genetic distances between fragments of a given primer pair (table 3) can be grouped into three distinct levels: (1) 0.005 for primer pair *Koi29-30*; (2) 0.033-0.061 for primer pairs *Koi89-90, MFW 23,* and *Koi3-4*; and (3) 0.099 to 0.114 for primer pairs *Koi35-36, Koi111-112*, and *Koi105-106*.

Genetic Relationships Within Loci

Two groups of trees were generated on the basis of the segregation patterns (fig. 3): those of diploid loci (*Koi3-4, MFW 23*, and *Koi29-30*) and those of duplicated

Table 3
Characteristics of the Sequenced Loci

Primer Pair	Type of Motif	Number of Observed Fragments	Number of Cloned Fragments	Range of Fragment Size (bp)	Range of Repeat Number	Range of Tamura-Nei's Distance
Koi3-4	CA+CT	4	5	156–190	18–35	0.000–0.061
Koi29-30	ACAA	3	3	219–255	3–8	0.000–0.005
Koi35-36	CA	4	4	172–202	4–13	0.000–0.099
Koi89-90	CT	4	4	164–194	5–20	0.000–0.033
Koi105-106	CA	6	5	141–198	7–34	0.000–0.114
Koi111-112	CA	7	6	134–171	6–24	0.000–0.109
MFW 23	CA	3	3	101–126	6–24	0.000–0.036

loci (*Koi35-36*, *Koi111-112*, and *Koi89-90*). In each tree, branches of fragments that were interpreted as alleles, based on their segregation pattern, are represented by heavy dark lines. In gene trees of duplicates, all bootstrap values that separate clades of alleles are highly significant, and the deepest split forms a distinct bifurcation. Primer pair *Koi3-4* had five fragments in two groups (fig. 3*a*). We suggest in table 1 that this locus is duplicated, but since alleles from one duplicate were not sequenced, we considered its tree as diploid. Primer pair *MFW 23* appears to be a tandem duplication. Allele 117 was found at both duplicates, and we cannot tell which of the duplicates is its origin; thus, this locus was treated as diploid (fig. 3*b*).

For all trees, we included scales of genetic distance and time since divergence, which vary among primer pairs (fig.

3). Primer pair *Koi29-30* has the smallest scale (up to 0.65 MYA), primer pairs *Koi3-4*, *MFW 23*, and *Koi89-90* have similar intermediate values (up to 5, 3.2, and 4.2 MYA, respectively), whereas primer pairs *Koi35-36* and *Koi111-112* are similar and have the largest scales (12 and 13 MYA, respectively). The deepest split of primer pair *Koi89-90* has a similar scale to the split between fragments 134 and 154 of the *Koi111-112* primer pair (about 4.2 MYA).

We included insertions and deletions (indels) and calculated the rate of total differences per base between all possible pairs of sequences. The distribution of these differences has three subgroups (fig. 4). There are 53 differences in two categories: 24 between alleles (diploid sets) and 29 between paralogs. At primer pairs for which segregation patterns were not fully informative, we could
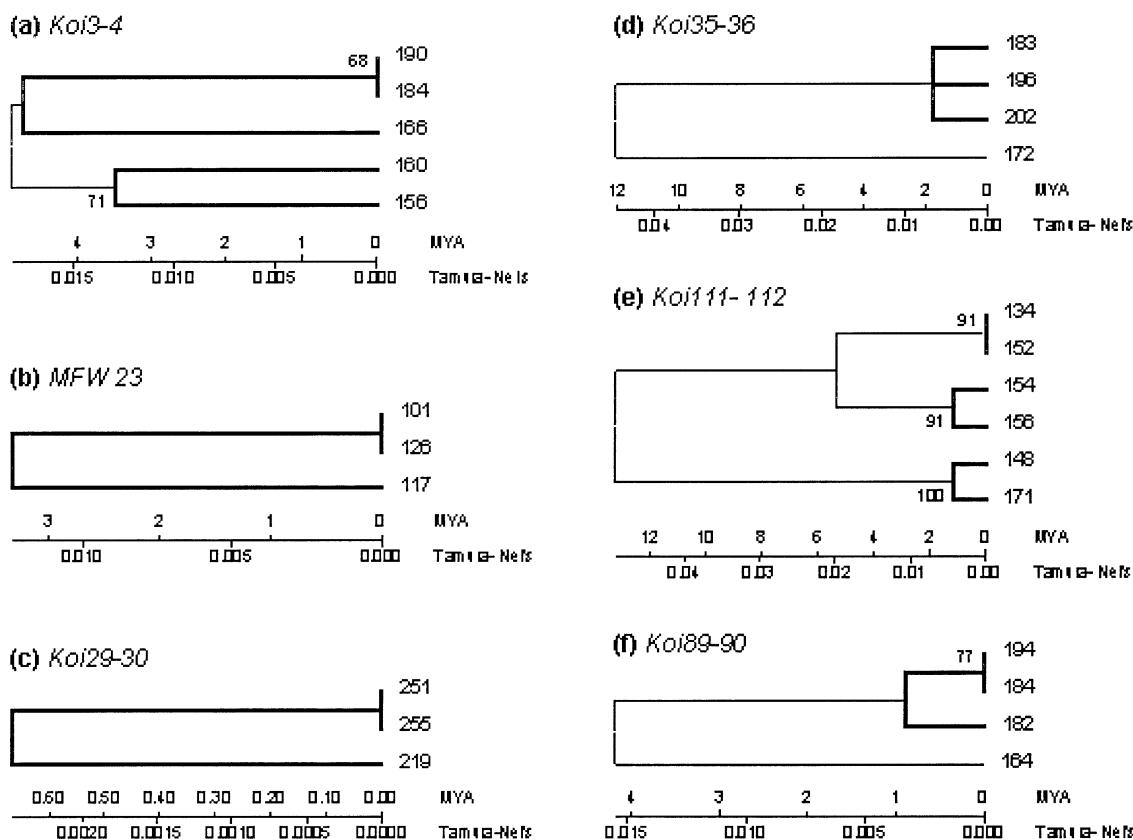

Fig. 3.—Gene trees of six primer pairs. The trees are based on sequences from flanking regions of the microsatellite alleles using Tamura-Nei's distance and Neighbor-Joining clustering. Scales are of both genetic distance and time from divergence (MYA). Darker branches connect edges that were determined as alleles by segregation pattern analysis. Numbers near nodes are significant bootstrap values (%). (*a*), (*b*), and (*c*) represent diploid loci; (*d*), (*e*), and (*f*) represent duplicated loci. Note the different scales among trees.
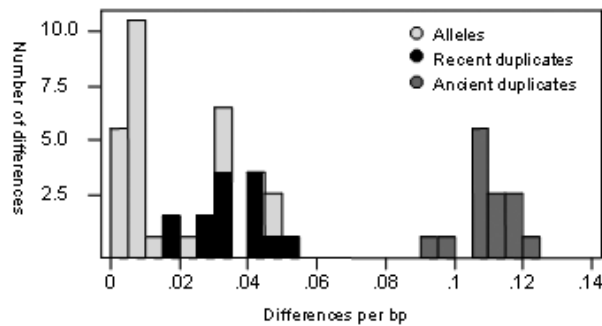
Fig. 4.—Frequency distribution of differences per base pair between sequences from duplicates and single loci. Based on 53 pairwise comparisons between fragments from seven primer pairs. Differences include substitutions and indels. Differences between sequences of the same locus (alleles) are in light gray. Differences between sequences of recent duplications are in black and differences between more ancient duplicates are in darker gray.

not define the allelic relationships between all sequences. In these cases, we used the clustering in the gene tree as an indication of allelic relationship. In figure 4, the distribution of allelic differences is skewed to the left (Shapiro-Wilk W test, $P < 0.0001$) and the mode is at 0.01. The average difference per base is 0.013, and the range is 0 to 0.047. The distribution of differences between duplicates has two distinct modes (fig. 4). The left distribution is normal ($P = 0.27$), with an average of 0.034 and a range of 0.017 to 0.050. The right distribution is also normal ($P = 0.46$), with an average of 0.109 and a range of 0.094 to 0.124. The range of allelic differences overlaps the left distribution, as expected from the scales for the gene trees (fig. 3) and the three groups of genetic distances (table 3).

The number of repeats in each allele at each locus was counted, and the $(\delta\mu)^2$ distance (Goldstein et al. 1995) between each pair of fragments was calculated separately for each locus. The average distance between alleles is 19.8, with a range of 0.25 to 60.7, whereas the average distance between paralogs is 38.5, ranging between 2.3 and 156.3 (table 4). The variance of $(\delta\mu)^2$ is 339.5 and 1987.2 for alleles and paralogs, respectively.

Divergence Time of Duplicates

Microsatellite loci in this study are in noncoding regions. We therefore used the high estimate of 3.71×10^{-9} substitutions per base per year, calculated by Li (1997, p. 91) for fourfold degenerate sites in mammals. Divergence times between pairs of sequences were first calculated using the substitution rate and then using the rate of total differences (including indels). In table 5, estimates of rates and divergence times are presented separately for the three subdistributions of figure 4, namely for alleles and for each of the two groups of duplicates. The average divergence time between alleles is 1.1 MYA, ranging between 0 and 4.6. For the younger duplicates, the average time is 4.5 MYA, with a range of 2.3 to 6.8. The average divergence time for the more ancient duplicates is 12.0 MYA, ranging between 9.5 and 13.5. The estimates based on total differences are slightly higher than those based on substitutions only (table 5). Time scales for the divergence between fragments (fig. 3) are compatible with

Table 4
Estimates of the $(\delta\mu)^2$ Genetic Distance Between Paralogs and Between Alleles Within Duplicates

	n	Mean	Variance	Range
Between paralogs	16	38.47	1987.20	2.25–156.25
Between alleles	12	19.78	339.51	0.25–60.67

the ranges found for the three inferred subdistributions (table 5). The deepest split dates the divergence to approximately 12 MYA (*Koi35-36* and *Koi111-112* [fig. 3]). The second class of more recent splits is dated to about four MYA (*Koi3-4*, *MFW 23*, *Koi89-90*, and *Koi111-112*). A third class of splits, present in all primer pairs, represents variation between alleles.

We had two microsatellite primer pairs downstream to gene loci, which enabled us to further study the duplicated nature of these regions. The first primer pair *Koi29-30* is located in the 5′-UTR of the *thyrotropin* β-*subunit* gene. We retrieved four homologous sequences from three other cyprinids, one each from *Ctenopharyngodon idella* (grass carp) and *Aristichthys nobilis* (bighead carp) and two from *Carassius auratus* (goldfish)—*thyrotropin* β-*subunit* and *gfTSHbeta* gene for *thyrotropin* β-*subunit*. The DNA sequences of the *Cyprinidae* family were aligned together with two fragments of the *Koi29-30* primer pair and a gene tree was constructed (fig. 5). The topology includes three clades: the common carp, the goldfish, and the Chinese carps (bighead and grass carp). Fragments of the *Koi29-30* primer pair are clustered with the gene sequence of the common carp. To estimate time of divergence, we used the synonymous substitution rate of 3.51×10^{-9} (Li 1997, p. 91). The common carp and the goldfish separated approximately 10.7 MYA, and both separated from the Chinese carps around 20.5 MYA. A second primer pair (*Koi35-36*) is about 30 bp downstream to the *urotensin II-γ* locus. The sequence of the *urotensin II-α* is available from GenBank. The genetic distance between the two sequences was found to be 0.0604 and corresponded to a divergence time of approximately 12.9 MYA, which is compatible with our estimates based on microsatellite fragments (12 MYA).

Discussion

Of the 54 segregation patterns, 40 (74%) had a ratio of genotypes that fits either a single or a duplicated locus. The other 14 (26%) deviated from the expected proportions of genotypes under disomic inheritance. The relatively small proportion of deviant patterns suggests that the segregation patterns are a result of reliable amplification. Some of the distorted segregation in diploid loci (type 5 in table 1) could be explained by null alleles (i.e., fragments that were not PCR amplified due to changes in primers' sites). However, segregation patterns of duplicates are more complicated to explain, especially those that deviate from expectation. Possible explanations for such deviations, other than null alleles, are selection against certain genotypes, hitchhiking of linked microsatellite alleles, or technical problems.

Higher likelihood for disomic segregation was found for duplicates. However, disomic inheritance can charac-

Table 5
Average Number of Differences and Divergence Time Between Alleles and Between Paralogs

		Alleles (n = 24 Comparisons)		Later Duplications[a] (n = 14 Comparisons)		Earlier Duplications[a] (n = 15 Comparisons)	
		K[b]	MYA[c]	K	MYA	K	MYA
Substitutions	Mean	0.008	1.079	0.033	4.469	0.089	12.031
	Minimum	0	0	0.017	2.264	0.071	9.515
	Maximum	0.034	4.596	0.050	6.792	0.100	13.477
	Standard error	0.002	0.310	0.003	0.349	0.003	0.348
Total Differences[d]	Mean	0.013	1.763	0.034	4.626	0.109	14.728
	Minimum	0	0	0.017	2.264	0.094	12.682
	Maximum	0.047	6.348	0.050	6.792	0.124	16.644
	Standard error	0.003	0.408	0.003	0.380	0.002	0.264

[a] Paralogs were divided according to the two distinct modes of figure 4.

[b] K = number of differences per site.

[c] MYA = million years ago.

[d] Total differences = substitutions plus indels.

terize both tandem and segmental duplications (e.g., primer pair *MFW 23* [table 3]). Segregation patterns of the duplicated loci, when differentially fixed or with only one segregating duplicate, can be equally explained as either tandem or unlinked duplicates (fig. 2). In fact, 20 of 21 segregation patterns of paralogs had this character. Tetraploidy however, predicts two separate duplicates rather than tandem duplication. In our study, segregation patterns strongly support the duplicated nature of loci but provide little support for a whole-genome duplication due to limited polymorphism within paralogs. However, support for tetraploidy was found in the proportion of paralogs.

Of the informative primer pairs, 60% have at least one parent with more than two fragments. We therefore offer 60% as an estimate of the extent of duplications at microsatellite loci in the common carp genome. Primer pairs such as *Koi17-18* and *Koi53-54* had up to two alleles in parents of one family and were found to be duplicated in the other family (table 1). Thus, 60% may be an underestimate of the proportion of duplicated loci since it is based on fragment number, although it agreed with the segregation patterns in 86% of the cases. In addition, a fraction of the noninformative loci (10%) could be found to be duplicated if studied in informative families. Our estimate is slightly higher than the estimate of 52% duplications based on expression of isozymes and allozymes
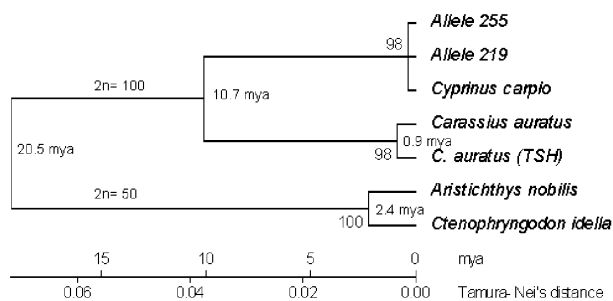

Fig. 5.—Gene tree of some cyprinids. Phylogeny is based on the DNA sequence of the *thyrotropin* gene using the Tamura-Nei distance. The tree includes flanking sequences of two microsatellite alleles that are located down stream to the gene. Numbers outside the nodes are bootstrap values (%), and those inside nodes are estimates for divergence time (MYA). Above the family branches are chromosome numbers (2n).

from 23 loci (Ferris and Whitt 1977*a*), which for the same reasons may also be an underestimate.

It has been hypothesized that *Homo sapiens*, yeast, and *Arabidopsis thaliana* are paleopolyploids, having a proportion of 12.5%, 16%, and 25% duplicated genes, respectively (Wolfe 2001). Therefore, the higher proportion of duplications in common carp may suggest a whole-genome duplication rather than several segmental duplications. In the past, the hypothesis of the common carp being a tetraploid was based mainly on its chromosome number (2n = 100 to 104 chromosomes) and high nuclear DNA content (Ohno et al. 1967). Moreover, if diploidization of duplicates occurs at a similar rate among species, one can assume that the genome duplication of the carp took place much later in evolution than the duplications in these three fully sequenced species. Therefore, our survey of microsatellite loci supports the hypothesis that this carp had relatively recent genome duplication.

The hypothesis that the genome duplication in this carp resulted from allotetraploidization has two supporting pieces of karyotipic evidence: no quadrivalents have been detected in meiotic nuclei and no chromosomes were lost in the duplication event (Ohno et al. 1967). Hybridization that results in a tetraploid species is more likely between closely related species. The carp has a doubled number of chromosomes in two distinct, although similar, sets as suggested by the identification of paralogs using the PCR method. Evolution of polyploid genomes makes the distinction between allotetraploidy and autotetraploidy, based on disomic inheritance, more difficult as time elapses (Wolfe 2001). In light of the relatively short time since tetraploidization of the carp, we suggest that the disomic inheritance in carp resulted from allotetraploidization rather than by diploidization of an autotetraploid genome. In addition, the phylogeny of cyprinids fits allotetraploidy in the sense that the clades of diploid and tetraploid species coalesce before the divergence of the diploid parents of the common carp.

Higher variation between paralogs and lower variation between alleles was suggested as evidence of the duplicated nature of four tested genes in the common carp by Larhammar and Risinger (1994). We established paralogous or allelic relationships between fragments by

segregation analysis and only then assessed their sequence variation. Our results confirm that sequence variation as well as variation in the number of repeats between paralogs is higher than that between alleles. The modal difference between alleles is 0.01, equivalent to 0.008 and ranging up to 0.034 substitutions per nucleotide. Thus, the frequency of potential SNPs in these regions was estimated to be 1/125 bases.

The distribution of variation between paralogous sequences formed two normally shaped distributions rather than one continuous distribution (fig. 4). The distributed values could result from different duplication times among loci and/or variation in evolutionary rate of these loci. Accordingly, two distinct distributions of duplications could be explained by (1) a continuous series of duplications over time that did not cover the gap between the observed modes due to small sample size, (2) two distinct duplication events with different evolutionary rates within each event, or (3) a combination of discrete and continuous events. The subdistribution of allelic differences implies that variation in evolutionary rate among loci exists. We suggest that the right subdistribution of differences between paralogs represents a whole-genome duplication with variation among loci, whereas the left subdistribution indicates a later surge of segmental duplications. The segmental duplications are represented by primer pairs *Koi89-90*, *MFW 23*, and *Koi3-4*, which show smaller genetic distances between paralogous fragments, and supported by primer pair *MFW 23*, whose segregation pattern suggested a tandem duplication. The normal subdistribution of recent duplications is different from the uniform distribution that would be predicted for a continuous series of segmental duplications. The apparently normal shape could be an artifact of small sample size that might prevent detection of younger or older representatives of this group. Alternatively, the segmental duplications could have occurred in bursts rather than continuously. Large-scale studies by McLysaght, Hokamp, and Wolfe (2002) and by Gu, Wang, and Gu (2002) suggest that at least one genome duplication took place in the evolution of humans (chordates) but small-scale duplications might have taken place as well. Gu, Wang, and Gu (2002) found two waves of duplications and interpreted these as evidence of an ancient genome duplication and a more recent expansion of gene families that resulted from tandem or segmental duplications. Our results (fig. 4), although on a much smaller scale, may be another example of a genome duplication followed by a more recent wave of segmental duplications.

The variance about the mean number of differences between paralogs is similar in magnitude to the corresponding variance between alleles (table 5), suggesting a single distribution of mutation rates for alleles and paralogs at microsatellite loci.

The average divergence age between microsatellite paralogs is 12 Myr, ranging between 9.5 and 13.5. Sequence comparison of the *urotensin II-γ* and *urotensin II-α* genes yielded a similar divergence time of 12.9 MYA, which supports the microsatellite estimate and suggests that these genes are paralogs. Our estimate based on microsatellite loci (i.e., 12 MYA) is a little lower than the previous estimate of 12 to 19 MYA based on only two genes (Larhammar and Risinger 1994). In calculating time estimates, we used substitution rates from mammals. Evidence from fish suggests lower mutation rates than in mammals (Rico, Rico, and Hewitt 1996; Krieger and Fuerst 2002) that might cause an underestimation of the times in our study. However, both divergence time and the proportion of duplicates in common carp place it among the few vertebrates in which a recent genome duplication occurred. This relatively young duplication of the genome (suggested for a few species of fish) is additional to the duplications that are proposed for all vertebrates and is also additional to that preceding the teleost fish radiation (Amores et al. 1998; Taylor et al. 2001; Wolfe 2001). These fish species, including this carp, may therefore be highly informative models for the study of genome evolution after duplication. For example, if differential fixation that was found in 28% of the microsatellite loci exists in duplicated genes, then the common carp might be a good model for the study of the functional consequences of such a state.

The region of the *thyrotropin* gene examined here suggests a coalescence time of 21 MYA for the Chinese carps (grass and bighead carps) that have $2n = 50$ chromosomes and the common carp (goldfish) that have $2n = 100$ chromosomes (fig. 5). This estimate is in agreement with the estimate of 19 MYA obtained by Larhammar and Risinger (1994). It is reasonable to assume one genome duplication for the common carp and the goldfish, which took place before their divergence at 11 MYA. Therefore, 11 to 21 MYA is the estimated interval at which genome duplication of the common carp took place. In theory, the two species that formed the ancestor of the common carp should have been close enough to be able to hybridize but different enough to generate disomic inheritance. The divergence time between paralogs provides an estimate of the time at which the two ancestral diploid species separated, and indeed 12 MYA is within the time interval that was suggested by the gene tree of *thyrotropin*. If the divergence of the common carp and the goldfish followed shortly after the genome duplication, as suggested by their phylogeny, then it may be an example of the role of polyploidization in speciation and diversification.

Literature Cited

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

Allendorf, F. W., and G. H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes. Pp. 1–53 *in* B. J. Turner, ed. Evolutionary genetics of fishes. Plenum, New York.

Amores, A., A. Force, Y. L. Yan et al. (13 co-authors). 1998. Zebrafish hox clusters and vertebrate genome evolution. Science **282**:1711–1714.

Angers, B., K. Gharbi, and A. Estoup. 2002. Evidence of gene conversion events between paralogous sequences produced by tetraploidization in *Salmoninae* fish. J. Mol. Evol. **54**:501–510.

Crooijmans, R. P. M. A., V. A. F. Bierbooms, J. Komen, J. J. Van Der Poel, and M. A. M. Groenen. 1997. Microsatellite markers in common carp (*Cyprinus carpio* L.). Anim. Genet. **28**:129–134.

Bailey, J. A., Z. Gu, R. A. Clark, K. Reinert, R. V. Samonte, S. Schwartz, M. D. Adams, E. W. Myers, P. W. Li and E. E. Eichler. 2002. Recent segmental duplications in the human genome. Science. **297**:1003–1007.

David, L., F. Jinggui, R. Palanisamy, J. Hillel, and U. Lavi. 2001. Polymorphism in ornamental and common carp strains (*Cyprinus carpio* L.) as revealed by AFLP analysis and a new set of microsatellite markers. Mol. Gen. Genomics **266**:353–362.

Ferris, S. D., and G. S. Whitt. 1977*a*. The evolution of duplicate gene expression in the carp (*Cyprinus carpio*). Experientia **33**:1299–1301.

———. 1977*b*. Duplicate gene expression in diploid and tetraploid loaches (Cypriniformes, *Cobitidae*). Biochem. Genet. **15**:1097–1112.

Friedman, R., and A. L. Hughes. 2001. Pattern and timing of gene duplications in animal genomes. Genome Res. **11**:1842–1847.

Goldstein, D. B., A. Ruíz-Linares, L. L. Cavalli-Sforza. and M. W. Feldman, 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. USA **92**:6723–6727.

Gu, X., Y. Wang, and J. Gu. 2002. Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution. Nat. Genet. **31**:205–209.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.

Holland, P. W., J. Garcia-Fernandez, N. A. Williams, and A. Sidow. 1994. Gene duplications and the origins of vertebrate development. Dev. Suppl. 125–133.

Hughes, A. L., J. da Silva, and R. Friedman. 2001. Ancient genome duplications did not structure the human *Hox*-bearing chromosomes. Genome Res. **11**:771–780.

Krieger, J., and P. A. Fuerst. 2002. Evidence for a slowed rate of molecular evolution in the order Acipenseriformes. Mol. Biol. Evol. **19**:891–897.

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**:1244–1245.

Larhammar, D., and C. Risinger. 1994. Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio*. Mol. Phylogenet. Evol. **3**:59–68.

Li, W. H. 1997. Molecular evolution. Sinauer Associates, Sunderland, Mass.

Llorente, B., A. Malpertuy, C. Neuveglise et al. (24 co-authors). 2000. Genomic exploration of the hemiascomycetous yeasts: 18. Comparative analysis of chromosome maps and synteny with *Saccharomyces cerevisiae*. FEBS Lett. **487**:101–112.

Ludwig, A., N. M. Belfiore, C. Pitra, V. Svirsky, and I. Jenneckens. 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics **158**:1203–1215.

McLysaght, A., K. Hokamp, and K. H. Wolfe. 2002. Extensive genomic duplication during early chordate evolution. Nat. Genet. **31**:200–204.

Meyer, A., and E. Malaga-Trillo. 1999. Vertebrate genomics: more fishy tales about *hox* genes. Curr. Biol. **9**:R210–R213.

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

Ohno, S. 1970. Evolution by gene duplication. Allen and Unwin, London.

Ohno, S., J. Muramoto, L. Christian, and N. B. Atkin. 1967. Diploid-tetraploid relationship among Old World members of the fish family *Cyprinidae*. Chromosoma (Berl.) **23**:1–9.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Annu. Rev. Genet. **34**:401–437.

Postlethwait, J. H., I. G. Woods, P. Ngo-Hazelett, Y. L. Yan, P. D. Kelly, F. Chu, H. Huang, A. Hill-Force, and W. S. Talbot. 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res. **10**:1890–1902.

Postlethwait, J. H., Y. L. Yan, M. A. Gates et al. (29 co-authors). 1998. Vertebrate genome evolution and the zebrafish gene map. Nat. Genet. **18**:345–349.

Pyatskowit, J. D., C. C. Krueger, H. L. Kincaid, and B. May. 2001. Inheritance of microsatellite loci in the polyploid lake sturgeon (*Acipenser fulvescens*). Genome **44**:185–191.

Rico, C., I. Rico, and G. Hewitt. 1996. 470 million years of conservation of microsatellite loci among fish species. Proc. R. Soc. Lond. B Biol. Sci. **263**:549–557.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Sidow, A. 1996. Gen(om)e duplications in the evolution of early vertebrates. Curr. Opin. Genet. Dev. **6**:715–722.

Skrabanek, L., and K. H. Wolfe. 1998. Eukaryote genome duplication—where's the evidence? Curr. Opin. Genet. Dev. **8**:694–700.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10**:512–526.

Taylor, J. S., Y. Van de Peer, I. Braasch, and A. Meyer. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. Philos. Trans. R. Soc. Lond. B Biol. Sci. **356**:1661–1679.

Uyeno, T., and G. R. Smith. 1972. Tetraploid origin of the karyotype of catostomid fishes. Science **175**:644–646.

Wolfe, K. H. 2001. Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. **2**:333–341.

Zhang, H., N. Okamoto, and Y. Ikeda. 1995. Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). Gene **153**:231–236.

Kenneth Wolfe, Associate Editor

Accepted March 27, 2003