Chinese Science Bulletin

The *c-myc* coding DNA sequences of cyprinids (Teleostei: Cypriniformes): Implications for phylogeny

KONG XiangHui[1,2], WANG XuZhen[1], GAN XiaoNi[1] & HE ShunPing[1†]

[1] Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China;
[2] College of Life Science, Henan Normal University, Xinxiang 453007, China

The family Cyprinidae is one of the largest fish families in the world, which is widely distributed in East Asian, with obvious difference in characteristic size among species. The phylogenetic analysis of cyprinid taxa based on the functionally important genes can help to understand the speciation and functional divergence of the Cyprinidae. The *c-myc* gene is an important gene regulating individual growth. In the present study, the sequence variations of the cyprinid *c-myc* gene and their phylogenetic significance were analyzed. The 41 complete sequences of the *c-myc* gene were obtained from cyprinids and outgroups through PCR amplification and clone. The coding DNA sequences of the *c-myc* gene were used to infer molecular phylogenetic relationships within the Cyprinidae. *Myxocyprinus asiaticus* (Catostomidae), *Misgurnus anguillicaudatus* (Cobitidae) and *Hemimyzon sinensis* (Homalopteridae) were assigned to the outgroup taxa. Phylogenetic analyses using maximum parsimony (MP), maximum likelihood (ML), and Bayesian retrieved similar topology. Within the Cyprinidae, Leuciscini and Barbini formed the monophyletic lineage respectively with high nodal supports. Leuciscini comprises Xenocyprinae, Cultrinae, East Asian species of Leuciscinae and Danioninae, Gobioninae and Acheilognathinae, and Barbini contains Schizothoracinae, Barbinae, Cypininae and Labeoninae. *Danio rerio*, *D. myersi* and *Rasbora trilineata* were supposed to separate from Leuciscinae and Barbini and to form another lineage. The positions of some Danioninae species were still unresolved. Analyses of both amino acid variation with parsimony information and two high variation regions indicated that there is no correlation between variations of single amino acid or high variation regions and characteristic size of cyprinids. In addition, the species with smaller size were usually found to be basal within clades in the tree, which might be the results of the adaptation to the primitive ecology and survival pressure.

Cyprinidae, *c-myc* gene, phylogeny, sequence variation

The family cyprinidae is one of the largest fish families in the world, which contains approximately 210 genera and 2010 species[1] and is widely distributed in Eurasia, East Indian Island, Africa and North America[2]. The Cyprinidae obtained its greatest species diversity in East Asia. For example, of the more than 600 species within 122 genera distributed in China, there are 384 endemic species[1]. The classification of the Cyprinidae based on the morphology has been subject to revisions since it was first established by Cuvier. Chen et al.[3] grouped the Cyprinidae into the two lineages (Leuciscini and Barbini) with 10 subfamilies, and merged Gobiobotinae and Schizothoracinae into Gobioninae and Barbinae, respectively. On the basis of barbels distribution, morphotype and innervations, Howes[4] recognized the following subfamilies within the Cyprinidae: Cyprininae, Gobioninae, Rasborinae (without or sporadically with barbels), Leuciscinae, Acheilognathinae, Cultrinae and Alburninae. The rearrangements for Danioninae, Leu-

Received December 26, 2006; accepted March 19, 2007
doi: 10.1007/s11434-007-0216-y
†Corresponding author (email: clad@ihb.ac.cn)
Supported by the Development Plan of the State Key Fundamental Research of China (Grant No. 2004CB117402), the National Natural Science Key Foundation of China (Grant No. 30530120) and Chinese Postdoctoral Science Foundation (Grant No. 2005037684)

ciscinae and Xenocyprinae by Chen et al. were different from that by Howes. According to the work of Chen et al., evolutionary relationships presented by Cavender and Coburn[5] were similar to that proposed by Chen et al.[3]. According to the latest taxonomic revision proposed by Chen[1], the family Cyprinidae was divided into 12 subfamilies: Danioninae, Leuciscinae, Cultrinae, Xenocyprinae, Hypophthalmichthyinae, Cobioninae, Gobiobotinae, Acheilognathinae, Barbinae, Labeoninae, Schizothoracinae, and Cyprininae. Furthermore, the placement and significance of East Asian group were emphasized on the phylogeny of the Cyprinidae.

With the development of molecular biology methods, more and more molecular data were used to reconstruct phylogenetic relationships within the family Cyprinidae. Complete mitochondrial cytochrome *b*, partial 16S rRNA mitochondrial DNA, control region were used to reconstruct phylogeny among European cyprinids[6−9]. The mitochondrial 12S rRNA and 16S rRNA genes were anylazed for North American cyprinids[10−12]. Mitochondrial cytochrome *b* and ND4 sequence were used to infer the phylogenetic relationships in the subfamily Xenocyprinae[13]. Phylogenetic relationships of the Cyprinidae in East Asia were inferred from the mitochondrial cytochrome *b* by He et al.[14,15]. Phylogeny of the lower-level cyprinids in East Asia was reconstructed by Wang et al.[16,17] based on cytochrome *b* and 1st intron of the *S7* ribosome protein gene. Up to date, though there still exist some controversies, phylogenetic relationships of the Cyprinidae were more and more clear. However, these previous investigations were focused on gene sequences with highly evolutionary rate, such as mitochondrial gene and intron sequence. Although few studies were conducted using the nuclear genes with important biological function, the phylogeny inferred from the functionally important nuclear gene could be helpful to understanding the evolutionary relationships, speciation, gene variation and functional divergence within the family Cyprinidae.

The cellular myelocytomatosis oncogene (*c-myc*) is an important member of the *myc* family, which is a transcript activity factor having crucial function in regulating development and growth of animal individual and is ubiquitously expressed in embryo or proliferation cell in adult tissues. The structure and nucleotide composition of the *c-myc* gene are conserved, and it generally comprises 3 exons and 2 introns. The *c-myc* gene is ubiquitous in mammal, avian, amphibian and fish, and the origin of the *c-myc* gene has been estimated to have occurred at least 600 Mya ago[18]. The *c-myc* gene was a single copy with an open reading frame existing in exons 2 and 3[19]. While mutations of the *c-myc* gene occurred, the adult size of *Drosophila* or mouse was reduced obviously, which was mainly controlled by cell size[20] or cell number[21]. Although nucleotide sequences of the *c-myc* gene are divergent between *Drosophila* and mouse, protein sequences are conserved and have similar biological functions. Furthermore, the *c-myc* gene in Drosophila can partially complement the deficiency in mouse[22]. The *c-myc* gene mutations resulted in the reduced individual size, and tumor[23] had already been proved to occur in animals. Recently, the *c-myc* gene was used to infer phylogeny in mammals[24], avians[25], crocodiles[26], and frogs[27], and served as a useful molecular marker to investigate the higher-level phylogenies such as order and family. However, as far as we know, phylogenetic studies of fishes based on the *c-myc* gene have not been available.

The Cyprinidae, with remarkable species size and food intake, occupying different ecological niches, is an ideal group for the study of gene evolution and phylogeny. In this work, the complete *c-myc* gene sequences were achieved from the 41 representative species from the Cyprinidae ingroup and outgroup using PCR amplification and clone, and then, the *c-myc* coding DNA was extracted to analyze the cyprinid phylogeny. In addition, each amino acid site with parsimony information in the translated protein sequence was analyzed to retrieve the correlation between variation of amino acid and characteristic size of cyprinids. The aims of investigation were (1) to reconstruct phylogenetic relationships in the family cyprinidae, (2) to resolve the systematic positions of the previous controversial species and (3) to explore the correlation between variations of the *c-myc* gene and the characteristic size of cyprinid species.

1 Materials and methods

1.1 Sample collection

In this study, novel *c-myc* DNA sequences were achieved from 41 representative species (Table 1) in the Fish Museum of the Institute of Hydrobiology, the Chinese Academy of Sciences. 38 cyprinid representative species were selected from various subfamilies. The following 3 species from 3 families belonging to the

same order Cypriniformes: *Myxocyprinus asiaticus* (Catostomidae), *Misgurnus anguillicaudatus* (Cobitidae) and *Hemimyzon sinensis* (Homalopteridae) were assigned to the outgroup taxa. The collected locations, deposited voucher and GenBank accession numbers of all species are listed in Table 1. Muscle or fin tissue preserved in 95% ethanol was used to extract the genome DNA.

1.2 Primer design

According to the conserved regions from the *c-myc* gene sequences deposited in GenBank, such as *Danio rerio* (NM_131412), *Cyprinus Carpio* (D37888) and *Carassius auratu* (D31729), primers were designed and optimized. At last, the following primers were used to amplify the *c-myc* DNA sequences: PF6: 5′-ATYAgTCTgTCCAgCAYCT-3′; PR7: 5′-SRAACTCgCTgACCATCTC-3′; PF8: 5′-KSGTTGWTTAYATTTTCCATCAC-3′; PR8: 5′-GRAACTCGSTSACYATCTC-3′; PF11: 5′-AATGCYGGTGAGTKCGAGTT-3′; PR11: 5′-GCTGMAGCYTGTGTTTTAACTGT-3′; PF12: 5′-TGGAGATRGTSAGCGAGTT-3′. Distribution of the various primers in the *c-myc* gene is illustrated in Figure 1.

Table 1 Samples of cyprinid ingroup and outgroup taxa in this study

Species	Sampling location	Voucher	Accession No.
Aristichthys nobilis	Wuhan, Hubei Prov.	IHBCYK0411001	EF194848
Hypophthalmichthys molitrix	Wuhan, Hubei Prov.	IHBCYK0411002	EF194849
Ctenopharyngodon idellus	Wuhan, Hubei Prov.	IHBCYK0411003	EF194850
Mylopharyngodon piceus	Wuhan, Hubei Prov.	IHBCYK0411004	EF194851
Ochetobius elongates	Tengxian, Guangxi AR	IHBCY0108003	EF194852
Elopichthys bambusa	Taoyuan, Hunan Prov.	NRMT2286	EF194853
Squaliobarbus curriculus	Wuhan, Hubei Prov.	IHBCY0407001	EF194854
Culter alburnus	Wuhan, Hubei Prov.	IHBCY0380494	EF194855
Megalobrama amblycephala	Wuhan, Hubei Prov.	IHBCY0305004	EF194856
Hemiculter leucisculus	Wuhan, Hubei Prov.	IHBCY2603026	EF194857
Pseudobrama simoni	Taoyuan, Hunan Prov.	IHBCY0405361	EF194858
Xenocypris argentea	Taoyuan, Hunan Prov.	IHBCY0405138	EF194859
Aphyocypris chinensis	Wuhan, Hubei Prov.	IHBCYK0411005	EF194860
Opsariichthys bidens	Taoyuan, Hunan Prov.	NRMT2358	EF194861
Saurogobio gracilicaudatus	Guizhou, Prov.	IHBCY0312012	EF194862
Saurogobio dabryi	Changyang, Hubei Prov.	IHBCY0405136	EF194863
Coreius heterodon	Wuhan, Hubei Prov.	IHBCY0312002	EF194864
Pseudorasbora parva	Mengla, Yunnan Prov.	IHBCY0312003	EF194865
Gobiocypris rarus	Wuhan, Hubei Prov.	IHBCYK0411006	EF194866
Tanichthys albonubes	Guangdong Prov.	IHBCYK0411007	EF194867
Rhodeus ocellatus	Wuhan, Hubei Prov.	IHBCYK0411008	EF194868
Rhodeus lighti	Wuhan, Hubei Prov.	IHBCYK0411009	EF194869
Paracheilognathus imberbis	Wuhan, Hubei Prov.	IHBCYK0411010	EF194870
Danio rerio	Wuhan, Hubei Prov.	IHBCYK0411011	EF194871
Danio myersi	Mengla, Yunnan Prov.	IHBCY0405411	EF194872
Rasbora trilineata	Wuhan, Hubei Prov.	IHBCYK0411012	EF194873
Schizothorax longibarbus	Xizang Prov.	IHBCY0510081	EF194874
Schizothorax oconnori	Xizang Prov.	IHBCY0510086	EF194875
Schizothorax lissolabiatus	Xizang Prov.	IHBCY0504193	EF194876
Gymnocypris waddelli	Qinghai Prov.	IHBCY0510092	EF194877
Puntius semifasciolatus	Mengla, Yunnan Prov.	IHBCY0405496	EF194879
Percocypris pingi pingi	Hejiang, Sichuan Prov.	IHBCY0205010	EF194881
Carassius auratus	Wuhan, Hubei Prov.	IHBCYK0411013	EF194880
Cyprinus Carpio	Wuhan, Hubei Prov.	IHBCYK0411014	EF194882
Spinibarbus sinensis	Nanchong, Sichuan AR	IHBCY0207036	EF194883
Ptychidio jordani	Yunnan Prov.	IHBCY0308004	EF194884
Epalzeorhynchus bicornis	Wuhan, Hubei Prov.	IHBCY0505291	EF194885
Garra kempi	Chayu,Xizang Prov.	IHBCY0309091	EF194886
Hemimyzon sinensis	Yunnan Prov.	IHBCYK0311012	EF194887
Misgurnus anguillicaudatus	Dayaoxian, Yunnan Prov.	IHBCYK0411015	EF194888
Myxocyprinus asiaticus	Wuhan, Hubei Prov.	IHBCY0305001	EF194889

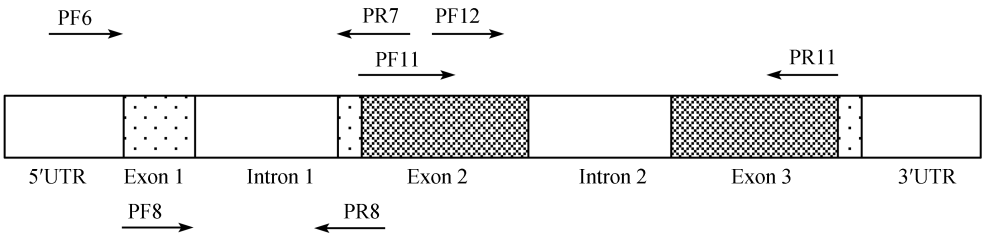

Figure 1 The *c-myc* gene structure and primers distribution.

1.3 DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from muscle or fin tissues using phenol/chloroform standard extraction procedure[28]. The *c-myc* gene was amplified from total DNA extracts using the polymerase chain reaction (PCR). The concentration and purity of DNA were determined by Beckman ultraviolet measurement and DNA concentration was diluted to 30—50 ng/μL. 60 μL PCR amplification mixtures (Various components from TaKaRa) contained 3 μL of DNA template, 6 μL of 10× Ex *Taq* PCR buffer, 4.8 μL dNTPs (each at 2.5 mmol/L, pH 8.0), 1.5 μL of each oligonucleotide primer (each at 15 μmol/L), 0.6 μL Ex *Taq* polymerase (5 U/μL), 42.6 μL dd H_2O. For oligonucleotide primer PF11 and PR11, PF12 and PR11, the PCR amplification profile included an initial denaturation step at 94℃ for 4 min, followed by 32 cycles of denaturation of 50 s at 94℃, annealing for 50 s at 57℃ (PF11 and PR11) or 56℃ (PF12 and PR11), 90 s extension at 72℃, and a final 6 min extension at 72℃. For primers PF6 and PR7, PF8 and PR8, reactions were carried out at an initial denaturation at 94℃ for 4 min, 32 cycles of denaturation at 94℃ for 30 s, annealing at 55℃ (PF6 and PR7) or 50℃ (PF8 and PR8) for 30 s and extension at 72℃ for 60 s, and a final extension at 72℃ for 5 min.

The amplified objective fragments were separated accurately by 1.2% agarose gel electrophoresis, purified using an OMEGA kit (From OMEGA bio-tek) and connected into a T-tailed pMD18-T vector (From TaKaRa). Enough positive clones were achieved with host DH5α bacteria, and the positive clones were determined by a PCR procedure. The triplicate positive clones carrying the intention fragments were sequenced in ABI3730.

1.4 The *c-myc* gene characteristics

The *c-myc* gene sequences have been deposited in GenBank (Table 1). Multiple alignments of sequences were performed using CLUSTAL X(1.83)[29] with a gap-opening penalty of 15.0 and a gap-extension penalty of 3.0. The aligned sequences with a manual correction were used to analyze gene characteristics. The *c-myc* coding DNA sequences base composition and substitution were calculated by MEGA3.1[30]. Mutation saturation analyses for nucleotide substitutions were estimated by DAMBE (V4.1.33)[31] from the slope of a linear regression of transversions and transitions against F84 distance.

1.5 Phylogenetic analysis

Phylogenetic relationships within the family Cyprinidae were reconstructed using maximum parsimony (MP), maximum likelihood (ML) and Bayesian method. The optimal bases substitution model and the optimized parameters for maximum likelihood analysis were estimated by Modeltest 3.7[32] according to Akaike information criterion (AIC). The best model is GTR+I+G ($-\ln L$=7760.8984; K=10; AIC=15541.7969). MP and ML trees were reconstructed by PAUP4.0b10[33]. Bayesian analysis was conducted by MrBayes3.1.2[34,35].

Maximum parsimony analyses were performed using the heuristic search. Bootstrap scores were calculated with 1000 random stepwise addition sequence replicates and tree-bisection-reconnection. The parameters for ML analyses were: Base=(0.2846 0.2567 0.2678), Rmat=(0.3476 3.8968 1.0766 0.3006 3.2010), Rates=gamma, Shape=0.7299, and Pinvar=0.4566. 100 bootstraps were conducted for ML. The parameters for Bayesian analysis were set: nst=6, rates=invgamma, Ngen=1000000, Nruns=2, Nchains=4, Burnin=600.

1.6 *c-myc* coding DNA variation analysis

With the reported and determined *D. rerio c-myc* coding DNA sequences, the aligned *c-myc* gene coding DNA sequences were translated into amino acid sequences. Each amino acid site with parsimony information was analyzed to search for the correlation between local mu-

tation and species characteristic size. In addition, two high variation regions in *c-myc* coding DNA were analyzed by calculating gap(–) number changes with Excel 2003, to explore the correlation between changes and species relationships or characteristic size.

2 Results

2.1 The *c-myc* gene characteristics

The alignment of the *c-myc* coding DNA sequences of 41 species were 1246 bp. Of these sites, 798 characters were conservative, 448 sites were variable, and 279 were parsimony informative. The mean base compositions of T, C, A, and G among all taxa were 18.1%, 26.4%, 28.5% and 27.0%, respectively. The ratio of transition to transversion was 2.2. Transition and transversion substitutions of the *c-myc* coding DNA dataset increased linearly against F84 distance, indicating that bases changes at these sites are not saturated.

2.2 Phylogenetic relationships

Unweighted maximum parsimony analysis yielded 111 equally parsimonious trees (tree length=1177, consistency index (CI)=0.5200, homoplasy index (HI)=0.4800, and retention index (RI)=0.6146). The 50% majority-rule consensus tree is shown in Figure 2. In the MP tree, the Cyprinidae formed a monophyletic lineage with 99% bootstrap support. Within the Cyprinidae, Leuciscini is a monophyletic lineage with 70% nodal support, which comprises Xenocyprinae, Cultrinae, Gobioninae, Acheilognathinae and East Asian species of Leuciscinae and Danioninae. Barbini is a monophyletic lineage with 95% bootstrap scores, which comprises Schizothoracinae, Barbinae, Cyprininae and Labeoninae. The cluster of *D. rerio* and *D. myersi* is a sister group of Leuciscinae; *Rasbora trilineata* is at the basal place of the family Cyprinidae; however, these nodal support are the lower scores.

The ML tree (Figure 3) achieved with heuristic search is mostly identical to the MP tree. Leuciscini and Barbini were grouped into monophyletic lineage with 80% and 96% bootstrap supports, respectively. Within Leuciscini, the monophyletic lineage (East Asian Clade + Gobioninae + *Tanichthys*) is a sister group of Acheilognathinae. *D. rerio*, *D. myersi* and *R. trilineata* were clustered into a monophyletic lineage, which was a sister group of the monophyletic Barbini; however, the bootstrap score is less than 50%.

Bayesian tree (Figure 4) was mostly similar to ML tree in topology. Within the family Cyprinidae, Leuciscini and Barbini were monophyletic lineages with 100% posterior probability support. Within Barbini, Barbinae, with some species appearing in Schizothoracinae and Cyprininae, is not a monophyletic lineage. The classification of Barbinae was still unresolved in this study.

2.3 Correlation between variation and species size

The *c-myc* gene coding DNA sequences are specially conserved. However, there exist two high variation regions (Figure 5). Variation region 1 is the tandem repeat of GAG(GAA)GAG. Variation region 2 is the tandem repeat of AGC(AAC)AGC. Changes of the gap number in each region indicated that there was no correlation between these changes and species size of cyprinids. After the *c-myc* coding DNA sequences were translated into protein, variations in each amino acid with parsimony information did not show the correlation relationship with species characteristic size.

3 Discussion

The *c-myc* gene, an important regulating gene in transcript activity, development, and individual growth, is first used to analyze the phylogeny of the Cyprinidae in this study. The phylogenetic relationships obtained from the *c-myc* coding DNA sequences are partly consistent with the previous phylogenies based on the mitochondrial genes[15−17]. Therefore, the *c-myc* coding DNA is a nuclear molecular marker suitable to the phylogenetic study of fishes at the level of family and subfamily.

In the present study, the well-resolved monophyletic lineages Leuciscini and Barbini agree with the suggestions based on morphology[3] and the mitochondrial cytochrome *b*[15]. Leuciscini comprises East Asian clade, Gobioninae and Acheilognathinae. Barbini comprises Schizothoracinae, Barbinae, Cyprininae and Labeoninae. *D. rerio*, *D. myersi* and *R. trilineata* should be treated as a group separated from Leuciscine and Barbini. The East Asian clade defined by He et al.[15] that contains Xenocyprinae, Cultrinae, East Asia species of Leuciscinae and Danioninae, are clustered into a monophyletic lineage with a high nodal support. Within leuciscini, Gobioninae and Acheilognathinae are two monophyletic lineages.

Even in terms of the morphological characteristics,

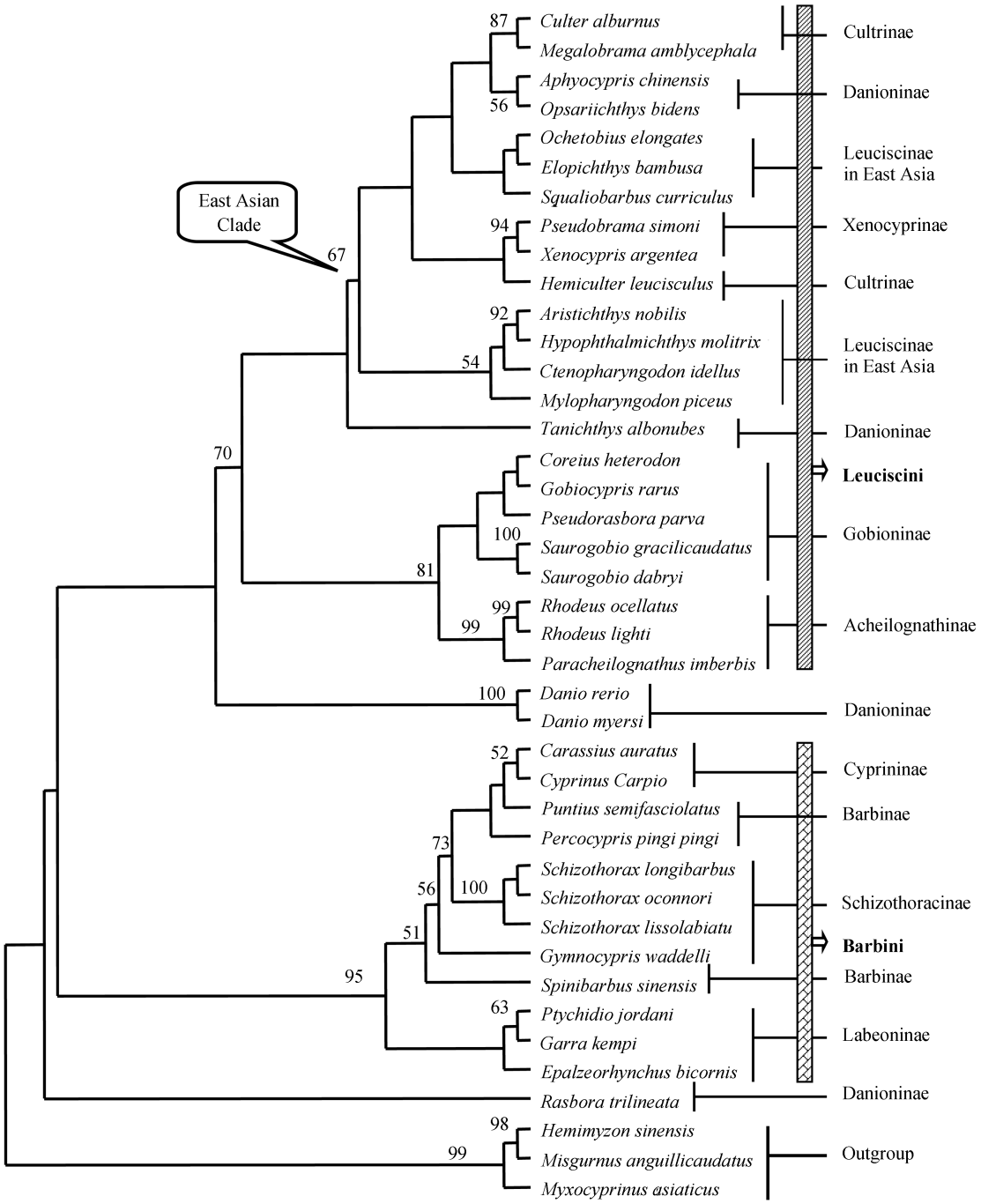

Figure 2 The 50% major-rule consensus tree of the 111 equally parsimonious trees obtained from most parsimony analysis of the *c-myc* coding DNA sequences. Tree Length=1177, CI=0.5200, RI=0.6146, and RC=0.3196. The numbers on the nodes refer to the bootstrap scores (above 50%).

the subfamily Danioninae was considered to be a complex assemblage[4]. In this study, it was also proved that Danioninae should not be a monophyletic lineage. *Aphyocypris chinensis*, *Opsariichthys bidens*, and *T. albonubes* were clustered into the Leuciscini lineage. *A. chinensis* and *O. bidens* belong to East Asian clade. *T. albonubes* was the closest to East Asian clade. The clade composed of *D. rerio* and *D. myersi* was a sister group of Leuciscini with lower bootstrap support in MP tree. In ML tree, *D. rerio*, *D. myersi* and *R. trilineata* were clustered into a lineage, which was a sister group with Barbini with lower nodal support. In Bayesian tree, *D. rerio*, *D. myersi* and *R. trilineata* were separated from Leuciscini and Barbini with higher nodal support. Ac-

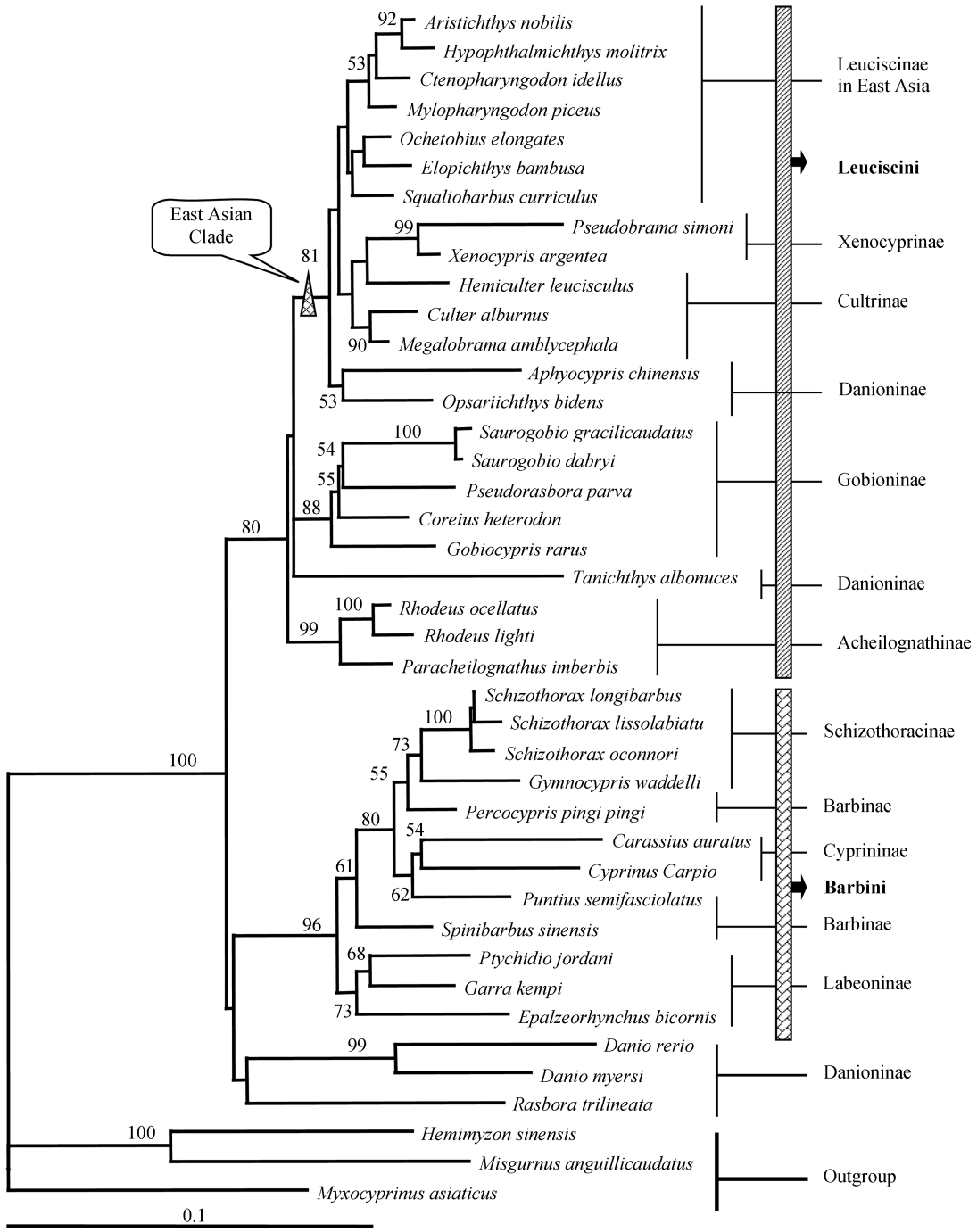

Figure 3 The tree obtained from maximum likelihood analysis of the *c-myc* coding DNA sequences. The model GTR+I+G used for ML analysis was estimated from Modeltest3.7. The numbers on the nodes refer to the bootstrap scores (above 50%).

cording to the present analyses, the family Cyprinidae could be grouped preferably into three lineages: Leuciscini, Barbini and primitive Danionini (*D. rerio*, *D. myersi* and *R. trilineata*). The positions of the species such as *T. albonubes*, *D. rerio*, *D. myersi*, and *R. trilineata* were still unresolved in this study; however, our analyses shed some light on these questions. The controversial *Gobiocypris rarus* was closely related with Gobioninae with high bootstrap score in this study.

In Bayesian tree, *Hemiculter leucisculus* was clustered into Xenocyprinae. This result indicated that *H. leucisculus* is close to Xenocyprinae. The same results

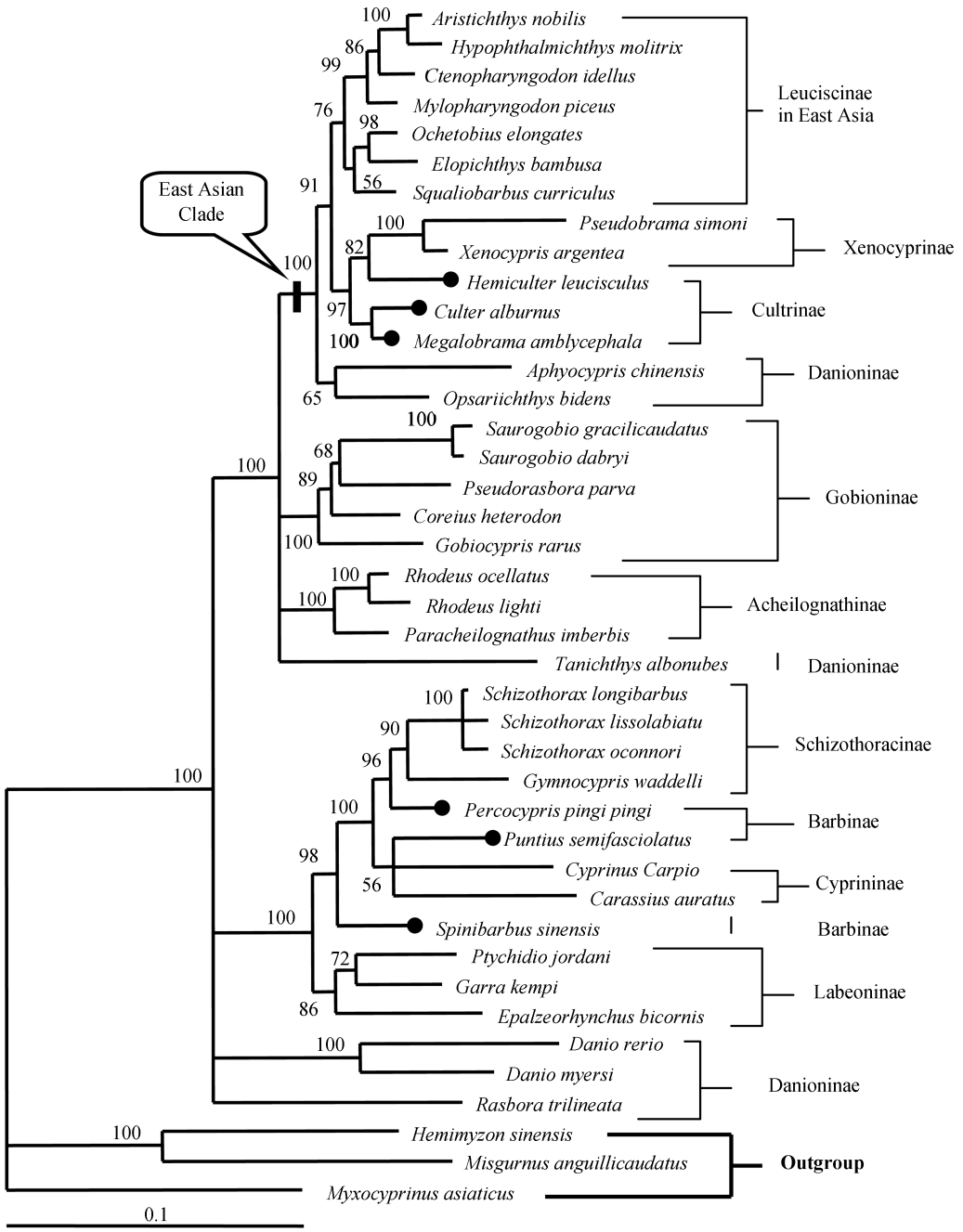

Figure 4 Bayesian tree obtained from analysis of the *c-myc* coding DNA sequences. The numbers on the nodes represent Bayesian posterior probabilities.

appeared in MP and ML topology, but the bootstrap scores were below 50%. Within Barbini, Barbinae, Cyprininae and Schizothoracinae were clustered into a monophyletic lineage with 98% nodal support. Of the Barbinae, some species were clustered into Schizothoracinae, and some merged into Cyprininae. These results indicated that there still existed some unresolved relationships in Barbinae.

The *c-myc* gene plays a crucial role not only in regulating cell growth, development, differentiation and apoptosis[18], but also in controlling individual final size[21]. In this study, variations of each amino acid with parsimony information were not correlated with species size. The *c-myc* coding DNA is a conserved sequence with two high variation regions. The first is the tandem repeat of GAG(GAA)GAG coding glutamic acid (Glu and E); the second is the tandem repeat of AGC(AAC)AGC coding Serine (Ser and S) or Aspar-

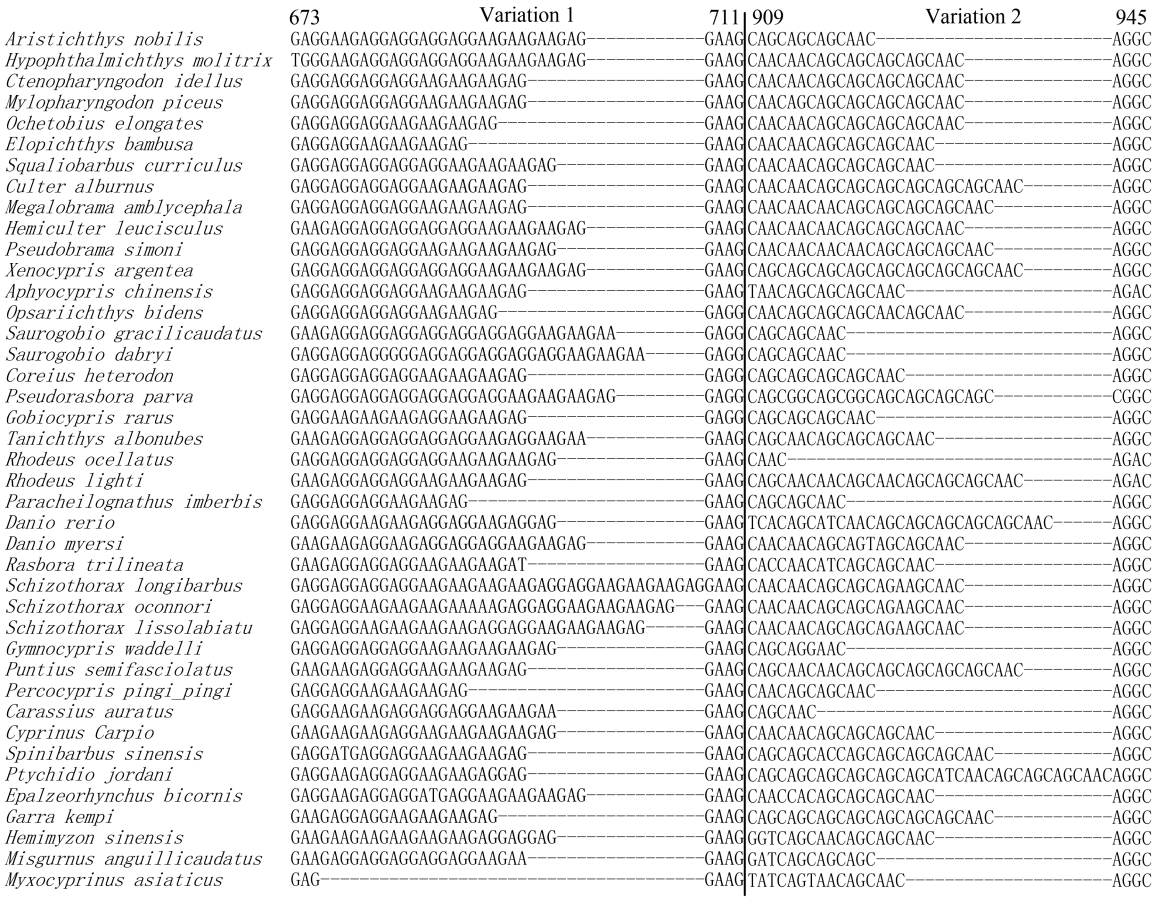

Figure 5 Two highly variable regions in the cyprinid *c-myc* coding DNA sequences.

agine (Asn and N). These two variation regions are Casein kinase II phosphorylation sites (CK-II)[36] with important biological function in regulating *c-myc* activity. Even between the closest species, there exists the elusory difference in high variation regions, so two variation regions could be deleted in phylogenetic analysis. In addition, these variation patterns are different in each region among the larger size species or the smaller size species, so these changes are incapable of explaining species divergence and final size control, which might be related to the activity of folded proteins. This hypothesis is to be tested by further study of protein function by altering variation region sequences.

In this study, phylogenetic analysis based on the *c-myc* coding DNA sequences showed that the smaller size species (such as Danioninae, Acheilognathinae and Labeoninae) cannot be clustered into a group. Therefore, the phylogenetic analysis of the *c-myc* coding DNA is not able to explain the riddle of species size control in the Cyprinidae. For each of molecular phylogenetic trees, the smaller size species are usually basal within a lineage, so the smaller size species are primitive. This supposal might be explained from the angle of species evolution and adaptation. When appearing in primitivity, due to survival pressure, the primitive fishes generally survived by the reproductive strategy of large quantity and the small species size. So the primitive lineages Danioninae, Acheilognathinae and Labeoninae are generally small. The diversification of East Asian clade of Leuciscinae is closely related with the formation of complex ecology system of rivers and lakes with rich food sources. The species with rich food and appropriate ecological environment are generally large in individual size. This deduction could be of certain help to the understanding of the difference in characteristic size between different species. However, further evidence is needed

The authors would like to thank LI JunBing for species collection and identification, HAN YaWei for suggestions in experiments, and HOU YuTing for partial works in experiment.

1 Chen Y Y. Fauna Sinica, Osteichthyes, Cypriniformes (II) (in Chinese). Beijing: Science Press, 1998. 1
2 Nelson J S. The Fish of the World. 3rd ed. New York: Wiley, 1995. 1－18
3 Chen X L, Le P Q, Lin R D. Major groups within the family Cyprinidae and their phylogenetic relationships. Acta Zootaxon Sin (in Chinese), 1984, 9(4): 424－440
4 Howes G J. Systematics and biogeography: An overview. In: Winfield I J, Nelson J S, eds. Cyprinid Fishes: Systematics, Biology and Exploitation. London: Chapman and Hall, 1991. 1－33
5 Cavender T, Coburn M. Phylogenetic relationships of North American Cyprinidae. In: Mayden R, ed. Systematics Historical Ecology, and North American Freshwater Fishes. Standford: Standford University Press, 1992. 293－327
6 Briolay J, Galtier N, Brito R M, et al. Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. Mol Phylogenet Evol, 1998, 9(1): 100－108
7 Gilles A, Lecointre G, Faure E, et al. Mitochondrial phylogeny of the European cyprinids: Implications for their systematics, reticulate evolution and colonization time. Mol Phylogenet Evol, 1998, 10(1): 132－143
8 Gilles A, Lecointre G, Miguelis A, et al. Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. Mol Phylogenet Evol, 2001, 19: 22－33
9 Zardoya R. Molecular evidence on the evolutionary and Biogeographical pattern of European cyprinids. J Mol Evol, 1999, 49: 227－237
10 Simons A M, Mayden R L. Phylogenetic relationships of the creek chubs and the spinefins: An enigmatic group of north American cyprinid fishes (Actinopterygii: Cyprinidae). Cladistics, 1997, 13: 187－205
11 Simons A M, Mayden R L. Phylogenetic relationships of the western North American phoxinins (Actinopterygii: Cyprinidae) as inferred from mitochondrial 12S and 16S ribosomal RNA sequences. Mol Phylogenet Evol, 1998, 9: 308－329
12 Simons A M, Mayden R L. Phylogenetic relationships of North American cyprinids and assessment of homology of the openposterior myodome. Copeia, 1999, 1: 13－21
13 Xiao W, Zhang Y, Liu H. Molecular systematics of xenocyprinae (Teleostei: Cyprinidae): Taxonomy, biogeography and coevolution of a special group restricted in East Asia. Mol Phylogenet Evol, 2001, 18: 163－173
14 He S, Chen Y, Nakajima T. Sequences of cytochrome *b* gene for primitive cyprinid fishes in East Asia and their phylogenetic concerning. Chin Sci Bull, 2001,46(8): 661－665
15 He S, Liu H, Chen Y, et al. Molecular phylogenetic relationships of Eastern Asian Cyprinidae (Pisces: Cyprinidformes) inferred from cytochrome *b* sequences. Sci China Ser C-Life Sci, 2004, 47(2): 130－138
16 Wang X, He S, Chen Y. Sequence variations of the S7 ribosomal protein gene in primitive cyprinid fishes: Implication on phylogenetic analysis. Chin Sci Bull, 2002, 47(19): 1638－1643.
17 Wang X, Liu H, He S, et al. Sequence analysis of cytochrome *b* gene indicated that East Asian group of cyprinid subfamily Leuciscinae (Teleostei: Cyprinidae) evolved independently. Prog Natl Sci, 2004, 14(2): 132－137
18 Atchley W R, Fitch W M. Myc and Max: Molecular evolution of a family of proto-oncogene products and their dimerization partner. Proc Natl Acad Sci USA, 1995, 92: 10217－10221
19 Marcu K B, Bossone S A, Patel A J. Myc function and regulation. Annu Rev Biochem, 1992, 61: 809－861
20 Johnston L A, Prober D A, Edgar B A, et al. Drosophila myc regulates cellular growth during development. Cell, 1999, 98: 779－790
21 Trumpp A, Refaeli Y, Oskarsson T, et al. *c-myc* regulates mammalian body size by controlling cell number but not cell size. Nature, 2001, 414: 768－773
22 Schreiber-Agus N, Stein D, Chen K, et al. *Drosophila* Myc is oncogenic in mammalian cells and plays a role in the diminutive phenotype. Proc Natl Acad Sci USA, 1997, 94: 1235－1240
23 Prendergast G C. *Myc* structure and function. In: Yaniv M, Ghysdael J, eds. Oncogenes As Transcriptional Regulators. Vol 1. Basel: Birkhauser Verlag, 1997. 1－28
24 Steppan S J, Storz B L, Hoffmann R S. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from *c-myc* and RAG1. Mol Phylogenet Evol, 2004, 30: 703－719
25 Irestedt M, Johansson U S, Parsons T J, et al. Phylogeny of major lineages of suboscines (Passeriformes) analysed by nuclear DNA sequence data. J Avian Biol, 2001, 32: 15－25
26 Harshman J, Huddleston C J, Bollback J P, et al. True and false gavials: A nuclear gene phylogeny of *Crocodylia*. Syst Biol, 2003, 52: 386－402
27 Crawford A J, Smith E N. Cenozoic biogeography and evolution in direct-developing frogs of Central America (Leptodactylidae: Eleutherodactylus) as inferred from a phylogenetic analysis of nuclear and mitochondrial genes. Mol Phylogenet Evol, 2005, 35: 536－555
28 Sambrook J, Fritsch E, Maniatis T. Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 1989
29 Thompson J D, Gibson T J, Plewniak F, et al. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res, 1997, 25: 4876－4882
30 Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform, 2004, 5: 150－163
31 Xia X, Xie Z. DAMBE: Data analysis in molecular biology and evolution. J Hered, 2001, 92: 371－373
32 Posada D, Crandall K A. Modeltest: Testing the model of DNA substitution. Bioinformatics, 1998, 14(9): 817－818
33 Swofford D L. PAUP: Phylogenetic analysis using Parsimony, Version4. Sunderland: Sinauer Associates, 2002.
34 Huelsenbeck J P, Ronquist F. MRBAYES: Bayesian inference of phylogeny. Bioinformatics, 2001, 17: 754－755
35 Ronquist F, Huelsenbeck J P. MRBAYES 3: Bayesian phylogenetic inference under mixed model. Bioinformatics, 2003, 19:1572－1574
36 Schreiber-agus N, Horner J, Torres R, et al. Zebra fish *myc* family and max genes: Differential expression and oncogenic activity throughout vertebrate evolution. Mol Cell Biol, 1993, 13: 2765－2277