

鲤科鱼类的 *c-myc* CDS 及其系统发育关系

孔祥会^{①②} 王绪祯^① 甘小妮^① 何舜平^{①*}

(^① 中国科学院水生生物研究所, 武汉 430072; ^② 河南师范大学生命科学学院, 新乡 453007. * 联系人, E-mail: clad@ihb.ac.cn)

摘要 鲤科鱼类在东亚的物种多样, 分布广泛, 物种特征尺寸差异明显, 弄清其功能基因的系统演变, 对于理解物种分化和功能进化具有重要意义. 以具有重要生长调控作用的 *c-myc* 基因为标记, 通过 PCR 扩增、克隆和测序, 共获得 41 种鲤科鱼类和外类群 *c-myc* 基因全序列, 发现并分析了 *c-myc* 编码区的两个高变异区. 基于 *c-myc* CDS 序列, 分别采用最大简约法(MP)、最大似然法(ML)和 Bayesian 法重建了鲤科鱼类的系统发育关系. 3 种方法所得系统发育关系较为相似. 当以亚口鱼科的胭脂鱼(*Myxocyprinus asiaticus*)、鳅科的泥鳅(*Misgurnus anguillicaudatus*)和平鳍鳅科的中华间吸鳅(*Hemimyzon sinensis*)作为外类群, 3 种方法重建的系统分支图均以较高的节点支持率支持鲤科鱼类中雅罗鱼系和鲃系的划分, 雅罗鱼系包括雅罗鱼亚科(Leuciscinae)的东亚种类、鲴亚科(Xenocyprinae)、鲃亚科(Cultrinae)、鲃亚科(Danioninae)东亚土著种类、鲃亚科(Gobioninae)和鲃亚科(Acheilognathinae). 鲃系包括裂腹鱼亚科(Schizothoracinae)、鲃亚科(Barbinae)、鲤亚科(Cyprininae)和野鲮亚科(Labeoninae). 斑马鱼(*Danio rerio*)、麦氏鲃(*D. myersi*)和三线波鱼(*Rasbora trilineata*)应从雅罗鱼系或鲃系中分出, 形成另外的一个系群. 鲃亚科是一个多起源的复合类群, 其中一些种类归属存在较多问题, 对鲃亚科一些种类有必要重新划分. *c-myc* CDS 翻译成蛋白质后, 逐个分析具有简约信息氨基酸变异发现氨基酸位点变异与鱼类特征尺寸之间无相关性. 同时分析并发现 *c-myc* 编码区内两个高变异区序列变异既不能体现物种分化关系, 也与物种特征尺寸无相关性. 本研究还发现每一类群位于分支图基部较为原始的种类一般鱼类物种尺寸较小, 可能是对原始生态和生存压力适应的结果.

关键词 鲤科鱼类 *c-myc* 基因 系统发育 基因变异

鲤科是现生鱼类中种类最多的一个科, 约 210 属 2010 种^[1], 广泛分布于欧亚大陆、东印度群岛、非洲和北美^[2]. 东亚是鲤科鱼类物种多样性最为丰富的地区之一, 其中中国特有种 384 种^[1]. 自从 1817 年 Cuvier 建立鲤科以来, 基于形态学特征的鲤科鱼类分类就不断地得以修订. 陈湘舜等人^[3]把鲤科鱼类分为 10 个亚科, 并分为两个支系: 雅罗鱼系(Leuciscini)和鲃系(Barbini), 并将鳅鲃亚科(Gobiobotinae)和裂腹鱼亚科(Schizothoracinae)分别并入了鲃亚科(Gobioninae)和鲃亚科(Barbinae). 基于胡须分布、形态类型和神经分布, Howes^[4]把鲤科鱼类分为鲤亚科(Cyprininae)、鲃亚科、波鱼亚科(Rasborinae)(没有或偶然具有胡须)、雅罗鱼亚科(Leuciscinae)、鲃亚科(Acheilognathinae)、鲃亚科(Cultrinae)和 Alburninae 亚科. 陈湘舜等人在鲃亚科(Danioninae)、雅罗鱼亚科和鲃亚科(Xenocyprinae)的处理上, 同 Howes 处理很不相同.

Cavender 和 Coburn^[5]在陈湘舜等人工作的基础上对鲤科鱼类进行了支序分析, 其结果与陈湘舜等人的划分大致相同. 陈宜瑜^[1]根据形态学特征将鲤科鱼类划分为 12 个亚科, 并强调了东亚鲤科鱼类的系统发育地位和进化上的意义.

目前, 鲤科鱼类的分子系统学研究已开展了部分工作. 基于线粒体的细胞色素 *b* 基因序列和线粒体 16S rRNA 基因部分序列、线粒体控制区 D-loop 序列重建了欧洲鲤科鱼类的系统发育关系^[6-9]; 基于线粒体 12S rRNA 和 16S rRNA 基因序列构建了北美鲤科鱼类系统发育关系^[10-12]; 线粒体中的细胞色素 *b* 和 ND4 序列也被用来分析鲃亚科的系统发育关系^[13]. 何舜平等^[14-17]曾利用细胞色素 *b* 基因序列对东亚鲤科鱼类开展了研究, 并利用细胞色素 *b* 和 S7 核糖体蛋白基因第 1 内含子序列研究了东亚低等鲤科鱼类的系统发生. 目前, 鲤科鱼类的系统发育关系虽

2006-12-26 收稿, 2007-03-19 接受

国家重点基础研究发展计划(批准号: 2004CB117402)、国家自然科学基金重点项目(批准号: 30530120)和中国博士后科学基金(批准号: 2005037684)资助

然还有许多争议的问题,但分子系统学研究使鲤科鱼类的系统发育框架越来越清晰。这些研究主要是基于进化速率较快的线粒体基因和内含子序列。利用核基因的研究还相对较少。利用有重要生物学意义的核功能基因进行构建鲤科鱼类的系统发育关系既是鲤科鱼类系统发育研究的重要内容,同时也为研究鲤科鱼类物种间的基因变异和功能分歧提供资料。

c-myc (cellular myelocytomatosis oncogene)是 *myc* 家族中重要的一员,在胚胎或成体组织的所有增生细胞中均有表达,是一类重要的转录激活因子,对动物生长发育具有重要的调控功能。*c-myc* 基因的结构和核苷酸组成非常保守,一般由 3 个外显子和 2 个内含子组成。它在哺乳动物、鸟类、两栖类和鱼类中广泛存在,估计至少有 6 亿年的演化历史^[18]。*c-myc* 基因是一个单拷贝基因,多肽可读框在第 2 和第 3 外显子^[19]。当 *c-myc* 基因发生突变时可使果蝇和老鼠个体尺寸变小,主要是通过调控细胞尺寸^[20]或细胞数目^[21]来调控个体尺寸。尽管果蝇和鼠中 *c-myc* 基因核苷酸序列同源性较低,但蛋白质序列上仍然有较高的同源性,具有相似的生物学功能,并且可以用果蝇 *c-myc* 部分弥补鼠 *c-myc* 表达量的不足^[22]。动物体中 *c-myc* 基因突变可以改变个体最终尺寸,同时它在肿瘤发生中具有重要作用^[23]。近年来,在研究哺乳动物^[24]、雀形目鸟类^[25]、鳄鱼^[26]和蛙类^[27]的系统发育关系时发现,*c-myc* 基因已作为目和科级等高阶元系统发育研究的分子标记。基于 *c-myc* 基因的鱼类系统发育分析还未见报道。

鲤科鱼类特征尺寸悬殊,食性各异,生活在不同的生态位,是研究物种分歧和系统发育的绝好材料。本研究在东亚鲤科鱼类中选取 38 种代表性物种,通过克隆和测序获得它们的 *c-myc* 基因全序列,对编码区的两个高变异区并进行了分析,并以编码区序列重建了鲤科鱼类的系统发育关系。目的是基于 *c-myc* 编码 DNA 序列(coding DNA sequence, CDS)重建鲤科鱼类的系统发育关系,并与线粒体基因重建的鲤科系统发育关系进行比较,探讨 *c-myc* CDS 变异与鲤科鱼类物种特征尺寸之间的相关性。

1 材料与方法

(i) 样品来源。本研究所用鲤科鱼类主要来自中国科学院水生生物博物馆的馆藏标本。所用标本

的采集地及标本号见表 1。外类群选择了亚口鱼科的胭脂鱼(*Myxocyprinus asiaticus*), 鳅科的泥鳅(*Misgurnus anguillicaudatus*), 平鳍鳅科的中华间吸鳅(*Hemimyzon sinensis*)。用于分子生物学实验的鱼类标本或组织均是经过梯度酒精脱水,保存于 95%酒精中。

(ii) 引物设计。根据从 GenBank 获取的斑马鱼(*Danio rerio*)(登录号: NM_131412)、鲤(*Cyprinus Carpio*) (D37888)和鲫 (*Carassius auratus*) (D31729) 的 *c-myc* 基因的保守序列,经过设计和优化后,最后确定 PCR 扩增所用的引物序列为: PF6, 5'-ATYAGTC-TGTCCAGCAYCT-3'; PR7, 5'-SRAACTCGCTGACCA-TCTC-3'; PF8, 5'-KSGTTGWTAYATTTTCCATCAC-3'; PR8, 5'-GRAACTCGSTSACYATCTC-3'; PF11, 5'-AATGCYGGTGAGTKCGAGTT-3'; PR11, 5'-GCTG-MAGCYTGTGTTTAACTGT -3'; PF12, 5'-TGGAG-ATRGTSAGCGAGTT-3'。引物在 *c-myc* 基因上的分布见图 1。

(iii) DNA 提取、扩增和测序。DNA 提取采用标准的酚/氯仿/异戊醇法^[28], DNA 的浓度及纯度测定采用 Beckman 紫外分析仪。DNA 浓度稀释并控制在 30~50 ng/ μ L 范围内。PCR 扩增反应各试剂(TaKaRa)在用前按比例配制。60 μ L PCR 反应液包括: 10 $\times$ Ex Taq 缓冲液 6 μ L, 2.5 mmol/L 4 种 dNTP 混合液(pH 8.0) 4.8 μ L, 15 μ mol/L 正、反向引物各 1.5 μ L, 5 U/ μ L Ex Taq 酶 0.6 μ L, dd H₂O 42.6 μ L, 模板 DNA 3 μ L。

利用引物 PF11 和 PR11 及 PF12 和 PR11 两对引物扩增时,反应条件为:反应体系 94 $^{\circ}$ C 预变性 4 min; 94 $^{\circ}$ C 变性 50 s, 57 $^{\circ}$ C(PF11 和 PR11)或 56 $^{\circ}$ C(PF12 和 PR11)复性 50 s, 72 $^{\circ}$ C 延伸 90 s, 32 个循环; 72 $^{\circ}$ C 延伸 6 min。利用引物 PF6 和 PR7, PF8 和 PR8 两对引物扩增时,反应条件为:反应体系 94 $^{\circ}$ C 预变性 4 min; 94 $^{\circ}$ C 变性 30 s, 55 $^{\circ}$ C(PF6 和 PR7)或 50 $^{\circ}$ C(PF8 和 PR8)复性 30 s, 72 $^{\circ}$ C 延伸 60 s, 32 个循环; 72 $^{\circ}$ C 延伸 5 min。

PCR 产物进行 1.2% 的琼脂糖电泳,然后准确割胶回收目的片段,用 OMEGA 试剂盒(OMEGA Bio-Tek)进行割胶回收。然后以 pMD18-T 载体试剂盒(TaKaRa)为连接载体,以 DH5 α 作为宿主菌,对目的片段进行克隆。目的片段的阳性克隆在 ABI3730 测序仪上进行测序,同一样品测 3 个克隆。

(iv) 基因特征分析。将获得的 *c-myc* 序列(GenBank 登录号见表 1)用 CLUSTALX (1.83)^[29]进行多重排定,设置参数 gap-opening penalty = 15, gap-extension penalty = 3。序列排好后辅以手工校正。

表 1 鲤科鱼类和外类群标本种类

鱼类物种	采集地	标本号	接受号
鳙 <i>Aristichthys nobilis</i>	湖北武汉	IHBCYK0411001	EF194848
鲢 <i>Hypophthalmichthys molitrix</i>	湖北武汉	IHBCYK0411002	EF194849
草鱼 <i>Ctenopharyngodon idellus</i>	湖北武汉	IHBCYK0411003	EF194850
青鱼 <i>Mylopharyngodon piceus</i>	湖北武汉	IHBCYK0411004	EF194851
鲢 <i>Ochetobius elongates</i>	广西藤县	IHBCY0108003	EF194852
鳊 <i>Elopichthys bambusa</i>	湖南桃源	NRMT2286	EF194853
赤眼鳟 <i>Squaliobarbus curriculus</i>	武汉金口	IHBCY0407001	EF194854
翘嘴鲌 <i>Culter alburnus</i>	武汉金口	IHBCY0380494	EF194855
团头鲂 <i>Megalobrama amblycephala</i>	湖北武汉	IHBCY0305004	EF194856
鳊 <i>Hemiculter leucisculus</i>	湖北武汉	IHBCY2603026	EF194857
似鳊 <i>Pseudobrama simoni</i>	湖南桃源	IHBCY0405361	EF194858
银鲴 <i>Xenocypris argentea</i>	湖南桃源	IHBCY0405138	EF194859
中华细鲫 <i>Aphyocypris chinensis</i>	中国科学院水生生物研究所	IHBCYK0411005	EF194860
马口鱼 <i>Opsariichthys bidens</i>	湖南桃源	NRMT2358	EF194861
细尾蛇鮈 <i>Saurogobio gracilicaudatus</i>	贵州	IHBCY0312012	EF194862
蛇鮈 <i>Saurogobio dabryi</i>	湖北长阳	IHBCY0405136	EF194863
铜鱼 <i>Coreius heterodon</i>	武汉金口	IHBCY0312002	EF194864
麦穗鱼 <i>Pseudorasbora parva</i>	云南勐腊	IHBCY0312003	EF194865
稀有鲫 <i>Gobiocypris rarus</i>	中国科学院水生生物研究所	IHBCYK0411006	EF194866
唐鱼 <i>Tanichthys albonubes</i>	广东	IHBCYK0411007	EF194867
高体鲮 <i>Rhodeus ocellatus</i>	湖北武汉	IHBCYK0411008	EF194868
彩石鲮 <i>Rhodeus lighti</i>	湖北武汉	IHBCYK0411009	EF194869
彩副橘 <i>Paracheilognathus imberbis</i>	湖北武汉	IHBCYK0411010	EF194870
斑马鱼 <i>Danio rerio</i>	湖北武汉	IHBCYK0411011	EF194871
麦氏鲮 <i>Danio myersi</i>	云南勐腊	IHBCY0405411	EF194872
三线波鱼 <i>Rasbora trilineata</i>	湖北武汉	IHBCYK0411012	EF194873
巨须裂腹鱼 <i>Schizothorax longibarbus</i>	西藏	IHBCY0510081	EF194874
异齿裂腹鱼 <i>Schizothorax oconnori</i>	西藏	IHBCY0510086	EF194875
光唇裂腹鱼 <i>Schizothorax lissolabiatus</i>	西藏	IHBCY0504193	EF194876
高原裸鲤 <i>Gymnocypris waddelli</i>	青海	IHBCY0510092	EF194877
条纹小鲃 <i>Puntius semifasciolatus</i>	云南勐腊	IHBCY0405496	EF194879
鲃鲤 <i>Percocypris pingi pingi</i>	四川合江	IHBCY0205010	EF194881
鲫 <i>Carassius auratus</i>	湖北武汉	IHBCYK0411013	EF194880
鲤 <i>Cyprinus Carpio</i>	湖北武汉	IHBCYK0411014	EF194882
中华倒刺鲃 <i>Spinibarbus sinensis</i>	四川南充	IHBCY0207036	EF194883
卷口鱼 <i>Ptychidio jordani</i>	云南	IHBCY0308004	EF194884
角鱼 <i>Epalzeorhynchus bicornis</i>	湖北武汉	IHBCY0505291	EF194885
西藏墨头鱼 <i>Garra kemp</i>	西藏察隅	IHBCY0309091	EF194886
中华间吸鳅 <i>Hemimyzon sinensis</i>	云南	IHBCYK0311012	EF194887
泥鳅 <i>Misgurnus anguillicaudatus</i>	云南大姚县	IHBCYK0411015	EF194888
胭脂鱼 <i>Myxocyprinus asiaticus</i>	湖北武汉	IHBCY0305001	EF194889

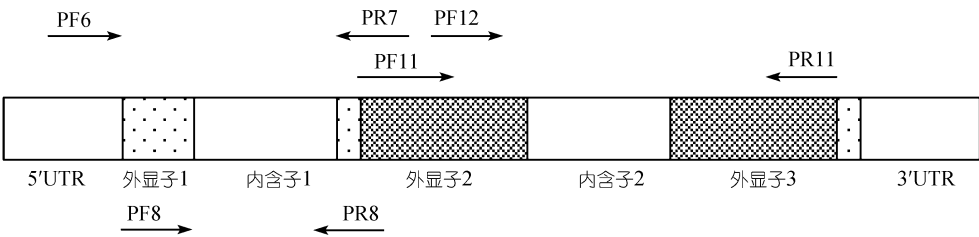


图 1 c-myc 基因结构及扩增引物的分布

c-myc CDS 序列变异、碱基组成及替代用 MEGA 3.1^[30]进行统计。基于转换和颠换与遗传距离的线性关系,采用 DAMBE (V4.1.33)进行碱基替代饱和性分析^[31]。

(v) 系统发育分析。鲤科鱼类的系统发育分析采用最大简约法(MP)、最大似然法(ML)和 Bayesian 法。最大似然分析利用 Modeltest 3.7^[32]进行模型和参数估计,基于 AIC 标准选择的最适碱基替代模型是 GTR + I + G(-lnL = 7760.8984; K = 10; AIC = 15541.7969)。MP 和 ML 分析采用 PAUP4.0b10^[33]。Bayesian 法采用 MrBayes3.1.2 软件分析^[34,35]。

最大简约性分析采用启发式搜索,最终在同等简约树中获得多数一致树。进行 1000 次 bootstrap 分析时参数设定为: nreps = 1000, ConLevel = 50, search = heuristic, brlens = yes。最大似然分析时参数设定为: Base = (0.2846 0.2567 0.2678), Rmat = (0.3476 3.8968 1.0766 0.3006 3.2010), Rates = gamma, Shape = 0.7299, Pinvar = 0.4566。ML 树进行 100 次 bootstrap 检验,分析参数为: nreps = 100, ConLevel = 50, search = heuristic, brlens = yes。Bayesian 分析时,设定参数 nst = 6, rates = invgamma, Ngen = 1000000, Nruns = 2, Nchains = 4, Temp = 0.2, Samplefreq = 100, Printfreq = 1000, Diagnfreq = 1000, Burnin = 600, Burninfrac = 0.25。

(vi) *c-myc* CDS 序列变异分析。排定的 *c-myc* CDS 矩阵,序列翻译成氨基酸序列后,逐个对具有简约信息的氨基酸变异与物种特征尺寸之间相关性进行分析。同时针对编码区中变异较大的两个区域进行分析,碱基缺失以 gap(-)表示,各种鱼类缺失碱基数利用 Excel 2003 统计分析。

2 结果

2.1 *c-myc* 编码区序列特征

41 种 *c-myc* 编码区序列排定后由 1246 bp 组成,其中保守位点 798 个,可变位点 448 个,非简约可变位点 169 个,简约信息位点 279 个。碱基组成 T, C, A 和 G 分别为 18.1%, 26.4%, 28.5%和 27.0%; 转换与颠换之比为 2.2。 *c-myc* CDS 转换和颠换与遗传距离的线性关系显示这些位点替代均没有达到突变饱和。

2.2 系统发育分析

最大简约性分析得到了 111 棵同等简约树,树长是 1177 步; CI (consistency index) = 0.5200; HI (homoplasy index) = 0.4800; RI (Retention index) = 0.6146。多数一致性树见图 2,系统分支图显示以

70%和 95%的节点支持率支持雅罗鱼系和鲃系的划分。雅罗鱼系包括雅罗鱼亚科的东亚类群、鲴亚科、鲃亚科、鮡亚科、鲮亚科和鲴亚科一些种类。鲃系包括裂腹鱼亚科、鲃亚科、鲤亚科和野鲮亚科(Labeoninae)。斑马鱼(*Danio rerio*)和麦氏鲃(*D. myersi*)形成的单系和雅罗鱼系形成姊妹群,节点支持率较低;三线波鱼(*Rasbora trilineata*)位于分支树基部,节点支持率也较低。

最大似然分析利用启发式搜索所得的 ML 树(图 3)和 MP 树拓扑结构大致相同,ML 树以 80%和 96%的节点支持率分别支持鲤科鱼类中雅罗鱼系和鲃系的划分。东亚类群和鮡亚科,唐鱼属形成 3 个系群,共同构成一个大的单系群,与鲮亚科一起形成雅罗鱼系。鲃系形成一个单系群。鲴亚科中的斑马鱼、麦氏鲃和三线波鱼聚为一个单系群,与鲃系互为姊妹群,节点支持率较低。

Bayesian 树(图 4)与 ML 树大致相同,系统分支图以 100%的节点支持率支持鲤科鱼类中雅罗鱼系和鲃系的划分;同时也以 100%的节点支持率支持斑马鱼、麦氏鲃和三线波鱼从雅罗鱼系和鲃系中分出。雅罗鱼系中鲴亚科和鮡亚科形成并系;鲃系中鲃亚科种类没有形成单系群,而是分散于裂腹鱼亚科和鲤亚科,说明鲃亚科的划分仍存在较多问题。

2.3 序列变异分析

c-myc 编码区序列较为保守,但有两个区域变异较大(图 5)。高变异区 1 (variation 1)主要是 GAG (GAA)GAG 的串联重复;高变异区 2 (Variation 2)主要是 AGC(AAC)AGC 的串联重复。不同鱼类之间差异以碱基缺失(gap)进行统计,即使亲缘关系较近的物种,碱基缺失也有较大的差异;同时这两个区域的变异与物种特征尺寸之间无相关关系。另外, CDS 序列翻译蛋白质后,逐个对具有简约信息的氨基酸变异分析发现,氨基酸变异与物种特征尺寸之间也无相关性。

3 讨论

c-myc 基因是生长调控的重要基因,首次应用于鱼类系统发育研究。基于 *c-myc* CDS 重建的分子系统树与以前构建的明晰的鲤科鱼类局部系统分支图^[15-17]较为一致,说明 *c-myc* CDS 序列是一个很好的遗传标记,适合研究鱼类科及亚科水平上的系统发育关系。

本研究中的 MP 树、ML 树和 Bayesian 树均以较

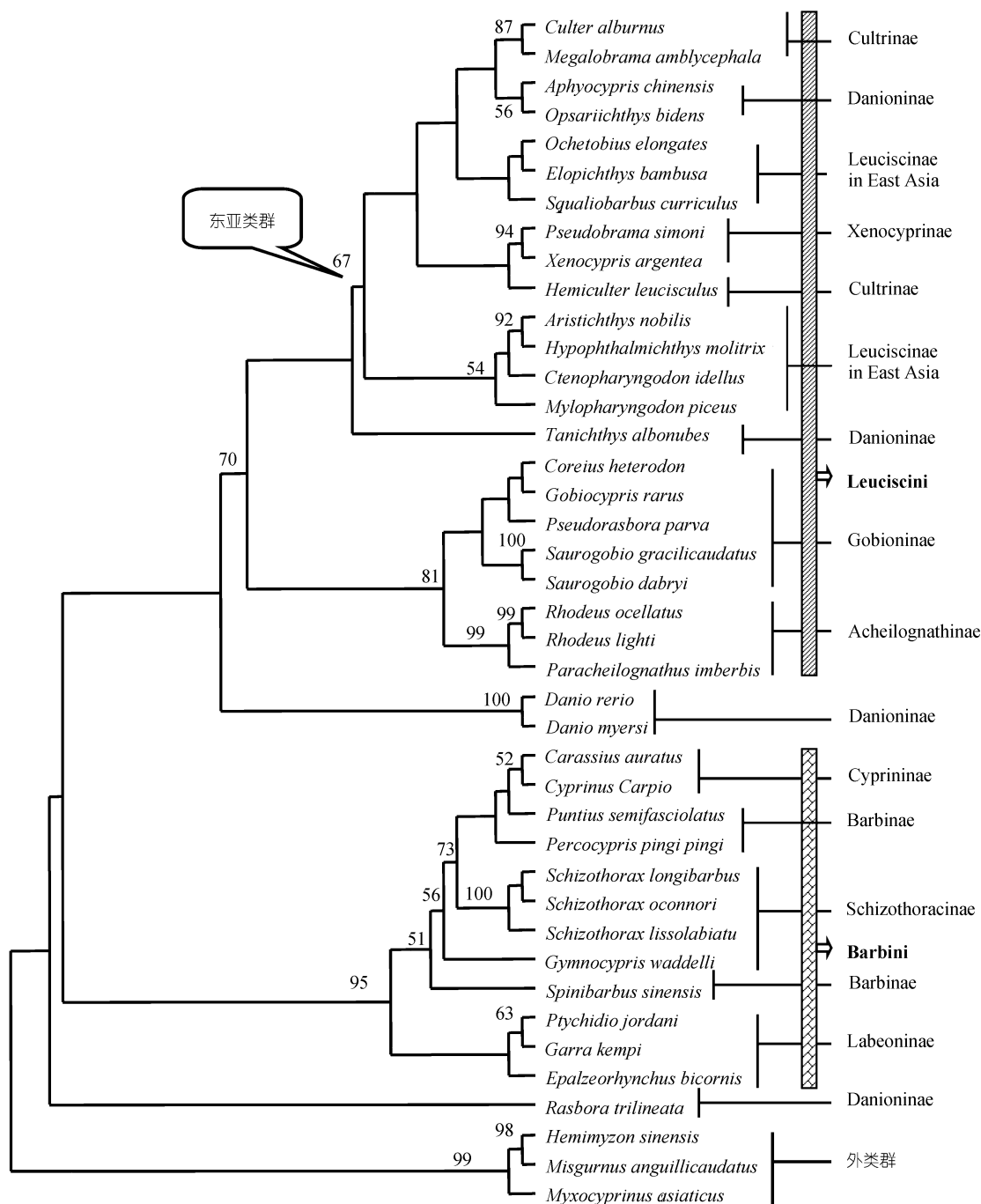


图2 通过启发式搜索(PAUP4.0b10)得到的 111 个同值最大简约的多数一致性树

树长是 1177 步; CI = 0.5200; RI = 0.6146; RC = 0.3196; 位于分支节点处的数值为大于 50% 的 1000 次 Bootstrap 重抽样支持率

高的节点支持率支持鲤科鱼类中雅罗鱼系和鲃系的划分, 基本支持陈湘舜等人^[3]基于形态以及何舜平等^[15]基于细胞色素 *b* 的两大类群的划分。雅罗鱼系包括东亚类群、鲃亚科和鲤亚科鱼类。鲃系包括裂腹鱼亚科、鲃亚科、鲤亚科和野鲮亚科。此外, 这些系统分支图显示斑马鱼、麦氏鲃和三线波鱼从雅罗鱼系

和鲃系中分出时, 其节点支持率较高; 与雅罗鱼系和鲃系聚在一起时, 其节点支持率较低, 所以本研究提议斑马鱼、麦氏鲃和三线波鱼从雅罗鱼系和鲃系中分出, 单独形成一个系群。何舜平等^[15]所定义的东亚类群, 包括鲃亚科、鲃亚科、东亚雅罗鱼亚科和鲃亚科东亚土著种类。本研究中, MP, ML 和 Bayesian 树

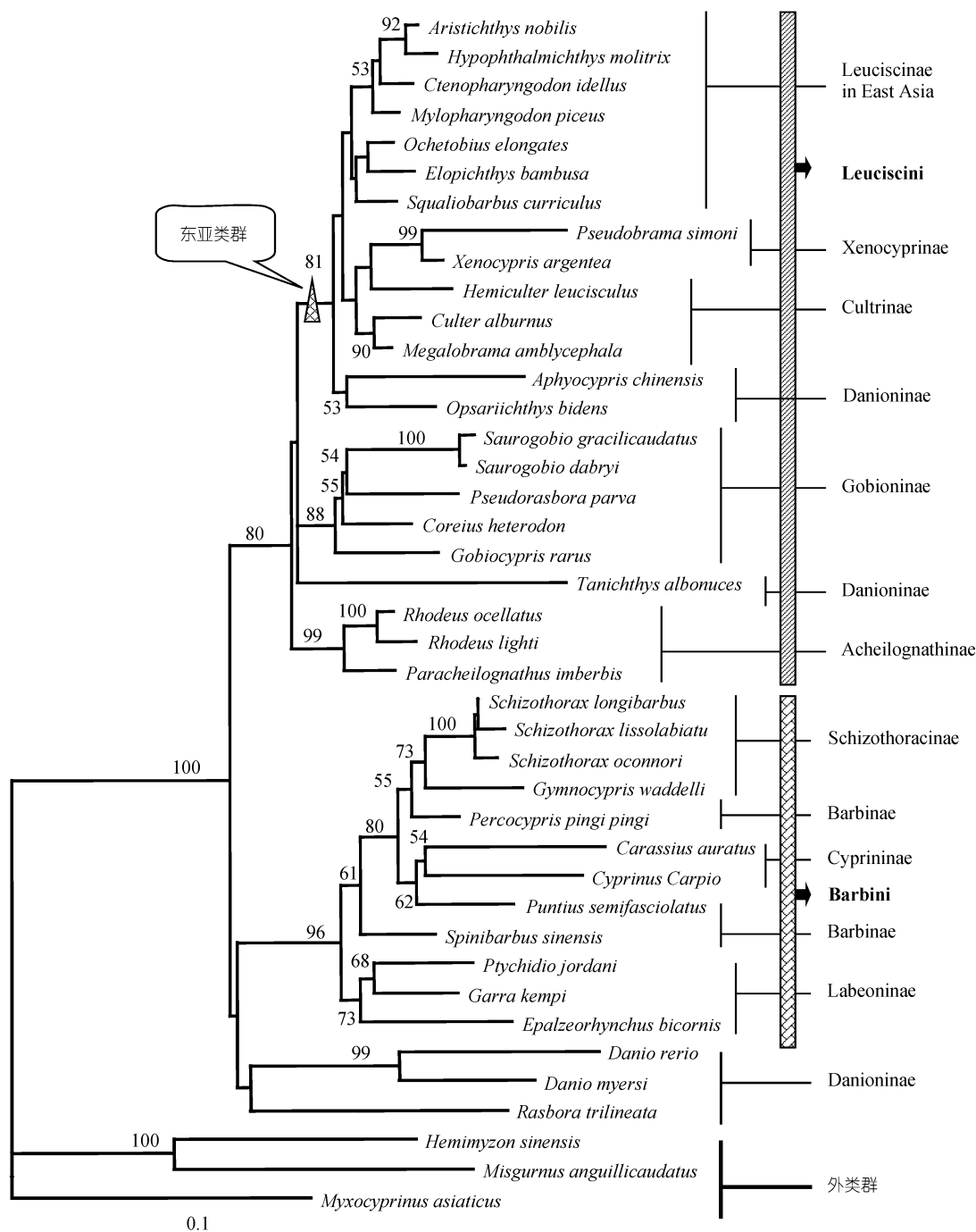


图 3 通过后发式搜索重建的最大似然树

分析基于 GTR+I+G 模型. 位于分支节点处的数值为大于 50% 的 100 次 Bootstrap 重抽样支持率

中均以较高的节点支持率支持东亚类群形成一个单系类群. 雅罗鱼系中鲃亚科和鲴亚科各形成单系群.

基于形态学特征, 鲃亚科被认为是多类群的复合体^[4]. 本研究也显示鲃亚科不是一个单系类群. 中华细鲫(*Aphyocypris chinensis*)和马口鱼(*Opsari-*

ichthys bidens)属于东亚类群, 而唐鱼(*Tanichthys albonubes*)在 MP 树中并入东亚类群, 在 ML 和 Bayesian 树中与东亚类群分开, 但这 3 种鱼类仍然属于雅罗鱼系. 斑马鱼、麦氏鲃始终聚在一起, 在 MP 树中与三线波鱼分开, 以较低的节点支持率与雅罗

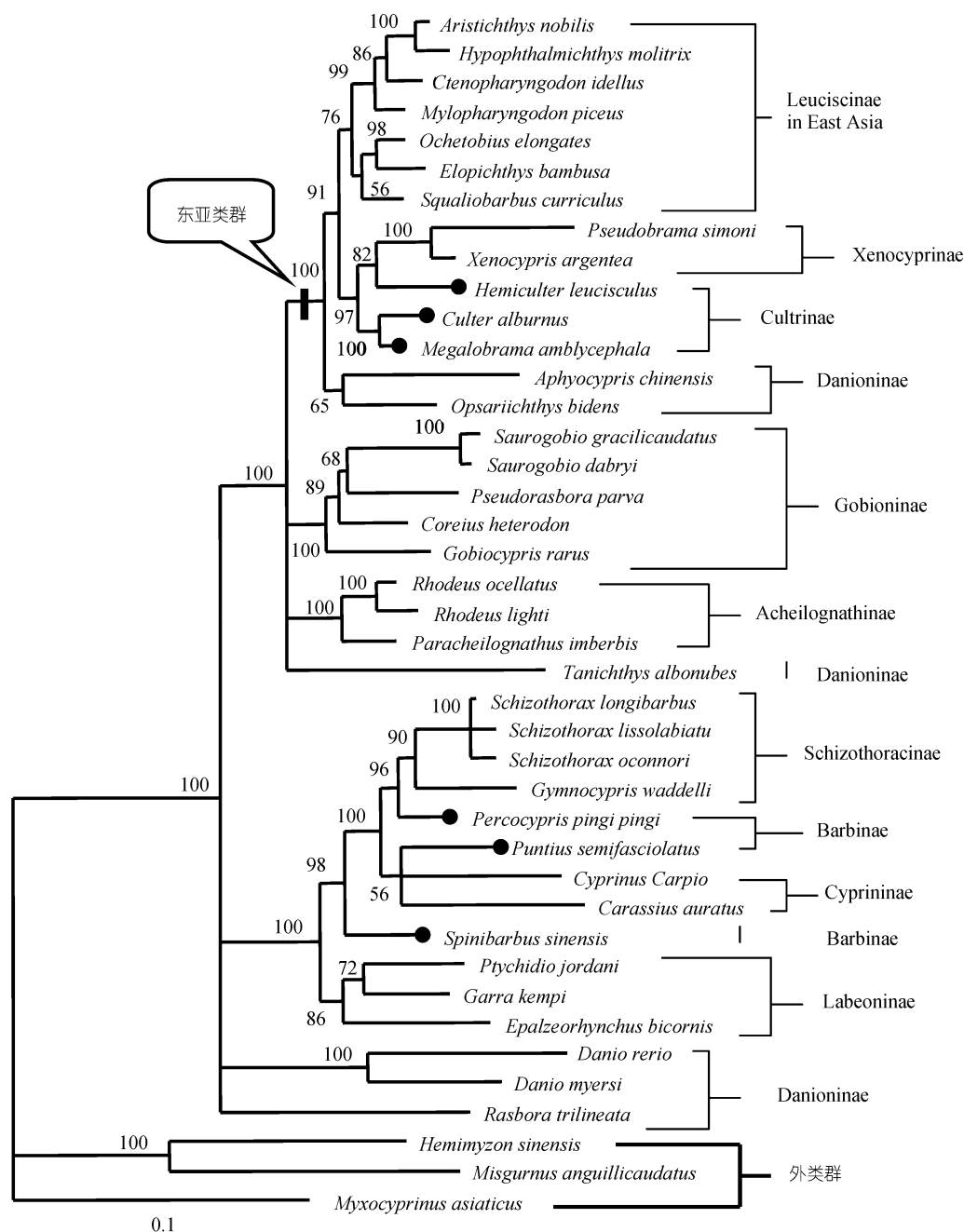


图4 Bayesian法构建的鲤科鱼类系统发育树
分支节点处的数值为大于50%的后验概率

鱼系形成姊妹群；在 ML 树中，斑马鱼、麦氏鲃与三线波鱼聚为单系，以较低的节点支持率与鲃系形成姊妹群。Bayesian 树则以 100% 的节点支持率支持斑马鱼、麦氏鲃和三线波鱼与雅罗鱼系和鲃系分开。因此，基于 *c-myc* CDS 的系统发育分析更倾向于鲤科鱼类划分为雅罗鱼系、鲃系和斑马鱼、麦氏鲃与三线波

鱼形成的系群。虽然鲤科鱼类基于线粒体基因系统发育研究遗留的问题，如唐鱼、斑马鱼、麦氏鲃和三线波鱼的归属问题，在本研究中也并没有得以很好的解决，但给出了解决的方向。引起争议较多的稀有鲫 (*Gobiocypris rarus*) (原属鲃亚科) 在本研究中以较高的节点支持率聚在鲫亚科。

	673	变异区1	711 909	变异区2	945
<i>Aristichthys nobilis</i>	GAGGAAGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAGCAGCAAC	-----AGGC
<i>Hypophthalmichthys molitrix</i>	TGGGAAGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAGCAAC	-----AGGC
<i>Ctenopharyngodon idellus</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAGCAAC	-----AGGC
<i>Mylopharyngodon piceus</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAGCAAC	-----AGGC
<i>Ochetobius elongates</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAGCAAC	-----AGGC
<i>Elopichthys bambusa</i>	GAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAAC	-----AGGC
<i>Squaliobarbus curriculus</i>	GAGGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAAC	-----AGGC
<i>Culter alburnus</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAGCAGCAAC	-----AGGC
<i>Megalobrama amblycephala</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAACAGCAGCAGCAAC	-----AGGC
<i>Hemiculter leucisculus</i>	GAAGAGGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAACAGCAGCAGCAAC	-----AGGC
<i>Pseudobrama simoni</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAACAGCAGCAGCAAC	-----AGGC
<i>Nenobrama argentea</i>	GAGGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAGCAGCAGCAGCAGCAAC	-----AGGC
<i>Aphyocypris chinensis</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	TAACAGCAGCAGCAAC	-----AGAC
<i>Opsariichthys bidens</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAGG	CAACAGCAGCAGCAACGCAAC	-----AGGC
<i>Saurogobio gracilicaudatus</i>	GAAGAGGAGGAGGAGGAGGAGGAAGAAGAA	-----	GAGG	CAGCAGCAAC	-----AGGC
<i>Saurogobio dabryi</i>	GAGGAGGAGGGGAGGAGGAGGAGGAAGAAGAA	-----	GAGG	CAGCAGCAAC	-----AGGC
<i>Coreius heterodon</i>	GAGGAGGAGGAGGAAGAAGAAGAG	-----	GAGG	CAGCAGCAGCAGCAAC	-----AGGC
<i>Pseudorasbora parva</i>	GAGGAGGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAGG	CAGCGGAGCGGAGCAGCAGCAGC	-----CGGC
<i>Gobiocypris rarus</i>	GAGGAAGAAGAAGAGGAAGAAGAG	-----	GAGG	CAGCAGCAGCAAC	-----AGGC
<i>Tanichthys albonubes</i>	GAAGAGGAGGAGGAGGAAGAAGAAGAA	-----	GAAG	CAGCAACAGCAGCAGCAAC	-----AGGC
<i>Rhodeus ocellatus</i>	GAGGAGGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAAC	-----AGAC
<i>Rhodeus lighti</i>	GAAGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAACAACAGCAACAGCAGCAGCAAC	-----AGAC
<i>Paracheilognathus imberbis</i>	GAGGAGGAGGAAGAAGAG	-----	GAAG	CAGCAGCAAC	-----AGGC
<i>Danio rerio</i>	GAGGAGGAAGAAGAGGAGGAAGAAGAGGAG	-----	GAAG	TCACAGCATCAACAGCAGCAGCAGCAGCAAC	-----AGGC
<i>Danio myersi</i>	GAAGAAGAGGAAGAGGAGGAGGAAGAAGAG	-----	GAAG	CAACAACAGCAGTAGCAGCAAC	-----AGGC
<i>Rasbora trilineata</i>	GAAGAGGAGGAGGAAGAAGAAGAT	-----	GAAG	CACCAACATCAGCAGCAAC	-----AGGC
<i>Schizothorax longibarbus</i>	GAGGAGGAGGAGGAAGAAGAAGAGGAGGAAGAAGAAGAGGAAG	-----	GAAG	CAACAACAGCAGCAGCAAGCAAC	-----AGGC
<i>Schizothorax oconnori</i>	GAGGAGGAAGAAGAAGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAAGCAAC	-----AGGC
<i>Schizothorax lissolabiatus</i>	GAGGAGGAAGAAGAAGAGGAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAAGCAAC	-----AGGC
<i>Gymnocypris waddelli</i>	GAGGAGGAGGAGGAAGAAGAAGAAGAG	-----	GAAG	CAGCAGGAAC	-----AGGC
<i>Puntius semifasciolatus</i>	GAAGAAGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAACAACAGCAGCAGCAGCAAC	-----AGGC
<i>Percocypris pingi pingi</i>	GAGGAGGAAGAAGAAGAG	-----	GAAG	CAACAGCAGCAAC	-----AGGC
<i>Carassius auratus</i>	GAGGAAGAAGAGGAGGAGGAAGAAGAA	-----	GAAG	CAGCAAC	-----AGGC
<i>Cyprinus Carpio</i>	GAAGAAGAAGAGGAAGAAGAAGAG	-----	GAAG	CAACAACAGCAGCAGCAAC	-----AGGC
<i>Spinibarbus sinensis</i>	GAGGATGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAGCAGCAGCAGCAGCAAC	-----AGGC
<i>Ptychidio jordani</i>	GAGGAAGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAGCAGCAGCAGCATCAACAGCAGCAGCAACAGGC	-----AGGC
<i>Epalzeorhynchus bicornis</i>	GAGGAAGAGGAGGATGAGGAAGAAGAAGAG	-----	GAAG	CAACCAACAGCAGCAGCAAC	-----AGGC
<i>Garra kempfi</i>	GAAGAGGAGGAAGAAGAAGAG	-----	GAAG	CAGCAGCAGCAGCAGCAGCAAC	-----AGGC
<i>Hemimyzon sinensis</i>	GAAGAAGAAGAAGAAGAAGAGGAG	-----	GAAG	GGTCAGCAACAGCAGCAAC	-----AGGC
<i>Misgurnus anguillicaudatus</i>	GAAGAGGAGGAGGAGGAAGAAGAA	-----	GAAG	GATCAGCAGCAGC	-----AGGC
<i>Myxocyprinus asiaticus</i>	GAG	-----	GAAG	TATCAGTAACAGCAAC	-----AGGC

图 5 *c-myc* CDS 编码区的两个高变异区

Bayesian 树中雅罗鱼系的鲴亚科和鲃亚科形成并系, 其中鲮(*Hemiculter leucisculus*)与鲴亚科聚在一起, 说明鲮与鲴亚科关系较近; MP 和 ML 中也有相同结果, 但节点支持率较低。鲃系中以 98% 的节点支持率支持鲃亚科、鲤亚科和裂腹鱼亚科共同形成单系群, 其中鲃亚科种类分散其中, 有的与裂腹鱼亚科聚为一类, 有的与鲤亚科聚为一类, 说明鲃亚科种类存在较多问题。

c-myc 基因不仅可以调控细胞的生长、发育、分化和凋亡^[18], 而且对动物个体尺寸有一定调控作用^[21]。本研究通过翻译成蛋白质后, 逐个分析具有简约信息的氨基酸变异与体型大小关系, 也未发现氨基酸变异与鱼类物种特征尺寸之间具有相关性。尽管 *c-myc* CDS 编码区非常保守, 但本研究发现编码区内有 2 个区域变异非常高。第 1 个区域主要是 GAG(GAA)GAG 的串联重复, 主要编码谷氨酸 (glutamic acid, Glu & E); 第 2 个区主要是 AGC(AAC)AGC 的串联重复, 主要编码丝氨酸 (serine,

Ser & S)和天冬酰胺 (asparagine, Asn & N)。这 2 个区域是酪氨酸激酶 II 磷酸化位点 (casein kinase II phosphorylation sites, CK-II)^[36], 对于 *c-myc* 活性发挥有调控作用。这 2 个变异区内碱基插入或缺失数目在亲缘关系很近的物种也有较大的差异, 不符合系统演化, 在研究鱼类系统发育关系时可以去除这 2 个高变异区; 另一方面在鱼类特征尺寸较小和较大的类群中也没有表现出相关性, 说明这 2 个高变异区不能用于解释物种分歧和鱼类尺寸控制问题, 其功能可能与蛋白折叠后活性的发挥有关, 这还需要通过研究该区域碱基剔除或增加对蛋白质功能的影响来加以确定。

本研究中基于 *c-myc* CDS 序列的系统发育分析显示个体尺寸较小的物种 (如鲃亚科、鲴亚科和野鲮亚科) 并没有聚在一起, 因此基于 *c-myc* CDS 序列的鲤科鱼类聚类分析对鱼类物种特征尺寸差异无法给出合理阐释。从分子系统树来看, 一个类群中特征尺寸较小的种类一般聚在基部, 说明物种特征尺寸

较小的种类一般为较原始的类群。这可以从物种进化适应上加以解释,较早出现的物种,由于生存环境和食物资源的压力较大,物种一般采用繁殖量大,个体尺寸小等进化策略使物种得以保存。所以较早出现的鲴亚科、鲮亚科和野鲮亚科等原始鱼类一般较小。东亚类群中后来分化出来的东亚雅罗鱼是在江湖生态系统食物资源丰富的条件下分化形成,具有充足的食物资源和良好的生态条件,一般个体较大。这种现象可以部分理解鱼类物种特征尺寸差异的原因,但还有待于更多的实验证据加以证明。

致谢 中国科学院水生生物研究所李俊兵博士对样品采集及鉴定给予的帮助,韩亚伟博士在实验上给予的帮助,侯玉婷参与了部分实验工作,在此一并表示感谢。

参 考 文 献

- 1 陈宜瑜. 中国动物志, 硬骨鱼纲, 鲤形目(中卷). 北京: 科学出版社, 1998. 1
- 2 Nelson J S. The Fish of the World. 3rd ed. New York: Wiley, 1995. 1—18
- 3 陈湘琳, 乐佩琦, 林人端. 鲤科的科下类群及宗系发生关系. 动物分类学报, 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[DOI]
- 5 Cavender T, Coburn M. Phylogenetic relationships of North American Cyprinidae. In: Mayden R, ed. Systematics Historical Ecology, and North American Freshwater Fishes. Stanford: Stanford 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[DOI]
- 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[DOI]
- 9 Zardoya R. Molecular evidence on the evolutionary and biogeographical pattern of European cyprinids. J Mol Evol, 1999, 49: 227—237[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 14 何舜平, 陈宜瑜, Nakajima T. 东亚低等鲤科鱼类细胞色素 *b* 基因序列测定及系统发育. 科学通报, 2000, 45(21): 2297—2302
- 15 何舜平, 刘焕章, 陈宜瑜, 等. 基于细胞色素 *b* 基因序列的鲤科鱼类系统发育研究(鱼纲: 鲤形目). 中国科学C辑: 生命科学, 2004, 34(1): 96—104
- 16 王绪桢, 何舜平, 陈宜瑜. S7 核糖体蛋白基因序列变异及其在低等鲤科鱼类中的系统发育意义. 科学通报, 2002, 47(14): 1089—1094
- 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[DOI]
- 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[DOI]
- 19 Marcu K B, Bossone S A, Patel A J. Myc function and regulation. Annu Rev Biochem, 1992, 61: 809—861[DOI]
- 20 Johnston L A, Prober D A, Edgar B A, et al. *Drosophila* myc regulates cellular growth during development. Cell, 1999, 98: 779—790[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 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[DOI]
- 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 Clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res, 1997, 25: 4876—4882[DOI]
- 30 Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform, 2004, 5: 150—163[DOI]
- 31 Xia X, Xie Z. DAMBE: Data analysis in molecular biology and evolution. J Hered, 2001, 92: 371—373[DOI]
- 32 Posada D, Crandall K A. Modeltest: Testing the model of DNA substitution. Bioinformatics, 1998, 14(9): 817—818[DOI]
- 33 Swofford D L. PAUP: Phylogenetic analysis using Parsimony, Version 4. Sinauer Associates. Sunderland, Massachusetts, 2002
- 34 Huelsenbeck J P, Ronquist F. MRBAYES: Bayesian inference of phylogeny. Bioinformatics, 2001, 17: 754—755[DOI]
- 35 Ronquist F, Huelsenbeck J P. MRBAYES 3: Bayesian phylogenetic inference under mixed model. Bioinformatics, 2003, 19: 1572—1574[DOI]
- 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—2775