沙鳅亚科鱼类线粒体 DNA 控制区结构分析及系统发育关系的研究

唐琼英[1,2]　刘焕章[1]　杨秀平[2]　熊邦喜[2]
(1. 中国科学院水生生物研究所,武汉　430072; 2. 华中农业大学水产学院,武汉　430070)

摘要:对沙鳅亚科鱼类 3 属 14 个代表种的线粒体 DNA 控制区序列的结构进行了分析。通过与鲤形目鱼类的控制区序列进行比较,将沙鳅亚科鱼类的控制区分为终止序列区、中央保守区和保守序列区三个区域。同时识别了沙鳅亚科中一系列保守序列,并给出了它们的一般形式。以胭脂鱼为外类群,对比条鳅亚科、花鳅亚科、以及平鳍鳅科的代表性种类,采用 NJ、MP 和 ML 法构建沙鳅亚科的分子系统树。分子系统发育分析表明,沙鳅亚科为一单系,包括 3 个属:沙鳅属、副沙鳅属和薄鳅属,各属均构成单系。根据分子系统学、形态学的结果及地理分布推断,沙鳅亚科中沙鳅属可能为最为原始的属,副沙鳅属其次,而薄鳅属最特化。

关键词: 沙鳅亚科;线粒体 DNA;控制区;系统发育
中图分类号:Q951　　**文献标识码**:A　　**文章编号**:1000-3207(2005)06-0645-09

鱼类线粒体 DNA 作为一种分子标记目前已被广泛应用于系统学研究[1—3]。线粒体 DNA 中的控制区即 D-loop 区为非编码区,不编码蛋白质,受到选择压力较小,因此积累了较多的突变,如碱基替换、插入、缺失,以及众多的串联重复序列等。控制区较为复杂,分为 3 个区段,终止序列区(Extended termination associated sequence, ETAS)、中央保守区(Central conserved dormain, CD)和保守序列区(Conserved sequence block, CSB),不同区段进化速度不同,故可适用于不同级别的系统发育分析研究[4]。

目前对控制区的研究已有较多报道[4—7],众多的保守序列也被识别。在鱼类方面,Buroker 等[5]和 Broughton 等[6]分别识别了鲟鱼和雅罗鱼线粒体控制区的终止相关序列 TAS 和保守序列 CSB1、CSB2、CSB3。Liu 等[7]识别了鲤形目鱼类的保守序列 ETAS1、CSB-F、CSB-E、CSB-D、CSB-1、CSB-2 和 CSB-3。同时,将控制区用于系统发育分析的报道也不断增加[8,9]。

从分子水平上来研究沙鳅亚科系统发育关系的报道较少,Liu 等[9]提及了本亚科的 3 个个体:*Botia sp.*, *Parabotia banarescui*, *Leptobotia elongata*。杨军山[10]对副沙鳅属的 5 个种从形态和分子水平上进行了较为详细的系统发育及生物地理学的研究。有关沙鳅亚科控制区序列的分析还未有报道,本研究拟通过对沙鳅亚科 3 属 14 个代表种控制区序列的研究,分析沙鳅亚科的系统发育关系,从而为整个鳅科分类及系统发育关系的研究提供分子基础。

1　材料和方法

1.1　材料　研究中选择了沙鳅亚科 3 属共 14 个代表种,进行控制区序列的结构分析。同时,选取了鳅科中条鳅亚科(Nemacheilinae)和花鳅亚科(Cobitinae)的 3 个种,以及平鳍鳅科(Homalopteridae)的 4 个种进行比较。所分析的 DNA 序列为 PCR 直接测序所得,部分从 GenBank 下载。为验证沙鳅亚科的单系性,选取胭脂鱼 *Myxocyprinus asiaticus* 作为外类群。实验所用标本均用 95 %的酒精固定,保存于中国科学院水生生物研究所淡水鱼博物馆。标本名称及来源见表 1。

1.2　基因组 DNA 的提取、PCR 扩增　基因组 DNA 的提取采用标准的蛋白酶 K 消化法。基因序列由 PCR 扩增所得。扩增所用引物为 DL1 和 DH2,其序列分别为 DL1:ACC CCT GGC TCC CAAAGC, DH2:ATC TTA GCA TCT TCA GTG。PCR 反应的总体积为 60μL,大约用 50 ng 的基因组 DNA 作为模板,反应体系包括:10 ×Buffer 6 μL, dNTPs 3 μL(2.5 mmol/L),

收稿日期:2004-04-20;**修订日期**:2004-05-20
基金项目:国家自然科学基金(30170137,404320037),中国科学院创新方向项目(KZCX3-SW-126)资助
作者简介:唐琼英:(1976—),女,四川简阳人;博士;研究方向:鱼类分子进化
通讯作者:刘焕章,E-mail:hzliu@ihb.ac.cn

引物各 3 μL (10 pmol/L) , *Taq* 酶 1.5U。PCR 反应条件为:94 ℃预变性 3 min ,然后循环包括:94 ℃变性 45s ,56 ℃退火 45s ,72 ℃延伸 1 min ,共 35 个循环 ,最后在 72 ℃延伸 8 min。

1.3 PCR 产物的纯化和测序 PCR 产物用 0.8 % 的琼脂糖电泳检测 ,在紫外透射仪上观察 ,并割下亮带。将割下的胶用 Glassmilk DNA 纯化回收试剂盒进行纯化回收 ,回收产物送上海基康生物公司测序。序列递交 GenBank , 获得的序列号为: AY600862-AY600880(表 1)。

1.4 数据处理及分析 DNA 序列的排列(Alignment)使用 Clustal X 软件[11] ,并在 SEAVIEW 程序[12] 中对序列进行手工调整。对比 GenBank 中缨口鳅线粒体 DNA 全序列 ,找到控制区的起点和终点。并以 CSB-F 和 CSB1 的起点分别作为终止序列区、中央保守区和保守序列区的分界线。在识别的保守序列中 ,简并密码子的代号分别为:W = A 或 T,R = A 或 G,M = A 或 C,Y = C 或 T,K = G 或 T。序列中各碱基

表 1 研究中所用标本及来源

Tab.1 Specimens and their localities in the present study

属名 Genus	种名及单倍型 Species	采集地 Locality	所属水系 Drainage	标本号或序列号 Specimen No. or Accession No.
	张氏薄鳅 *L. tchangi*	江西上饶	长江下游	0000024/ AY600871
	天台扁尾薄鳅 1 *L. tientaiensis tientainensis*1	江西婺源	长江下游	0000027/ AY600865
	天台扁尾薄鳅 2 *L. tientaiensis tientainensis*2	广西阳朔	珠江水系	0000028/ AY600866
薄鳅属 *Leptobotia*	薄鳅 *L. pellegrini*	福建福安	闽江水系	0000029/ AY600873
	红唇薄鳅 *L. rubrilabris*	四川彭县	长江上游	0000021/ AY600872
	长薄鳅 *L. elongata*	四川泸洲	长江上游	0000019/ AY600875
	汉水扁尾薄鳅 *L. tientaiensis hansuiensis*	湖南桃源	沅江	0307110/ AY600874
	紫薄鳅 *L. taeniops*	四川彭县	长江上游	0000020/ AY600870
	花斑副沙鳅 *P. fasciata*	湖北武汉	长江中下游	0000038/ AY600868
副沙鳅属 *Parabotia*	武昌副沙鳅 *P. banarescui*	湖南辰溪	沅江水系	0000037/ AY600869
	漓江副沙鳅 *P. lijiangensis*	湖南辰溪	沅江水系	0000036/ AY600867
	点面副沙鳅 *P. maculosa*			AY091508 *
	江西副沙鳅 *P. kiangsiensis*			AY091509 *
	中华沙鳅 1 *B.(Sinibotia) superciliaris*1	四川彭县	长江上游	0000030/ AY600862
沙鳅属 *Botia*	中华沙鳅 2 *B.(S.) superciliaris*2	长江中上游	长江中上游	0000031/ AY600863
	壮体沙鳅 1 *B.(Hymenophysa) robusta*1	广西南宁	珠江水系	0000033/ AY600864
	壮体沙鳅 2 *B.(H.) robusta*2			AY091510 *
条鳅属 *Nemacheilus*	北方须鳅 *N. barbatulus*			AJ388422 *
泥鳅属 *Misgurnus*	泥鳅 *M. anguillicaudatus*	广西	珠江水系	0000005/ AY600879
花鳅属 *Cobitis*	中华花鳅 *C. sinensis*	广西	珠江水系	0000012/ AY600880
缨口鳅属 *Crossostoma*	台湾缨口鳅 *C. lacustre*			M91245 *
	台湾缨口鳅 *C. lacustre*			NC001727 *
华平鳅属 *Sinohomaloptera*	广西华平鳅 *S. kwangsiensis*	海南乐东		0307116/ AY600878
金沙鳅属 *Jinshaia*	短身金沙鳅 *J. abbreviata*	四川乐山		0307117/ AY600876
华吸鳅属 *Sinogastromyzon*	四川华吸鳅 *S. szechuanensis*	四川乐山		0307119/ AY600877
胭脂鱼属 *Myxocyprinus*	胭脂鱼 *M. asiaticus*			AY017140 *

注: *表示从 GenBank 下载的序列

www.cnki.net

的含量及变异情况应用 Mega2.1 软件[13]进行统计。分子系统发育关系的重建采用 Mega2.1 中的邻接法(Neighbor-joining, NJ),PAUP*4.0b10 中的最大简约法(Maximum parsimony, MP),以及 TREE-PUZZLE5.0[14]软件中的最大似然法(Maximum likelihood, ML)。其中 NJ 树的构建采用 Kimura 双参数法,MP 树采用启发式搜索(Heuristic search),树二等分再连接选项(Tree-bisection-reconnection, TBR)。NJ 树和 MP 树分支的置信度以 1000 次自展分析(Bootstrap analysis)进行重复检验。ML 法采用 HKY85 替代模型,采用 Quartet puzzling 搜索,重复 10 000 次估计各分支的支持率。

2　结　果

2.1　沙鳅亚科鱼类线粒体 DNA 控制区序列及变异

沙鳅亚科鱼类线粒体 DNA 控制区的长度变化范围为 909bp—932bp,A、T、G、C 的平均含量为 35.5%,31.7%,19.5%,13.2%。A+T 的含量(67.2%)明显高于 C+G 的含量(32.7%),这与其他鱼类 A、T 含量高,G、C 含量低的特点相似。转换颠换比(Ts/Tv)为 2.60。对比 Liu 等[6]中鲤形目线粒体 DNA 控制区的保守序列,将沙鳅亚科的控制区初步分为三个区段:终止序列区(ETAS)、中央保守区(CD)和保守序列区(CSB)。各区段的碱基组成情况见表 2。比较三个区段的碱基组成可以得出:A 的平均含量在终止序列区(39.6%)和保守序列区(37.4%)显著高于中央保守区(30.4%),这与一般认为的终止序列区和保守序列区的 A 的含量应明显高于中央保守区的观点是一致的[7]。中央保守区中 G 的含量(17.8%)明显高于终止序列区(9.6%)和保守序列区(11.6%)。研究中还发现,序列变异率在终止序列区最高(18.2%),保守序列区次之(14.1%),中央保守区最低(12.0%)。这种现象也在其他脊椎动物中发现,这说明终止序列区变异最大,中央保守区最为保守。对所有研究鱼类的控制区序列进行简约分析,结果显示在 963 个位点中,共有 601 个变异位点,其中有 409 个为简约信息位点。

2.2　终止序列区(ETAS)

终止序列区是控制区变异最大的区域,因含有与 DNA 复制终止相关的序列而得名。沙鳅亚科的终止序列区长度约在 234bp—255bp 间。在此区段发现一段保守的终止结合序列 TAS(TACATA),还有一段与 TAS 互补的序列(ATGTAT),分别位于 19—25nt 处和 62—72nt 处,对比其他鱼类中的保守序列[11],在终止序列区识别了沙鳅亚科的 ETAS1 这一保守序列,其序列为 TACATAWTATGCATRATATTACAT,与 Liu 等[7]在鲤形目中识别出的略有不同。

2.3　中央保守区(CD)

控制区的中央保守区是最为保守的区域。在哺乳动物中识别了保守序列 B、C、D、E、F[15]。Lee 等[16]在鱼类中则仅识别了 CSB-D。而后 Liu 等[7]和张燕等[17]又分别识别了鲤形目和[illegible]科鱼类中的 CSB-D,CSB-E 和 CSB-F 三个保守序列。类似地,沙鳅亚科中的这几个保守序列也被识别出来,近 5′端的是 CSB-F,其普遍形式为:ATGTAGTAAGARMC-CACCAACYRGTTTA。紧接其后的是 CSB-E 和 CSB-D,它们的形式分别为:TCARGGRCAAYAATYGTRRRGK 和 TGAACTATTACTGGCATYTGGTTCCTATTTCAGG。其中 CSB-F 和 CSB-D 较为保守,很容易识别出来,并与 Liu 等[7]鲤形目中的这两段序列几乎相同。在沙鳅亚科的大多数种类中发现有一个 GTGGG-box 存在于 CSB-E 中。

2.4　保守序列区(CSB)

哺乳动物的保守序列区一般存在 CSB-1,CSB-2,CSB-3。众多研究发现,在鱼类的保守序列区中同样存在这三个保守序列。与哺乳动物不相一致[18],鱼类中 CSB-2 最为保守,最易被识别出来,CSB-3 比 CSB-2 变异稍大,也较易识别,而 CSB-1 在鱼类中的变异很大,不易识别,但有一保守片段 GACATA 可以帮助识别。在沙鳅亚科中成功地识别了 CSB-1、CSB-2 和 CSB-3,它们的普遍形式分别为:AGRTKAATGMTWRAAWGACATAMCC(-)AYAA-GA,CGCGACAAACCCCCC(-)C(-)TTACCCCC(-)TACGC 和 CCTTGTCAAACCCCRAAACCAAGGARRRC。另外在 CSB-3 之后大概 50nt 处发现有一个极其保守的片段 CATATTATATATAA。

2.5　系统发育关系分析

对本研究的所有类群控制区序列进行分析,转换已达饱和,但颠换仍未饱和,因此仍具有系统发育意义。以胭脂鱼为外类群,采用三种不同的方法构建分子系统发育树。其中 NJ 法只采用颠换数进行分析。1000 次自展分析所得的支持率大于 50%的一致树列于图 1。此图显示沙鳅亚科鱼类(包括沙鳅属、副沙鳅属和薄鳅属)聚在一起构成一个单系,且具有较高的支持率(88%)。其他非沙鳅鱼类(包括条鳅亚科、花鳅亚科和平鳍鳅科鱼类)构成另一个分支(支持率为 75%)。在沙鳅亚科这一大的分支内部,又明显分为三个支系:沙鳅属、副沙鳅属和薄

表 2 沙鳅亚科控制区序列碱基组成情况表

Tab. 2 Nucleotide compositions of the mitochondrial DNA control region in the subfamily Botiinae

	终止序列区(ETAS)				中央保守区(CD)				保守序列区(CSB)				整个控制区				总长度
	T(U)	C	A	G	T(U)	C	A	G	T(U)	C	A	G	T(U)	C	A	G	
B. (*S.*) *superciliaris* 1	32.9	18.9	38.3	9.9	34.4	15.6	32.5	17.5	28.1	24.6	37.5	9.7	31.6	20.0	36.0	12.5	912
B. (*S.*) *superciliaris* 2	32.9	18.9	37.9	10.3	34.1	15.9	32.5	17.5	28.2	24.7	37.6	9.5	31.5	20.1	35.9	12.5	911
B. (*H.*) *robusta* 1	36.4	15.1	38.9	9.6	32.8	18.3	30.0	18.9	27.8	24.9	38.8	8.5	31.8	20.0	35.8	12.4	909
B. (*H.*) *robusta* 2	36.2	15.2	39.9	8.6	33.1	18.0	30.3	18.6	27.7	25.5	38.1	8.7	31.8	20.2	35.9	12.1	917
L. tientaiensis tientaiensis 1	31.7	17.1	39.8	11.4	33.3	18.8	30.2	17.6	29.3	22.1	36.6	12.0	31.4	19.6	35.2	13.8	928
L. tientaiensis tientaiensis 2	31.3	17.9	39.0	11.8	33.0	19.1	29.6	18.2	29.1	21.8	37.4	11.7	31.0	19.8	35.1	14.0	928
P. maculosa	35.6	15.8	38.1	10.5	33.3	18.5	30.2	17.9	29.9	20.7	37.7	11.7	32.6	18.6	35.2	13.6	929
P. kiangsiensis	36.0	15.4	39.7	8.9	32.4	19.1	29.9	18.5	28.9	21.4	36.7	13.1	32.0	19.0	35.1	13.9	931
P. lijiangensis	33.5	17.7	39.5	9.3	33.1	18.7	29.4	18.7	28.2	22.3	38.0	11.5	31.3	19.8	35.4	13.4	932
P. fasciata	35.0	15.8	40.2	9.0	33.8	18.5	29.8	17.8	29.1	21.2	37.6	12.1	32.3	18.8	35.5	13.4	913
P. banarescui	33.6	17.8	39.3	9.3	32.3	19.7	30.5	17.5	29.1	21.2	37.4	12.3	31.4	19.8	35.5	13.3	930
L. taeniops	33.1	18.0	39.6	9.4	33.6	18.8	30.2	17.3	29.8	20.3	37.0	12.8	32.0	19.2	35.3	13.5	928
L. tchangi	33.6	17.0	40.1	9.3	33.2	18.5	31.1	17.2	30.2	20.4	36.9	12.6	32.2	18.8	35.7	13.3	930
L. rubrilabris	33.6	17.4	40.1	8.9	32.9	19.4	29.5	18.2	28.7	21.7	37.6	12.0	31.5	19.8	35.4	13.3	931
L. pellegrini	32.8	17.8	41.3	8.1	32.9	19.4	31.1	16.6	28.8	21.2	37.4	12.6	31.3	19.7	36.2	12.8	930
L. tientaiensis hansuiensis	34.0	17.0	40.1	8.9	32.2	19.8	30.0	18.0	28.8	21.9	36.6	12.7	31.4	19.9	35.2	13.5	931
L. elongata	31.4	17.6	41.2	9.8	34.0	18.2	30.6	17.3	29.3	20.9	36.6	13.1	31.5	19.1	35.7	13.7	927
Avg.	33.7	17.1	39.6	9.6	33.2	18.5	30.4	17.8	28.9	22.2	37.4	11.6	31.7	19.5	35.5	13.2	

鳅属,它们分别构成单系群,支持率分别为 100 %、65 %、78 %。这三个支系中,沙鳅属处于根部位置,属于最原始的类群,其次是副沙鳅属,薄鳅属最为特化。

对于 ML 法,采用 Quartet puzzling 搜索,进行 10 000 次重复分析,得到的沙鳅亚科鱼类系统发育关系与 NJ 法的基本一致(图 2),强烈支持沙鳅亚科鱼类构成一个单系(支持率为 94 %)。沙鳅属、副沙鳅属和薄鳅属各属构成单系的支持率分别为 100 %、75 %、70 %。各属间的系统发育关系在此也能得到很好的反映。其他非沙鳅鱼类分成了两个分支,分别为鳅科的类群和平鳍鳅科的类群,但平鳍鳅科中的台湾纓口鳅和鳅科中的花鳅亚科鱼类聚在了一起,支持率较低(58 %)。

MP 法中,赋予转换、颠换不同的权重(分别设置为 8∶1,4∶1,2∶1,1∶1),得到的拓扑结构基本一致,仅在分支的支持率上有所差异,故只在文中列出 1∶1 这组的结果(图 3)。MP 分析得到 1 棵最简约树(Length = 1736, CI = 0.547, RI = 0.663)。在 MP 树中,沙鳅亚科鱼类也能聚在一起构成一个单系,但支持率略低,为 57 %。在沙鳅亚科的内部,沙鳅属构成单系可以得到很高的支持率,薄鳅属构成单系的支持率仅为 56 %,而副沙鳅属鱼类不能聚在一起,它们与薄鳅属的鱼类成多歧的拓扑结构。所以在 MP 树中,副沙鳅属和薄鳅属的系统发育关系未能得到很好的解决。至于其他非沙鳅鱼类的关系与 ML 法得到的一致。

总的来说,三种不同方法构建的分子系统发育树都支持沙鳅亚科构成一个单系。在沙鳅亚科内部,明显分为三个类群,对应于沙鳅属、副沙鳅属和薄鳅属。三者间的系统发育关系为:沙鳅属最原始,其次是副沙鳅属,薄鳅属最为特化。

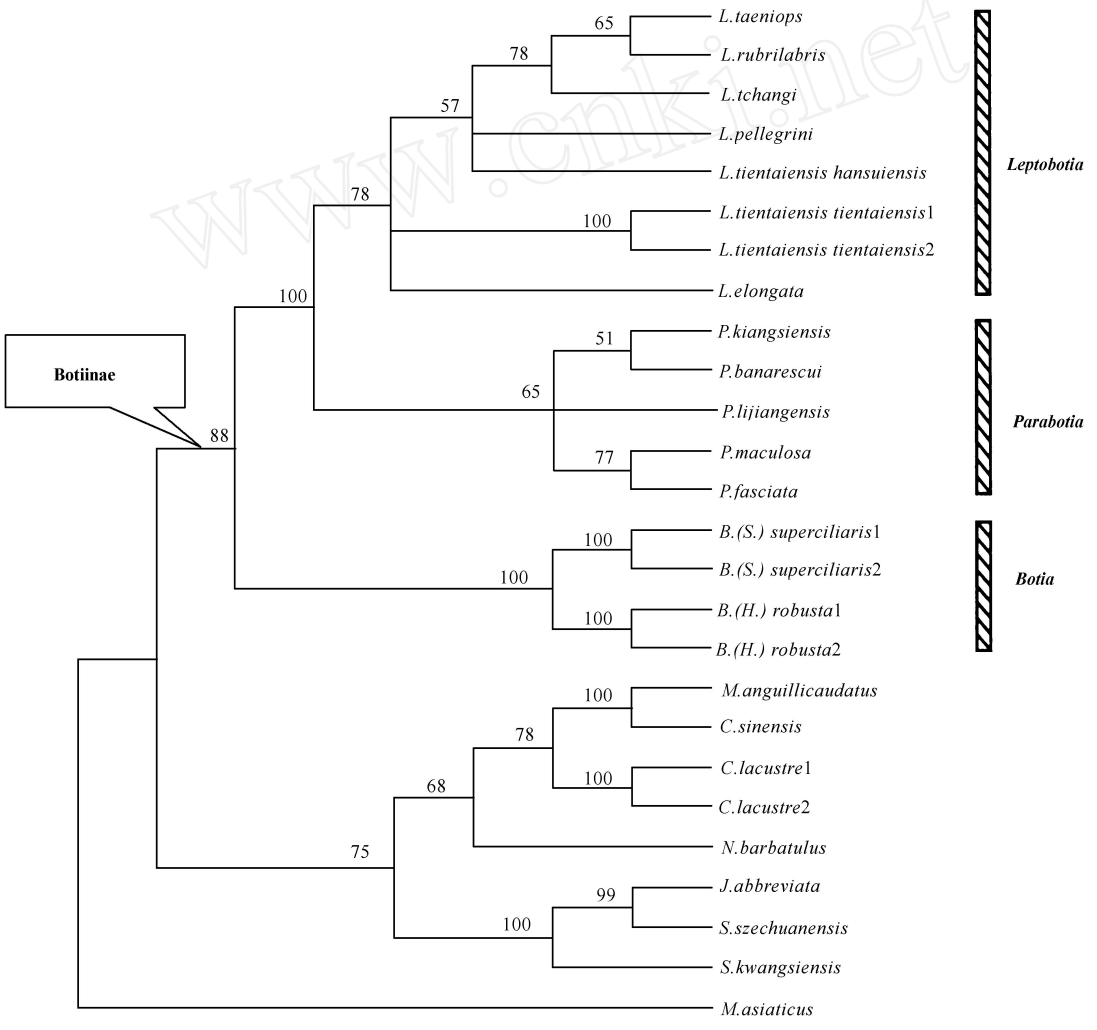


图 1　基于控制区序列构建的沙鳅亚科的 NJ 树。枝上数值表示 1000 次重复抽样所得的大于 50 %的支持率(只采用了颠换数)

Fig. 1　NJ tree of Botiinae based on the control region. The value at the nodes present the bootstrap value above 50 % using 1000 bootstrap replications. (Only transversions included)

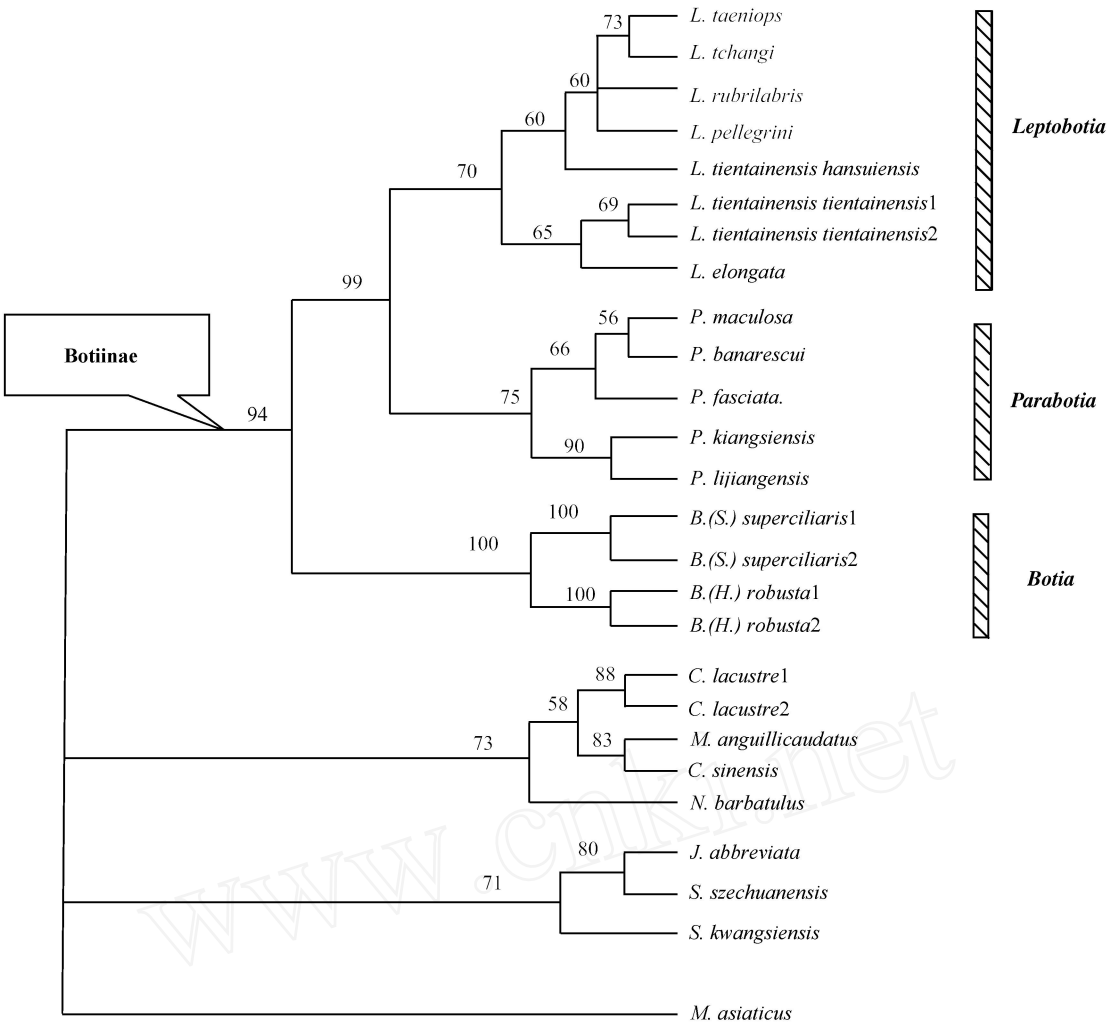


图 2 基于控制区序列构建的沙鳅亚科的 ML 树,替代模型为 HKY85,节点的数值为 10 000 次 Quartet puzzling 搜索所得的大于 50 %的支持率

Fig. 2 ML tree of Botiinae based on the control region. HKY85 was chose as the evolution model. The value at the nodes present the confidence level above 50 % using 10 000 Quartet puzzling

3 讨 论

3.1 沙鳅亚科控制区特点及用于系统发育研究的可行性

同其他鱼类一样,沙鳅亚科的控制区分为终止序列区、中央保守区和保守序列区,其中终止序列区为富含 AT 的区域,A + T 的含量高达 73.3 %,此区变异最大,包含了较多的系统发育信息。而中央保守区为富含 GC 的区域,G + C 的含量(36.3 %)远高于终止序列区(26.7 %),这或许是中央保守区最为保守的一个原因,此区适合种属以上阶元间的系统发育分析。在沙鳅亚科的控制区中,笔者成功地识别了一系列的保守序列,包括 ETAS1、CSB-F、CSB-E、CSB-D、CSB-1、CSB-2 和 CSB-3,其中 CSB-F 和 CSB-1 的起点是分开终止序列区、中央保守区和保守序列区的分界点。

对研究类群的控制区序列进行分析,转换已达到饱和,但颠换仍未饱和,转换/颠换值为 1.597,因此仍具有系统发育意义。由于转换达到饱和,受进化噪音影响的可能性较大,所以在用 NJ 法进行系统发育树的重建时,只采用颠换数进行分析。在 MP 法中赋予转换和颠换不同的权重,但拓扑结构一样,且均具有较高的 CI 值,故采用了最简单的一组,即转换和颠换具有相同权重,得到的树长最短。至于 ML 法,设定转换/颠换为 2.0 进行分析。

3.2 沙鳅亚科的系统发育关系

以胭脂鱼作为外类群,与鳅科中的条鳅亚科、花鳅亚科以及平鳍鳅科鱼类作比较,沙鳅亚科可以构成一个单系,这在三种方法构建的分子系统发育树中均可得到反映。在形态特征上,陈景星等[19]以胭

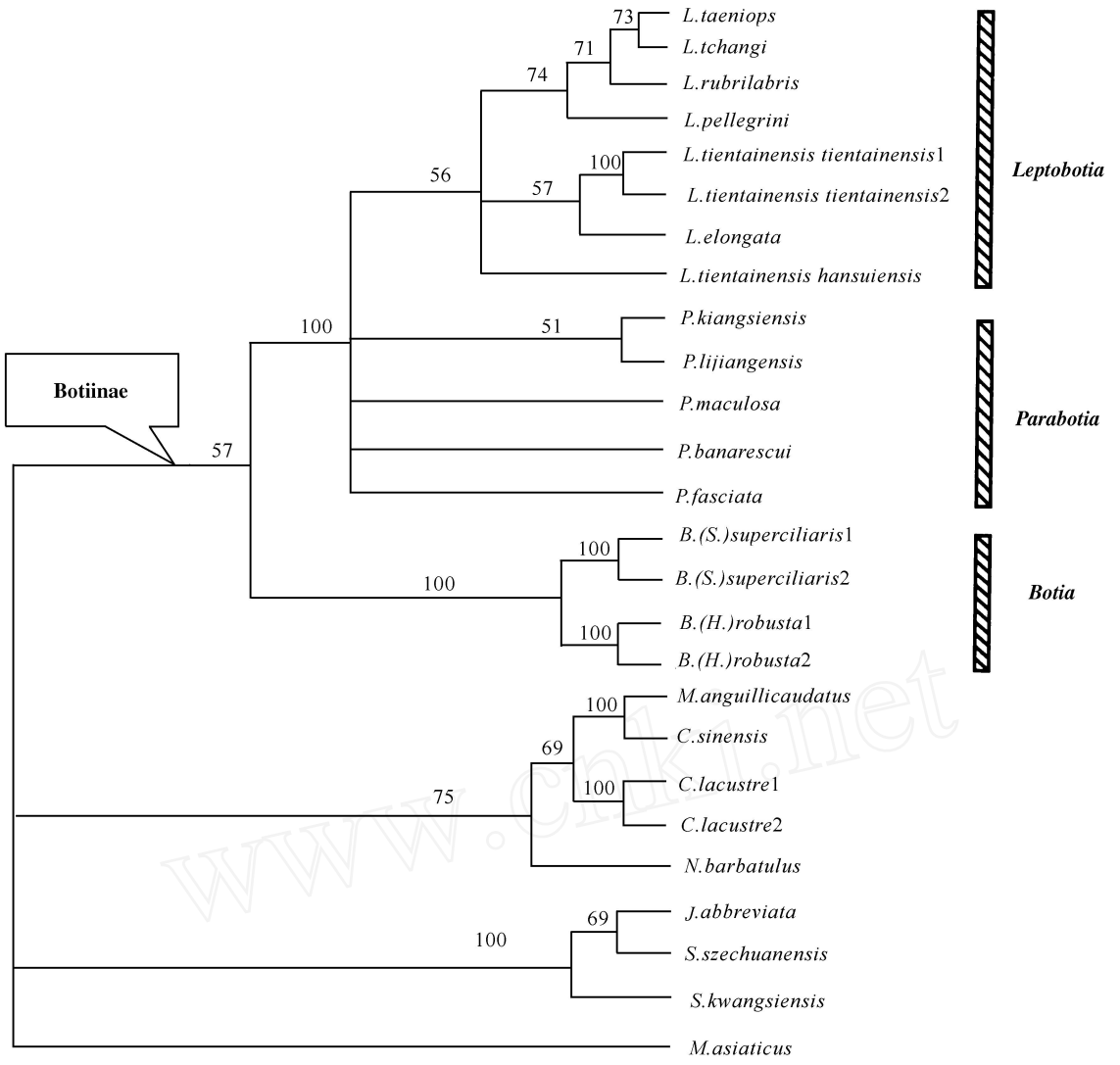


图 3　基于控制区构建的沙鳅亚科的 MP 树,节点上的数值为 1000 次重复抽样所得的大于 50 %的支持率

Fig. 3　MP tree of Botiinae based on the control region. The value at the nodes present the bootstrap value above 50 % using 1000 bootstrap replications

脂鱼为外类群,对鳅科鱼类的系统发育关系进行了研究。沙鳅亚科以具眼下刺,2 对吻须聚生于吻端,尾鳍深分叉,侧线完全,骨质鳔由第二脊椎横突的腹支向后伸展与第四脊椎横突、肋骨和悬器构成等特征而明显地区别于鳅科的其他两个类群。形态和分子的结果都显示沙鳅亚科构成一个单系。

同时,分子系统树显示:沙鳅亚科由三个主要的支系组成,分别为沙鳅属、副沙鳅属和薄鳅属,它们各自形成单系。在 MP 树中,副沙鳅属鱼类聚在一起的支持率低于 50 %。杨军山[10]在对副沙鳅属的 6 个种进行详细的分子和形态特征研究后认为副沙鳅属为一单系;作者在用线粒体 cyt *b* 基因进行研究时也得到同样的结果。因此作者认为副沙鳅属为一个有效属,即沙鳅亚科应包括 3 个属。这与陈景星形态学上的分类观点基本一致,同时也说明了 Fang 和 Nalbant 将副沙鳅属作为沙鳅属中膜鳔沙鳅亚属的同物异名的观点是不正确的。

另外,分子系统树还显示:沙鳅亚科中沙鳅属处于根部位置,是最为原始的类群,副沙鳅属其次,薄鳅属最为进化。陈景星等[23]在推测鳅科鱼类的起源时曾提出:鳅科鱼类的起源可能是在现今华来氏线以北的大巽他群岛、泰国和缅甸这一区域,后来向北和向西扩散和演化。后因喜马拉雅山脉隆起,周围地文引起变迁,沙鳅亚科鱼类的一支往西扩散致印度和巴基斯坦,演化为现今的沙鳅属,而另一支往北扩散致珠江、长江,远至黑龙江和日本,演化为现今的副沙鳅属和薄鳅属。中国鳅科中的一些原始类群分布在云南和广西南部。从目前沙鳅属的分布来

看,除少数几个种外大部分种均分布在云南,因此推断沙鳅属应为沙鳅亚科中最原始的一个属。陈景星[1]根据颊部有无鳞片将沙鳅亚科分为两大类群,沙鳅属以颊部裸露首先分离出来。然后薄鳅属再以眼下刺不分叉与副沙鳅属区别开来。陈景星等[19]以胭脂鱼的特征作为原始类型,分析了鳅科鱼类的系统发育关系,其研究中未提及胭脂鱼的颊部是否裸露。作者检视了保存在淡水鱼类博物馆中的胭脂鱼标本,发现其颊部裸露。因此,沙鳅属的颊部裸露为祖征,副沙鳅属和薄鳅属颊部具鳞为离征,副沙鳅属比沙鳅属更特化。另外,沙鳅属和副沙鳅属的眼下刺分叉,而薄鳅属眼下刺不分叉。薄鳅属的眼下刺不分叉为一离征。

根据上述分子、形态及地理分布的分析,可以推断,沙鳅亚科为一单系,包括 3 个属:沙鳅属、副沙鳅属和薄鳅属,它们均各自构成单系。其中沙鳅属最为原始,副沙鳅属较为特化,而薄鳅属最进化。

参考文献:

[1] Xiao W H, Zhang Y P. Genetics and evolution of mitochondrial DNA in fish [J]. *Acta hydrobiologica sinica*, 2000, **24**(4):384—391[肖武汉,张亚平. 鱼类线粒体 DNA 的遗传与进化. 水生生物学报,2000,**24**(4):384—391]

[2] Yang J Q, Liu H Z. Mitochondrial cytochrome *b* gene sequence variations of two Bagrid fishes in Changjiang River and Zhujiang River [J]. *Acta Hydrobiologica Sinica*, 2003, **27**(3):253—257[杨金权, 刘焕章. 两种[illegible]科鱼类在长江和珠江流域 cytb 基因序列变异性分析. 水生生物学报,2003,**27**(3):253—257]

[3] Tang Q Y, Yang X P, Liu H Z. Biogeographical process of *Spinibarbus caldwelli* revealed by sequence variations of mitochondrial cytochrome *b* gene [J]. *Acta Hydrobiologica Sinica*, 2003, **27**(4):352—365[唐琼英,杨秀平,刘焕章. 刺[illegible]基于线粒体细胞色素 b 基因的生物地理学过程. 水生生物学报,2003,**27**(4):352—365]

[4] Liu H Z. The structure and evolution of mitochondrial DNA control region of fish: a case study to bitterlings [J]. *Prog. Nat. Sci.*, 2002, 12(3): 266—270[刘焕章. 鱼类线粒体 DNA 控制区的结构和进化:以[illegible]鱼类为例. 自然科学进展,2002,**12**(3):266—270]

[5] Buroker N E, Brown J R, Gilbert T A, *et al*. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model [J]. *Genetics*, 1990, **124**:157—163

[6] Broughton R E, Dowling T E. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera* [J]. *Genetics*, 1994, **138**:179—190

[7] Liu H, Teng C S, Teng H Y. Sequence variations in the mtiochondrial DNA control region and their implications for the phylogeny of the Cypriniformes [J]. *Can. J. Zool.*, 2002, **80**:569—581

[8] Gilles A, Lecointre G, Miquelis A, *et al*. Partial Combination Applied to Phylogeny of European Cyprinids Using the Mitochondrial Control Region [J]. *Mol. Phylogenet. Evol.*, 2001, **19**(1):22—33

[9] Liu H, Chen Y. Phylogeny of the East Asian cyprinids inferred from sequences of the mitochondrial DNA control region [J]. *Can. J. Zool.*, 2003, **81**:1938—1946

[10] Yang J S. Studies on the phylogeny and biolgeography of Parabotia fishes (Pisces: Cobitidae) Master degree Thesis [D] Institue of Hydrobiology, Chinese Academy of Sciences. 2002 [杨军山. 副沙鳅属鱼类的系统发育与生物地理学研究. 硕士论文. 中国科学院水生所]

[11] Thompson J D, Gibson T J, Plewniak F. The Clustal X windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools [J]. *Nucleic. Acids. Res.*, 1997, **25**(4): 4876—4882

[12] Galtier N, Gouy M, Gautier C. SEAVIEW and PHYLO-WIN: two graphic tools for sequence alignment and molecular phylogeny [J]. *Comput. Appl. Biosci.*, 1996, **12**:543—548

[13] Kumar S, Tamura K, Jakobsen I B, *et al*. MEGA2: Molecular Evolution Genetics Analysis Software [M]. Arizona State University, Tempe. 2001

[14] Strimmer K, von Haeseler A. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies [J]. *Mol. Biol. Evol.*, 1996, **13**: 964—969

[15] Southern S O, Southern P J, Dizon A. E. Molecular characterization of a clone dolphin mitochondrial genome [J]. *J. Mol. Evol.*, 1988, **28**: 32—42

[16] Lee W, Conroy J, Howell W H, *et al*. Structure and evolution of teleost mitochondrial control regions [J]. *J. Mol. Evol.*, 1995, **41**: 54—66

[17] Zhang Y, Zhang E, He S P. Study on the structure of the control region of the Bagridae in China and its phylogenetic significance [J]. *Acta Hydrobiologica Sinica*, 2003, **27**(**5**): 463—467[张燕,张鹗,何舜平. 中国[illegible]科鱼类线粒体 DNA 控制区结构及其系统发育分析. 水生生物学报, 2003, **27**(**5**): 463—467]

[18] Sbis à E, Tanzariello F, Reyes A, *et al*. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications [J]. *Gene*, 1997, **205**: 125—140

[19] Chen J X, Zhu S Q. Phylogenetic relationships of the subfamily in the loach family Cobitidae (Pisces) [J]. *Acta Zootaxon. Sin.*, 1984, **9**(**2**): 201—207[陈景星,朱松泉,鳅科鱼类亚科的划分及其宗系发生的相互关系[J]. 动物分类学报,1984,**9**(2):201—207]

STUDIES ON THE STRUCTURE OF THE MITOCHONDRIAL DNA CONTROL REGION AND PHYLOGENETIC RELATIONSHIPS OF THE SUBFAMILY BOTIINAE

TANG Qiong-Ying[1,2], LIU Huan-Zhang[1], YANG Xiu-Ping[2] and XIONG Bang-Xi[2]

(1. *Institute of Hydrobiology, The Chinese Academy of Sciences, Wuhan* 430072;
2. *College of Fishery, Huazhong Agricultural University, Wuhan* 430070)

Abstract: Belonging to Cobitidae of Cypriniformes, the subfamily Botiinae is a group of small or middle freshwater fishes dwelling in slow to moderately swift waters, and many of them are common aquarium species. The classification of Botiinae fishes is controversial among different authors. Fang (1936) and Nalbant (1963) divided Botiinae into two genera, *Botia* and *Leptobotia*, while Chen (1980) separated it into three, *Botia*, *Parabotia* and *Leptobotia*. Up to now, there was no study conducted on the structure of mitochondrial DNA control region and the molecular phylogenetic relationship of the subfamily Botiinae. In the present study, the structure of the control region sequences was analyzed for fourteen species representing the three genera of Botiinae. Similar to other fishes, the control region of Botiinae is composed of three domains, the extended termination associated sequence (ETAS) domain, the Central conserved domain (CD) and the Conserved sequence block (CSB). The ETAS domain is very rich (73.3 %) in the contents of A and T, while the CD domain is much rich (36.3 %) in G and C. A termination-associated sequence TAS (TACATA) and its reverse complementary sequence (ATGTAT) were found in the ETAS domain, and the conserved sequence in this domain is ETAS 1. The CD domain is highly conserved, and three conserved sequences were identified, CSB-F, CSB-E and CSB-D. Three conserved sequences were identified in the CSB domain, CSB-1, CSB-2 and CSB-3.

With *Myxocyprinus asiaticus* as outgroup, compared with members of Nemacheilinae, Cobitinae and Homalopteridae as well, the phylogenetic relationship of Botiinae was analyzed using the neighbor-joining (NJ) method in MEGA2.1, the maximum parsimony (MP) method in PAUP* 4.0b10 and the maximum likelihood (ML) method in TREE-PUZZLE 5.0. The results showed that the subfamily Botiinae is monophyletic and the confidence levels in the three analysis methods were 88 % in NJ, 57 % in MP and 97 % in ML respectively. It also supported that Botiinae contains three genera, *Botia*, *Parabotia* and *Leptobotia*, since each of these genera formed a monophyly, which were supported with 100 %, 65 % and 78 % of bootstrap values in NJ method and 100 %, 75 % and 70 % of Quartet puzzling values in ML method. In MP analysis, the bootstrap value supporting monophyly of the genus *Parabotia* was less than 50 %, while for *Botia* and *Leptobotia*, the values were 100 % and 56 %.

Based on the results of molecular phylogenetic analysis, morphological comparison and geographical distribution of Botiinae fishes, it was suggested that *Botia* is the most primitive genus, *Parabotia* the second, and *Leptobotia* the most specialized.

Key words: Botiinae; Mitochondrial DNA; Control region; Phylogeny