

4种鲷科鱼类的线粒体细胞色素 b 基因 序列及分子系统学分析

江世贵¹, 刘红艳^{1,2}, 苏天凤¹, 龚世园²

(1. 中国水产科学研究院南海水产研究所, 广东 广州 510030;

2. 华中农业大学水产学院, 湖北 武汉 430070)

摘要:从鲷科鱼类中黄鳍鲷 (*Sparus latus*)、黑鲷 (*S. macrocephalus*)、真鲷 (*Pagrosomus major*) 和平鲷 (*Rhabdosargus sarba*) 的线粒体 DNA 扩增出约 430 bp 的细胞色素 b (Cytb) 基因, 经 Clustal X 同源排序后得 408 bp 序列, 对该序列进行分析并结合 GeneBank 中黄鲷 (*Taius tumifrons*)、灰鳍鲷 (*S. berda*)、犁齿鲷 (*Evynnis japonica*) 和高背四长棘鲷 (*Argyrops spinifer*) 的该区段 DNA 序列比较, 共发现 126 个核苷酸位点存在变异 (30.7%), 序列中的转换大于颠换, 碱基替换多发生在密码子的第 3 位点。黑鲷和灰鳍鲷亲缘关系最近, 序列差异为 3.8%; 而黄鲷和平鲷的亲缘关系最远, 序列差异为 21.7%。用 MEGA2.1 软件中的 NJ 法构建的分子系统树, 将上述 8 种类分为 4 组, 其中黄鲷属、平鲷属和鲷属分类的结果与形态分类学的观点一致; 而传统分类上把犁齿鲷、真鲷和高背四长棘鲷归为 3 个不同属, 在本研究分子系统树中, 这 3 个属却聚成一支。本研究目的在于从分子水平上阐明鲷科鱼类的分类与系统进化关系。

关键词: 鲷科鱼类; 细胞色素 b; 分子系统学

中图分类号: Q951.3; Q959.483

文献标识码: A

文章编号: 1005-8737-(2003)03-0184-05

线粒体 DNA 为单倍型, 无重组, 呈母系遗传。由于受选择的压力小, 其进化速度较核 DNA 快 5~10 倍, 故线粒体 DNA 成为系统遗传学理想的分子标记^[1]。近年来, DNA 测序技术的发展, 使分子系统学进入了一个新时期, 而且在所有的方法中, DNA 序列的比较是最直接揭示遗传变异的方法。细胞色素 b 基因为重要的蛋白质编码基因, 其进化速度适中, 适于进行种间和属间亲缘关系的研究, 已在不少动物类群的系统进化和分类研究中得到了有效的应用^[2-5]。

鲷科 (Sparidae) 属鲈形目, 鲈亚目, 为世界性的鱼类^[6]。世界上有记载的鲷科鱼类大致分为 30 多个属^[6], 而我国的鲷科鱼类大致分为 6~8 个属, 其

属间和种间的系统分类关系混乱^[7-9]。虽然在形态学水平上做了不少工作, 但有关其分类和系统进化关系存在较多争议。因此, 本研究测定了分布于我国的鲷科鱼类中黄鳍鲷 (*Sparus latus*)、黑鲷 (*S. macrocephalus*)、真鲷 (*Pagrosomus major*) 和平鲷 (*Rhabdosargus sarba*) 细胞色素 b 基因部分片段序列, 并结合 GeneBank 中的黄鲷 (*Taius tumifrons*)、灰鳍鲷 (*S. berda*)、犁齿鲷 (*Evynnis japonica*) 和高背四长棘鲷 (*Argyrops spinifer*) 及外群牙鲆 (*Paralichthys olivaceus*) 的相应序列进行比较和分析, 旨在从分子水平上探索鲷科鱼类分类和系统进化关系。

1 材料和方法

1.1 材料来源

黄鳍鲷、黑鲷、真鲷和平鲷 4 种 DNA 序列通过 PCR 产物直接测序获得, 黄鲷、灰鳍鲷、犁齿鲷和高背四长棘鲷及外群牙鲆的 DNA 序列来自 GeneBank。材料和 DNA 数据的具体来源见表 1。

收稿日期: 2002-08-15; 修订日期: 2002-12-05.

基金项目: 国家基础性工作专项; 广东省科技兴海重大项目 (A200099A01); 中国水产科学研究院人才基金资助项目。

作者简介: 江世贵 (1964-), 男, 研究员, 主要从事海洋生物技术研究。E-mail: jiangsg@21cn.com

1.2 总 DNA 的提取

取 100 mg 左右的鲷科鱼背部肌肉,加 TEN9 细胞裂解液(50 mmol/L Tris · HCl pH 9.0, 100 mmol/L EDTA, 200 mmol/L NaCl),终质量分数为 2% SDS 和终质量浓度为 1 mg/mL Protein K,混匀后,56 ℃ 消化过夜,经饱和酚、酚-氯仿(酚:氯仿:异戊醇 = 25:24:1,体积比)、氯仿(氯仿:异戊醇 = 24:1,体积比)抽提,2 倍体积无水乙醇和 1/10 体积 NaAc 沉淀,70% (体积分数)乙醇洗涤后,用灭菌蒸馏水溶解,4 ℃ 存放^[10]。

1.3 线粒体 DNA 细胞色素 b 基因片段的扩增

目标片段为鲷科鱼类细胞色素 b (Cyt b) 编码区中约 430 bp 的一段 DNA 序列,采用的是一对通用引物: L14724: 5'-GACTTGAAAAACCACCGTTG-3'; H15149: 5'-CCTCAGAAGGATATTTGTCCTC-3'^[11]。PCR 反应总体积为 50 μL,其中 10 × Ex Taq

Buffer (Takara) 5 μL, Ex Taq polymerase (Takara) 0.3 μL (1.5 U), dNTP (2.5 mmol/L) (Takara) 4 μL, 10 pmol/μL 引物 2 μL, 1 μL 模板 DNA, 其余为灭菌蒸馏水。扩增条件为: 94 ℃ 预变性 2 min, 之后进行 35 个循环(94 ℃ 45 s, 55 ℃ 45 s, 72 ℃ 60 s), 最后 72 ℃ 延伸 7 min。扩增产物经 1.5% (质量分数) TAE 琼脂糖凝胶回收纯化后直接进行序列测定。

1.4 序列测定及数据分析

回收纯化后的 PCR 产物送至上海基康公司 (Genecore) 测序。使用 Clustal X (Ver. 5.0) 排定 DNA 序列, 并经人工核查。在此基础上, 用 DNASTAR 软件中的 Megalign 计算碱基转换/颠换率和序列百分差异。用 MEGA2.1^[12] 中的“Neighbor-Joining”构建分子系统树, 置信度由“Bootstrap”1 000 次循环检验^[13], 以牙鲆作为外群来比较。

表 1 鲷科鱼类的 DNA 数据或材料来源

Table 1 The origins of DNA data or tissue samples of Sparidae fishes

种类 Species	形态分类 Morphological taxonomy	鲷科鱼或序列来源 Origin of Sparidae fishes or sequences
黄鳍鲷 <i>S. latus</i>	鲷科, 鲷属 Sparidae, <i>Sparus</i>	珠海东澳岛 Dongao Island, Zhuhai
黑 鲷 <i>S. macrocephalus</i>	鲷科, 鲷属 Sparidae, <i>Sparus</i>	珠海东澳岛 Dongao Island, Zhuhai
真 鲷 <i>P. major</i>	鲷科, 真鲷属 Sparidae, <i>Pagrosomus</i>	广州黄沙水产市场 Guangzhou
平 鲷 <i>R. sarba</i>	鲷科, 平鲷属 Sparidae, <i>Rhabdosargus</i>	广州黄沙水产市场 Guangzhou
黄 鲷 <i>T. tumifrons</i>	鲷科, 黄鲷属 Sparidae, <i>Taius</i>	GeneBank (AF240786)
灰鳍鲷 <i>S. berda</i>	鲷科, 鲷属 Sparidae, <i>Sparus</i>	GeneBank (AF240715)
犁齿鲷 <i>E. japonica</i>	鲷科, 犁齿鲷属 Sparidae, <i>Erynnis</i>	GeneBank (AF240725)
高背四长棘鲷 <i>A. spinifer</i>	鲷科, 四长棘鲷属 Sparidae, <i>Argyrops</i>	GeneBank (AF240717)
牙 鲆 <i>P. olivaceus</i>	鲆科, 牙鲆属 Bothidae, <i>Paralichthys</i>	GeneBank (AB028664)

2 结果

2.1 鲷科鱼类 Cyt b 基因片段序列及其差异

从黄鳍鲷、黑鲷、真鲷和平鲷总 DNA 中成功的扩增出约 430 bp 的 Cyt b 基因片段(图 1)。PCR 产物直接测序得到的碱基序列在 GeneBank 中的登录号如下: 黄鳍鲷 AF539743、黑鲷 AF539742、真鲷 AF539741、平鲷 AF539740。结合 GeneBank 中的黄鲷、灰鳍鲷、犁齿鲷和高背四长棘鲷及牙鲆的相应序列, 经 Clustal X 同源排序后的 408 bp 序列见图 2。在所有 408 bp 序列中, 没有发现碱基的缺失或插入, 其中共检测到 126 个核苷酸位点存在变异, 约占 30.7%, 序列中的转换大于颠换, 碱基替换多发生在

密码子的第 3 位点。

由表 2 可知, 鲷科属之间细胞色素 b 序列的平均差异为 18.13%, 鲷属 3 个种(黄鳍鲷、黑鲷和灰鳍鲷)之间的平均差异为 8.73%, 真鲷、犁齿鲷和高背四长棘鲷之间的平均差异为 9.13%。序列差异最小是黑鲷和灰鳍鲷, 仅为 3.8%, 序列差异最大的是黄鲷和平鲷, 达到了 21.7%。

2.2 系统进化树

用 MEGA2.1 中的“NJ”法构建分子系统树, 其拓扑结构见图 3。Bootstrap 1000 给出各支的置信度, 最低为 79%, 最高为 99%。分子系统树可分为 4 组, 黄鲷和平鲷各 1 支, 鲷属 3 个种聚为 1 支, 真鲷、犁齿鲷和高背四长棘鲷聚为 1 支。

表2 鲷科鱼类细胞色素 b DNA 序列差异百分比(对角线以上)和转换/颠换数(对角线以下)

Table 2 Percentage divergence(above diagonal) and numbers of transitions/transversions(below diagonal) for cytochrome b sequences(408 bp) of Sparidae fishes

	黄鳍鲷 <i>S. latus</i>	黑鲷 <i>S. macrocephalus</i>	真鲷 <i>P. major</i>	平鲷 <i>R. sarba</i>	黄鲷 <i>T. tumifrons</i>	灰鳍鲷 <i>S. berda</i>	犁齿鲷 <i>E. japonica</i>	高背四长棘鲷 <i>A. spinifer</i>
黄鳍鲷 <i>S. latus</i>	—	9.9	20.6	19.7	19.2	12.5	21.0	17.3
黑鲷 <i>S. macrocephalus</i>	33/4	—	18.6	18.8	17.0	3.8	18.6	14.8
真鲷 <i>P. major</i>	53/19	49/17	—	15.8	16.1	18.6	6.4	10.9
平鲷 <i>R. sarba</i>	44/26	45/22	33/25	—	21.7	19.1	17.7	18.3
黄鲷 <i>T. tumifrons</i>	47/21	44/17	44/14	49/27	—	18.6	16.5	14.6
灰鳍鲷 <i>S. berda</i>	40/6	13/2	47/19	44/24	47/19	—	19.9	16.7
犁齿鲷 <i>E. japonica</i>	54/19	49/17	21/4	39/25	47/12	49/19	—	10.1
高背四长棘鲷 <i>A. spinifer</i>	46/16	40/14	32/9	40/26	42/11	44/16	31/7	—

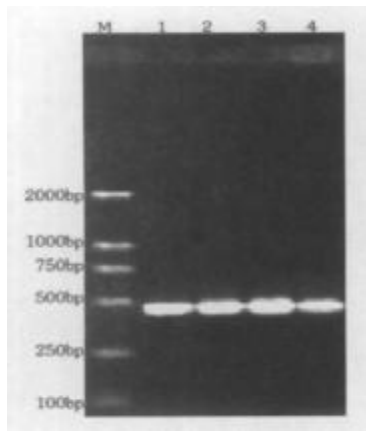


图1 线粒体 DNA Cyt b(引物 L14724 和 H15149)约 430 bp 基因片段的扩增结果(1.5%TAE 琼脂糖凝胶)

Fig. 1 An about 430 bp gene segment amplified with primers L14724 and H15149 from mtDNA Cyt b region (1.5% TAE agarose gel)

1-4: 分别对应黄鳍鲷、黑鲷、真鲷和平鲷的 PCR 扩增带;
M: Marker(DL 2000)

1-4: The corresponding PCR products of *S. latus*, *S. macrocephalus*, *P. major* and *R. sarba* separately. M: Marker(DL 2000)

3 讨论

成庆泰等^[7]编著的《中国鱼类系统检索》把分布于我国的鲷科鱼类分为黄鲷属、真鲷属、平鲷属、二长棘鲷属、四长棘鲷属、单列齿鲷属、犁齿鲷属和鲷属 8 个属;随后,伍汉霖等^[8]把二长棘鲷属并入犁齿鲷属;而且真鲷和平鲷的归属问题长期以来存在很大争议。有关它们的系统发育尚不十分清楚。早期 Forsskal 把平鲷划归 *Sparus*^[8],而多数学者认为平鲷应自成一属 *Rhabdosargus*^[6-7,9]。其实在形态特征上,平鲷属与鲷属已有了很大的分化,如头部形状、臼齿特征、背鳍和臀鳍数目有很大的差别^[6-7];

本研究结果中平鲷也自成单独的一支,从而说明把平鲷从鲷属中划分出来是合理的。多年来,把黄鳍鲷、黑鲷和灰鳍鲷归为一个属并无太大争议,只是不同的学者所用的属名不同^[6-9],本系统进化研究也印证了这个观点,而且得出灰鳍鲷和黑鲷的亲缘关系最近,它们的序列差异仅为 3.8%。

从表 2(上三角)可知,黄鲷与其他几种鲷的序列差异为 16.1%~21.7%,平鲷与其他几种鲷的序列差异为 14.6%~19.7%,鲷属与其他几种鲷的序列差异为 14.8%~21%,Cronin^[14]曾报道属间的线粒体 DNA 的序列差异在 14%~21%之间,假定以此为参考,那么分子系统树的根部 4 支应分别对应 4 属。黄鲷属、平鲷属和鲷属的分子系统分析结果与形态分类学上的研究结果相吻合。值得注意的是,传统的分类学把犁齿鲷、真鲷和高背四长棘鲷归为 3 个不同的属(虽然真鲷归属问题变动很大,但是以往的归类将之与四长棘鲷属和犁齿鲷属都视为不同的属),而本研究分子系统中,这 3 种鲷却聚成一支,而且真鲷和犁齿鲷的序列差异仅为 6.4%,3 种鲷间序列差异最大也仅为 10.9%,处于种的分化水平,序列变化处在属变化范围之内。另外,在幼鱼阶段,二长棘鲷背鳍鳍棘还未延长时,真鲷的体形、体色和身体上兰色的斑点与其都极为相似,从外形上很难把它们区分开^[7],从而说明真鲷和二长棘鲷具有较近的亲缘关系,而最近伍汉霖等^[8]根据形态分类特征,把二长棘鲷属并入犁齿鲷属,说明真鲷跟犁齿鲷也应具有较近的亲缘关系。四长棘鲷与二长棘鲷最根本的区别在于背鳍鳍棘延长的数目不同,从而似乎也可以推断出真鲷与四长棘鲷具有较近的亲缘关系。至于这 3 种鲷是否可以归为一个属,我们建议分类学家们对其系统分类地位作进一步考证。

```

#S. macro. ATG GCA AGC CTT CGA AAA ACA CAC CCC CTA TTA AAA ATT GCT AAC CAC GCA GTA GTT GAC CTA CCT
#S. berda. ....
#S. latus. ....T..T.T.G.C....
#P. major. ....G..C....T....C....G....
#E. japon. ....G..C....T....A....
#A. spini. ....G..C....C....
#T. tumif. ....C..G....T..C....
#R. sarba. ....G....T....C..C....C..C....
#P. oliva. ....C....A....T....T....C..A..T.G.T....T....C..T.T....A

#S. macro. GCA CCC TCC AAC ATT TCC GTT TGA TGA AAT TTT GGA TCC CTC CTC GGT CTC TGC TTA ATT TCC GAA
#S. berda. ....G..T....C....
#S. latus. ....T....T....C....
#P. major. ....A..G..T....T..C....C..T..A....C....G....C..T..G
#E. japon. ....G..A..G..T....T..C....C....A..T..C....C....C..T..G
#A. spini. ....A..A..T....T..C....C....C..C....C....C....C..T..G
#T. tumif. ....T....C..T....C..C..T....C..C..T....C....C.G....T..G
#R. sarba. ....C....T....A..C....T....T.A..G..G..A..T.C....C..T...
#P. oliva. ....C....T..T..C..T..C....C....G....T.A..A..A..T....C.A..G

#S. macro. CTT CTC ACA GGA CTA TTT CTT GCT ATA CAC TAT ACT TCC GAT ATT GCC ACA GCC TTC TCC TCC GTA
#S. berda. ....T....G....C....C....T..T....
#S. latus. ....T....G....C....C....T..T....
#P. major. A.C....C....C....C..A..C....T....T..T..T..T
#E. japon. ..C....G..C....C....T....C..G..C....T..T..T..T
#A. spini. ..C..T....C..C....C....C..C....C....C....T..T..T..T
#T. tumif. ..C....C....C..C....C....C..A..C....AA....T..T..T..T
#R. sarba. ..C..A..G..G....G....A..T....C..T....T..T..T..C
#P. oliva. A..T.A..C..C....C.T.A..C....C..C..A..C..C..A..T....A.T....G

#S. macro. GCT CAC ATT TGC CGA GAT GTA AAT TAT GGG TGG CTA ATC CGA AAC CTT CAC GCC AAC GGA GCA TCT
#S. berda. ..C..T....C..T....C....T..A....T....T....G....
#S. latus. ..C....C..T....C....C..C..A..T....C..T..C..T..T..T..G....
#P. major. ..C..T..C....C....C..C..C..A..T....C..T..C..T..T..T..T..T
#E. japon. ..C..T..C..T....C....C..C..A..T....C..T..C..T..T..T..T..T
#A. spini. ..C....C....C....C..C..C..A..T....T..T..T..T..T..T..C
#T. tumif. ..A....C..T....C....C..A..A..T....C....C..T..T..T..C
#R. sarba. ..C..T..C....C....C..C..A..A..T....C....T..T..T..T..C
#P. oliva. ..C....C....G..C....C..C..T..A..T....A.C....T..T..C....A

#S. macro. TTC TTC TTC ATC TGT ATT TAC CTT CAC ATC GGA CGG GGA CTC TAC TAC GGC TCC TAC CTT TAT AAA
#S. berda. ..T....T....T....T....T..G....C....
#S. latus. ....T....T....T....T....G..T....G....T....C....
#P. major. ....T....T..C..C..T....A..C....T..T..T..C....
#E. japon. ....T....C..C....C....C..T....C..T..T..T..T..C....
#A. spini. ....T..T....C....C....C..A..G....T....T..T..T..C....
#T. tumif. ....T..T..T....G..A..A..T....C....C..T..T..C....
#R. sarba. ....T....T..C..C..T..G....T..A..C....T..T..T..C..C....
#P. oliva. ....T....C....C....C..C..A..T..A....T..T..G....

#S. macro. GAA ACC TGA AAT ATT GGG GTC GTT CTC CTT CTC CTA GTT ATA GCA ACA GCC TTT GTA GGT TAT GTA
#S. berda. ....A....G....C....
#S. latus. ....C..C..A..T..A....C..T..C....C....G....
#P. major. ....A....T....C....C..T..T..C....C....C..C..T
#E. japon. ....G..A....C....T....C....C..T..A....C....C..C..C..C
#A. spini. ....A....C..C..T....C..T..C....T..A....C....C..C..C..C
#T. tumif. ....A....C..C..T..C..T..C....T..A....G....T..C....C..C..T
#R. sarba. ....A....C....T....T....G..T....G..T....C....C..C..C..C
#P. oliva. ....A....G..C....A..C..G..G..T..C..A..G..ATG..C....T....C..C..C

#S. macro. CTC CCT TGA GGA 408
#S. berda. ....T
#S. latus. ....
#P. major. ..T
#E. japon. ..T
#A. spini. ..T..A
#T. tumif. ....A
#R. sarba. ..T..A
#P. oliva. ..T..C

```

“.”表示此位点与最上排核苷酸相同 Dot indicates the nucleotide identity to that at the top row

图2 鲷科鱼类和牙鲆线粒体细胞色素b 408 bp的DNA序列

Fig. 2 Mitochondrial cytochrome b 408 bp DNA sequences of Sparidae fishes and *P. olivaceus*

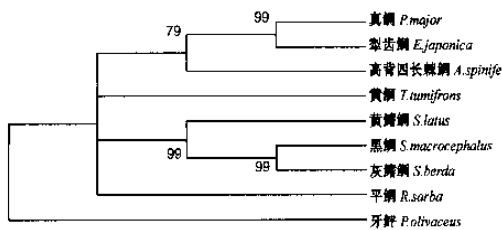


图3 NJ法构建鲷科鱼类的分子系统树,数字表示 Bootstrap 1 000 个循环的置信度

Fig. 3 Molecular phylogenetic tree of Sparidae fishes constructed by NJ method, the numbers showing the confident values of Bootstrap 1 000

参考文献:

- [1] Brown W M. Evolution of animal mitochondrial DNAs [A]. Nei M, Koehn R K. Evolution of genes and proteins [C]. Sunderland MA: Sinauer, 1983. 62-88.
- [2] Harrison R G. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology [J]. Trend Ecol evol, 1989, 4: 6-11.
- [3] Meyer A, Koehler T D, Basasibwaki P, et al. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences [J]. Nature, 1990, 347: 550-553.
- [4] Zhang Y P, Ryder O A. Phylogenetic relationships of bears (the Ursidae) inferred from mitochondrial DNA sequences [J]. Mol Phylogenet Evol, 1994, 3: 351-359.
- [5] Yoder A D, Cartmill M, Ruvolo M, et al. Ancient single origin for Malagasy primates [J]. Proc Natl Acad Sci USA, 1996a, 93: 5122-5126.
- [6] 孟庆闻, 苏锦祥 (译). 世界的鱼类 [M]. 北京: 农业出版社, 1982.
- [7] 成泰庆, 郑葆珊. 中国鱼类系统检索 [M]. 北京: 科学出版社, 1987.
- [8] 伍汉霖, 邵广昭, 赖春福. 拉汉世界鱼类类名典 [M]. 台北: 水产出版社, 1999.
- [9] 黄宗国. 中国海洋生物种类与分布 [M]. 北京: 海洋出版社, 1994.
- [10] Wang W, Shi L M. An improved method for isolation of animal mitochondrial DNA [J]. Zoological Research, 1993, 14(2): 197-198.
- [11] 孔庆鹏, 罗静, 黄顺友, 等. 从线粒体细胞色素 b 探讨长臂鲃属三个种分类与进化的关系 [J]. 遗传, 2000, 22(6): 379-384.
- [12] Kumar S, Tamura K, Nei M. MEGA: Molecular evolutionary genetics analysis [CP/DK]. Ver 1.02. Pennsylvania: University Park, the Pennsylvania State University, 1994.
- [13] Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap [J]. Evolution, 1985, 39: 783-791.
- [14] Cronin A M. Mitochondrial DNA phylogeny of deer (Cervidae) [J]. J Mamm, 1991, 72: 553-566.

Molecular phylogeny of mitochondrial cytochrome b gene sequences from four Sparidae fishes

JIANG Shi-gui¹, LIU Hong-yan^{1,2}, SU Tian-feng¹, GONG Shi-yuan²

(1. South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510300, China;

2. Fisheries College, Huazhong Agriculture University, Wuhan 430070, China)

Abstract: The 430 bp fragments of Cytochrome b gene were obtained by PCR method from the mitochondrial DNA of *Sparus latus*, *Sparus macrocephalus*, *Pagrosomus major* and *Rhabdosargus sarba*, and a mitochondrial Cyt b DNA of 408 bp sequence was successfully sequenced by ClustalX homologous sequence analysis. By comparing with the same DNA fragments of *Sparus berda*, *Taius tumifrons*, *Evynnis japonica* and *Argyrops spinifer* recorded in the GeneBank, the sequences of 408 bp DNA have 126 nucleotide sites substituted (30.7%), and the transition times are more than the transversion times; most of the base substitution happens in the third locus of the codon. The closest relationship is between *S. macrocephalus* and *S. berda*, and *T. tumifrons* is highly divergent from *R. sarba*. The molecular phylogenetic tree constructed by NJ method of MAGE2.1 indicates the eight fishes are divided into four groups, in which *R. sarba*, *T. tumifrons* and the species of *Sparus* are divided into three branches, respectively, which is consistent with the conclusion of morphological analysis. But *P. major*, *E. japonica* and *A. spinifer* are clustered together, which is different from the result of traditional taxonomy which divided the three species into three different genera. The purpose of this study is to clarify the taxonomy and phylogenetic evolution relationship of Sparidae fishes from the view of molecular biology.

Key words: Sparidae; cytochrome b; molecular phylogeny