

文章编号: 1000-0615(2017)03-0321-09

DOI: 10.11964/jfc.20160410359

5种鲷科鱼类核糖体ITS1序列比较

龚理^{1,2,3}, 时伟², 杨敏^{2,3}, 司李真^{2,3}, 孔晓瑜^{2*}

(1. 广西科学院广西红树林研究中心, 广西红树林保护与利用重点实验室, 广西北海 536000;

2. 中国科学院南海海洋研究所, 热带海洋生物资源与生态重点实验室, 广东广州 510301;

3. 中国科学院大学, 北京 100049)

摘要: 核糖体基因在很长一段时间内被认为严格遵循协同进化方式, 但是在很多种类中都发现了明显的序列多态性, 表明其是非协同进化。为了检测鲷科鱼类核糖体内转录间隔区1 (ITS1)序列是否存在多态性, 并探究其能否作为种类鉴定的分子标记, 本研究克隆获得了5种鲷科鱼类共118条ITS1全序列。结果显示, 眼斑豹鲷具有两种差异显著的片段类型, 表明其在基因组中遵循非协同进化方式; 而在其余4种鲷科鱼类中均没有发现序列多态性, 表明其为协同进化。序列分析显示ITS1具有明显的种间长度异质性, 最短的序列出现在蛾眉条鲷(412 bp), 最长的为眼斑豹鲷(585 bp)。碱基分析显示ITS1序列在5种鲷科鱼类中都呈现出相同的趋势: C>G>A>T, 且GC含量为69.5%, 远高于AT含量。聚类分析显示除眼斑豹鲷外, 所有鲷类均单独聚为一支, 种类区分度非常明显, 表明ITS1序列能够作为种类鉴定的分子标记。但是眼斑豹鲷的一个克隆和东方箬鲷聚为一支, 这种序列多态性对种类鉴定产生了干扰, 因此用具有多态的ITS1序列作为分子标记时一定要有足够的克隆数量, 避免因数据不充分而得到不正确的结论。

关键词: 鲷科; 内转录间隔区; 种类鉴定; 分子标记; 假基因; 非协同进化; 聚类分析
中图分类号: Q 522; S 917.4 **文献标志码:** A

真核生物核糖体DNA (rDNA)是由高度串联重复序列组成的多基因家族, 真核生物基因组中一般有几十个甚至上万个拷贝, 分布于一条或者多条染色体上。每个转录单元由18S、5.8S、28S rRNA编码基因和内部转录间隔区(internal transcribed spacer, ITS)组成。ITS包括ITS1和ITS2, 前者位于18S与5.8S之间, 后者位于5.8S与28S之间。一般核糖体18S、5.8S和28S序列比较保守, ITS区序列变异相对较大^[1]。

由于核糖体ITS1属于非编码区, 在进化中承受的选择压力较小, 进化速率相对较快, 遗传信息位点相对丰富, 在种内基本趋于相似, 而在种间则表现出显著的差异, 加之其两端的18S和5.8S序列高度保守, 便于设计引物进行扩增,

因而除了线粒体条形码(Barcoding)常用于种类鉴定之外, 利用核糖体ITS1序列作为种类鉴定的研究也非常广泛, 在动植物、真菌等类群都有报道^[2-5], 但在鱼类种类鉴定中报道相对较少^[6-8], 如Pérez等^[6]扩增了12种鳕(*Merluccius* sp.)的ITS1序列, 直接利用酶切图谱或者依靠测序能够有效将其鉴别; Kumar等^[7]研究了5种印度鲷25个个体的核糖体基因, 利用ITS1序列构建的系统树能有效将不同种类区分开。

核糖体串联重复ITS1能够成为种类鉴定的分子标记, 是基于以下一个重要的前提, 即ITS1通过染色体的不等交换和基因转换, 使不同ITS1拷贝基本趋于一致或完全一致^[9-10]。这样, ITS1在同一基因组中所有拷贝单元都趋于一致, 而

收稿日期: 2016-04-14 修回日期: 2016-09-06

资助项目: 国家自然科学基金(31272273); 广西红树林保护与利用重点实验室开放基金(GKLMC-201502)

通信作者: 孔晓瑜, E-mail: xykong@scsio.ac.cn

在种间或者更高阶元之间则表现出差异, 适合作为种类区分的分子标记。但是, 近年来很多研究发现, ITS1拷贝在许多种内甚至个体内存在明显的多态性, 例如在真菌、植物、海绵、双壳类、甲壳动物、鱼等方面都有报道^[11-16], 有些研究还表明这种现象主要是由于假基因造成的^[12, 14, 17-19]。ITS1序列多态性可能会给种类鉴定产生一定的影响, 不能有效区分不同种类, 同时相应的系统演化分析也可能会因此而无法构建出可靠的进化树^[16, 19-21]。如龚理等^[21]在分析雌牙鲆(*Paralichthys olicaceus*)、雄大西洋牙鲆(*P. dentatus*)及其杂交子一代之间的ITS1序列特征时, 发现只有大西洋牙鲆具有一种序列类型(X型), 而牙鲆有两种类型(X和Y型), 子代有三种类型(X、Y及重组类型)。利用这些ITS1序列并对它们进行聚类分析显示, 相同类型的序列混在一起聚为一支, 不能作为区分这三个种的分子标记。

目前对鲽形目(Pleuronectiformes)鱼类核糖体基因的研究鲜有报道, 仅在舌鳎科(Cynoglossidae)^[22]、鲽科(Pleuronectidae)^[12]和牙鲆科(Paralichthyidae)^[21]各有一例报道。鳎科(Soleidae)作为鲽形目鱼类最丰富的类群(35属175种), 很多种类都具有非常重要的经济价值, 但是迄今关于鳎科鱼类核糖体序列的报道仍为空白。本团队在研究鳎科的眼斑豹鳎(*Pardachirus pavoninus*) ITS1序列时, 发现其存在两种差异显著的片段类

型, 呈现明显的多态性。为了检测其他鳎科鱼类ITS1是否存在序列多态性, 并初步探究其能否在多态状态下作为种类鉴定的分子标记, 又选择了3个近缘属的种类作为研究对象, 包括条鳎属(*Zebrias*)的2个种类[带纹条鳎(*Zebrias zebrius*)和蛾眉条鳎(*Z. quagga*)]以及箬鳎属(*Brachirus*)的东方箬鳎(*B. orientalis*)和鳎属(*Solea*)的卵鳎(*S. ovata*)共4个种类, 进一步分析比较了这5个种类的序列特征, 为今后该类群的分子系统分析提供科学依据。

1 材料与方法

1.1 实验材料

眼斑豹鳎、东方箬鳎、带纹条鳎、蛾眉条鳎和卵鳎的采集地详见表1。每个种取一个个体, 于背部取少许肌肉保存于95%的酒精中待用。

1.2 DNA提取、PCR扩增和测序

取样品肌肉约30 mg切碎, 置于1.5 mL的离心管, 加入200 mL匀浆缓冲液和20 mL蛋白酶K, 至肌肉组织完全裂解, 然后用海洋动物组织基因组提取试剂盒提取DNA; 用双蒸水溶解基因组DNA, 保存于-20 °C冰箱中。根据已有近缘种类的18S及5.8S序列, 在保守位点设计通用引物用于扩增ITS1片段。正向引物Z-18S-1720: 5'-TCGCTACTACCGATTGGATGGTTTA-3'; 反向引物F-5.8S-150: 5'-AGCGACCCTCAGA

表1 5种鳎科鱼类ITS1序列多态特征信息表

Tab. 1 The polymorphism information of ITS1 sequence from five Soleidae species

种名及缩写 species	采集地 location	克隆数/个 no. of clone	基因型/个 genotype	多态性指数(Hd) haplotype diversity	序列 长度/ bp length	CS/ 个	VS/ 个	PS/ 个	微卫星 SSR	碱基含量/% base content				
										T	C	A	G	GC
眼斑豹鳎 <i>P. pav</i> (A)	海南三亚	36	29	0.9794	561~585	563	25	7	(CT) ₁₀₋₁₅ 、(CCG) ₅ 、 (CGTG) ₄ 、(CCGG) ₂	15.3	39.8	14.5	30.4	70.2
眼斑豹鳎 <i>P. pav</i> (B)	海南三亚	1	1		433				(CT) ₁₉	15.5	35.8	17.6	31.2	67.0
东方箬鳎 <i>B. ori</i>	广东湛江	14	8	0.8901	427~435	431	7	4	(CT) ₁₉₋₂₁ 、(GTA) ₃ 、 (AGAGGGC) ₂ 、 (CTCCACC) ₂	15.4	37.5	17.5	29.6	67.1
带纹条鳎 <i>Z. zeb</i>	浙江台州	19	5	0.5263	418	414	4	1	(CT) ₃ 、(GC) ₃ 、(GA) ₃ 、 (TCCCC) ₃	13.3	38.6	18.2	29.9	68.5
蛾眉条鳎 <i>Z. qua</i>	广东湛江	23	22	0.9960	412~441	360	80	47	(GGGGA) ₂ 、(GCGCCGG) ₂	10.8	37.6	20.0	31.6	69.2
卵鳎 <i>S. ova</i>	广东深圳	25	9	0.8067	520	514	9	5	(GCCC) ₃ 、(GGGCGC) ₂ 、 (AGAAAG) ₃	13.8	40.3	14.2	31.8	72.1
平均 average										13.6	38.8	16.9	30.7	69.5

注: *P. pav* (A)和*P. pav* (B) 分别代表眼斑豹鳎A、B两种类型序列。CS代表保守位点, VS代表变异位点, PS代表简约位点, 下同

Notes: *P. pav* (A) and *P. pav* (B) denotes Type A and B in *P. pavoninus*, respectively. CS denotes conserved sites, VS denotes variable sites, and PS denotes parsimonious informative sites, the same below

CAGGCGTAG-3'。反应总体积为25 μ L, 其中包括2.5 μ L 10 \times 缓冲液, 2 μ L MgCl₂(25 mmol/L), 2 μ L dNTP (分别为2.5 mmol/L), 每个引物1 μ L (10 μ mol/L), 1单位rTaq酶, 1 μ L模板DNA (50 ng/ μ L), 灭菌双蒸水补足至25 μ L。根据以往研究结果^[23-25], 一定浓度(一般不超过10%)的变性剂二甲基亚砜(DMSO)能够提高扩增效率, 因此本研究在常规PCR基础上添加了8%的DMSO。PCR反应程序: 94 $^{\circ}$ C预变性3 min, 94 $^{\circ}$ C变性1 min, 50 $^{\circ}$ C退火50 s, 72 $^{\circ}$ C延伸50 s, 进行35个循环后72 $^{\circ}$ C延伸10 min。扩增产物使用1%琼脂糖凝胶电泳检测, 产物用GenClean柱式琼脂糖凝胶DNA回收试剂盒纯化回收后与pMD 19-T载体连接, 转化至大肠杆菌DH5 α 。随机挑取单克隆测序, 选择峰图准确清晰的序列用于分析。

1.3 数据处理与系统发育分析

测定的序列用Codoncode Aligner软件(vers. 3, CodonCode Corporation, USA)进行拼接, 然后BLAST (<http://www.ncbi.nlm.gov/BLAST/>)检索, 扩增片段两侧的18S和5.8S序列与近缘种均具有极高的相似性, 证明扩增产物为目的片段。用CLUSTAL X^[26]软件对序列进行多重比对, 辅以BioEdit^[27]软件进行手工校准。通过MEGA 5.0^[28]软件统计碱基含量、保守位点、变异位点和简约位点, 采用Kimura双参数模型计算遗传距离, 用DnaSP^[29]软件检测单克隆序列的基因型以及多态性指数, 用SSRHunter3.1^[30]软件查找微卫星序列。因为目前鲷形目中仅公布了石鲷(*Kareius bicoloratus*)的ITS1序列, 因此选择石鲷作为外类群, 在MEGA 5.0中采用最大似然法(maximum likelihood, ML)和邻接法(neighbor-joining, NJ)构建5种鲷科鱼类的分子系统进化树, 置信度用自举分析(Bootstrap)进行1000次重复检验。

2 结果

2.1 ITS1序列长度变异、微卫星与碱基组成特征

分别获得了眼斑豹鲷、东方箬鲷、带纹条鲷、蛾眉条鲷和卵鲷5种鲷类的37、14、19、23、25个单克隆序列。去除两侧的18S (162 bp)和5.8S (158 bp)序列, 5种鲷科鱼类的ITS1长度变异

非常大, 具有明显的种间甚至个体内长度差异(表1)。种间长度异质性尤为明显, 相差173 bp, 最短的序列出现在蛾眉条鲷中(412 bp), 最长的为眼斑豹鲷(585 bp)。5种鲷类中有3种存在比较明显的个体内长度差异, 眼斑豹鲷序列长度差异最大, 为152 bp (433~585 bp); 其次是蛾眉条鲷和东方箬鲷, 分别相差29 bp (412~441 bp)和8 bp (427~435 bp)。而在带纹条鲷和卵鲷中, 则不存在个体内的长度差异, 分别为418和520 bp。

比较后还发现5种鲷科鱼类ITS1序列在不同位点存在不同类型的微卫星序列, 序列长度为2~7 bp, 重复次数2~5次(表1); 但是在眼斑豹鲷和东方箬鲷序列相同位点发现大量微卫星(CT)_n, 重复次数分别为10~19和19~21, 而其余3种鱼类在相应位点则未出现该微卫星序列。进一步分析了序列长度与微卫星之间的关系, 结果显示, 除眼斑豹鲷外, 同一种类不同序列长度差异主要是由微卫星拷贝次数不同引起的, 眼斑豹鲷序列长度差异则是由于具有两种完全不同类型序列引起的; 而不同种类之间的序列长度异质性则是由不同位点的插入/缺失造成的(图1)。

碱基分析显示5种鲷科鱼类所有ITS1序列中均呈现C>G>A>T的趋势(眼斑豹鲷中A<T), 含量最高的为C(平均38.8%), 其次为G(30.7%), A(16.9%), 含量最低为T(13.6%); 所有序列中GC含量(69.5%)远高于AT含量(30.5%), 这一结果与其他研究中ITS1序列高GC含量的特点相类似^[11, 31-32]。

2.2 序列多态性分析

将5种鲷科鱼类ITS1序列比对后发现眼斑豹鲷具有两种差异显著的片段类型, 其余4种鱼类ITS1序列均只有一种类型, 且不同种类之间序列差异非常显著, 存在大量的插入和缺失位点, 序列相似性低(图1)。眼斑豹鲷两种类型分别命名为A和B型, 其中A类型序列占绝大多数(36/37), 长度为561~585 bp, GC含量为69.9%~70.5%(平均70.2%), 二级结构最小自由能为-1757.6 kJ/mol; 而B类型仅有一条序列(标记为P. pav01), 长度仅为433 bp, 比A类型短了至少128 bp, GC含量为67.0%, 最小自由能为-1412.3 kJ/mol。眼斑豹鲷B类型序列(P. pav01)和东方箬鲷(B. ori)序列具有极高的相似性(约96.2%)(图1), 通过比较它们两侧的18S和5.8S序列可以排除相互污染的可能性。

```

P.pav25 --ATTGCTCC GGGCAAAGCG GCCACGCCGG CGAGTCTCCG GCTTCCCGCG CGCCGCGCGC GCGCTC-CG AGCCGCCCCG CTCCC-----AGGGCCG ATGGACTCGG CGCGGTGAGC [480]
P.pav29 --ATTGCTCC GGGCAAAGCG GCCACGCCGG CGAGTCTCCG GCTTCCCGCG CGCCGCGCGC GCGCTC-CG AGCCGCCCCG CTCCC-----AGGGCCG ATGGACTCGG CGCGGTGAGC [480]
P.pav31 --ATTGCTCC GGGCAAAGCG GCCACGCCGG CGAGTCTCCG GCTTCCCGCG CGCCGCGCGC GCGCTC-CG AGCCGCCCCG CTCCC-----AGGGCCG ATGGACTCGG CGCGGTGAGC [480]
P.pav33 --ATTGCTCC GGGCAAAGCG GCCACGCCGG CGAGTCTCCG GCTTCCCGCG CGCCGCGCGC GCGCTC-CG AGCCGCCCCG CTCCC-----AGGGCCG ATGGACTCGG CGCGGTGAGC [480]
Z.zeb01 ---TTTGGTCC GGG-GCCGCG TTCCCGCGGAG AGAGCGCCCC TCTCTCCACC GGGGGCTG-----GCCCCA CTCC-----AACGA CCGACACCAA [480]
Z.zeb05 ---TTTGGTCC GGG-GCCGCG TTCCCGCGGAG AGAGCGCCCC TCTCTCCACC GGGGGCTG-----GCCCCA CTCC-----AACGA CCGACACCAA [480]
Z.zeb15 ---TTTGGTCC GGG-GCCGCG TTCCCGCGGAG AGAGCGCCCC TCTCTCCACC GGGGGCTG-----GCCCCA CTCC-----AACGA CCGACACCAA [480]
Z.zeb03 ---TTTGGTCC GGG-GCCGCG TTCCCGCGGAG AGAGCGCCCC TCTCTCCACC GGGGGCTG-----GCCCCA CTCC-----AACGA CCGACACCAA [480]
Z.qua02 ---TTGCTCC GGGGGCCGCG CCGGAGACGG TGGGGAGGG GACCTCAACC CCTCCCTCT GTCTCCACCG GGGCGGCCCC CCCCC--A AAAAAAAAAA CCGGAACCTAG [480]
Z.qua08 ---TTGCTCC GGGGGCCGCG CCGGAGACGG TGGGGAGGG GACCTCAACC CCTCCCTCT GTCTCCACCG GGGCGGCCCC CCCCC--AA AAAAAAAAAA CCGGAACCTAG [480]
Z.qua05 ---TTGCTCC GGGGGCCGCG CCGGAGACGG TGGGGAGGG GACCTCAACC CCTCCCTCT GTCTCCACCG GGGCGGCCCC CCCCC--AA AAAAAAAAAA CCGGAACCTAG [480]
Z.qua09 ---TTGCTCC GGGGGCCGCG CCGGAGACGG TGGGGAGGG GACCTCAACC CCTCCCTCT GTCTCCACCG GGGCGGCCCC CCCCCCAA AAAAAAAAAA CCGGAACCTAG [480]
B.ori01 TCTCTCTCCC ACCTAGTCTG G-----GGGGTAGTA ----GTAGT CCC-----GGTGGCCCC CCCC-----ACCG GGCTTCCCC CCACCTCGGA [480]
B.ori08 TCTCTCTCCC ACCTAGTCTG G-----GGGGTAGTA ----GTAGT CCC-----GGTGGCCCC CCCC-----ACCG GGCTTCCCC CCACCTCGGA [480]
B.ori02 TCTCTCTCCC ACCTAGTCTG G-----GGGGTAGTA ----GTAGT CCC-----GGTGGCCCC CCCC-----ACCG GGCTTCCCC CCACCTCGGA [480]
B.ori05 TCTCTCTCCC ACCTAGTCTG G-----GGGGTAGTA ----GTAGT CCC-----GGTGGCCCC CCCC-----ACCG GGCTTCCCC CCACCTCGGA [480]
▲P.pav01 TCTCTCTCCC ACCTAGTCTG G-----GGGGTAGTA ----GTAGT CCC-----GGTGGCCCC CCCC-----ACCG GGCTTCCCC CCACCTCGGA [480]
S.ova01 CAACCTCCCC ACCAAACCTG GCTCTAGTC- TGGGCTACTA CGTCCCGGGA CCCCCTCCCC GCC--C-ACG GGGAGACCCG GCCTG-----GACCG TGCCCCCGCC CCACACTCGG [480]
S.ova06 CAACCTCCCC ACCAAACCTG GCTCTAGTC- TGGGCTACTA CGTCCCGGGA CCCCCTCCCC GCC--C-ACG GGGAGACCCG GCCTG-----GACCG TGCCCCCGCC CCACACTCGG [480]
S.ova02 CAACCTCCCC ACCAAACCTG GCTCTAGTC- TGGGCTACTA CGTCCCGGGA CCCCCTCCCC GCC--C-ACG GGGAGACCCG GCCTG-----GACCG TGCCCCCGCC CCACACTCGG [480]
S.ova04 CAACCTCCCC ACCAAACCTG GCTCTAGTC- TGGGCTACTA CGTCCCGGGA CCCCCTCCCC GCC--C-ACG GGGAGACCCG GCCTG-----GACCG TGCCCCCGCC CCACACTCGG [480]

```

图1 5种鲷科鱼类ITS1序列比对(部分)

▲指眼斑豹鲷B类型序列

Fig. 1 The alignment of ITS1 sequences of five Soleidae species (partial)

▲denotes *P. pavoninus* Type B sequence

5种鲷科鱼类ITS1序列多态性分析显示多态性指数(Hd)最高的为蛾眉条鲷(0.9960),其次为眼斑豹鲷(0.9802)、东方箬鲷(0.8901)、卵鲷(0.8067),最低为带纹条鲷(0.5263)。5种鲷科鱼类序列中,带纹条鲷最为保守,19条序列仅有4个变异位点;除去眼斑豹鲷B类型序列,变异最大的是蛾眉条鲷,23条序列中有80个变异位点。眼斑豹鲷序列变异最大,存在198个变异位点;但A类型的36条序列仅有25个变异位点,由此可以看出该种类不同类型序列是造成序列多态性显著的原因。

2.3 遗传距离分析

对种间及种内不同基因型之间的遗传距离进行分析,结果显示种间最大的遗传距离出现在眼斑豹鲷A类型和卵鲷之间(0.475),最小遗传距离出现在眼斑豹鲷A类型和带纹条鲷之间(0.218),而不是出现在同属的带纹条鲷和蛾眉条鲷之间(0.277),说明ITS1序列种间遗传距离不足以反映这5种鲷类之间的亲缘关系;至于种内不同基因型之间的遗传距离,最高出现在眼斑豹鲷,高达0.429,远大于其他4种鲷类,而其他4种鲷类种内遗传距离最大为蛾眉条鲷,也只有0.015(与眼斑豹鲷相差约28倍),最小为带纹条鲷和卵鲷(0.003);但是除去眼斑豹鲷特异的B类型序列,A类型种内遗传距离仅为0.002,甚至低于最小的带纹条鲷和卵鲷,说明眼斑豹鲷具有的两种完全不同的序列类型是导致其种内遗传距离非常高的原因(表2)。

2.4 系统发育分析

为了探究ITS1序列能否作为种类鉴定的分

表2 5种鲷科鱼类种内(对角线上)及种间(对角线下)K2P遗传距离

Tab. 2 The Kimura 2-parameter genetic distance of within (on diagonal) and between species (below diagonal) in five Soleidae species

种类及缩写 species abbr.	1	2	3	4	5	6
1.眼斑豹鲷 <i>P. pav</i> (A)	0.002					
2.眼斑豹鲷 <i>P. pav</i> (B)	0.429					
3.带纹条鲷 <i>Z. zeb</i>	0.218	0.427	0.003			
4.蛾眉条鲷 <i>Z. qua</i>	0.328	0.411	0.277	0.015		
5.东方箬鲷 <i>B. ori</i>	0.438	0.011	0.427	0.413	0.005	
6.卵鲷 <i>S. ova</i>	0.475	0.237	0.417	0.434	0.238	0.003

子标记,从种内相同的基因型中选取一条序列作为代表,以石鲷作为外类群构建5种鲷科鱼类的最大似然树(ML)和邻接树(NJ)。两种方法得出的系统发育树拓扑结构一致,本研究仅以最大似然树进行系统分析(图2)。系统分析结果显示,所有种类明显分为5大支(外类群除外),除眼斑豹鲷(*P. pav*01)外,同种类的不同基因型都单独聚为一支,且均具有很高的置信度(≥99%)。*P. pav*01序列不与眼斑豹鲷其他的序列聚在一起,而是与东方箬鲷聚为一支,且置信度为100%。此外,眼斑豹鲷(*P. pav*)其他序列先与带纹条鲷(*Z. zeb*)聚为一支,再与蛾眉条鲷(*Z. qua*)聚为一支;卵鲷(*S. ova*)和东方箬鲷(*B. ori*)聚为一支,作为以上3个种类的姐妹支,最后5个类群合并成一大支。该结果与种间遗传距离完全吻合。

3 讨论

在研究的5种鲷科鱼类的ITS1序列中,只在

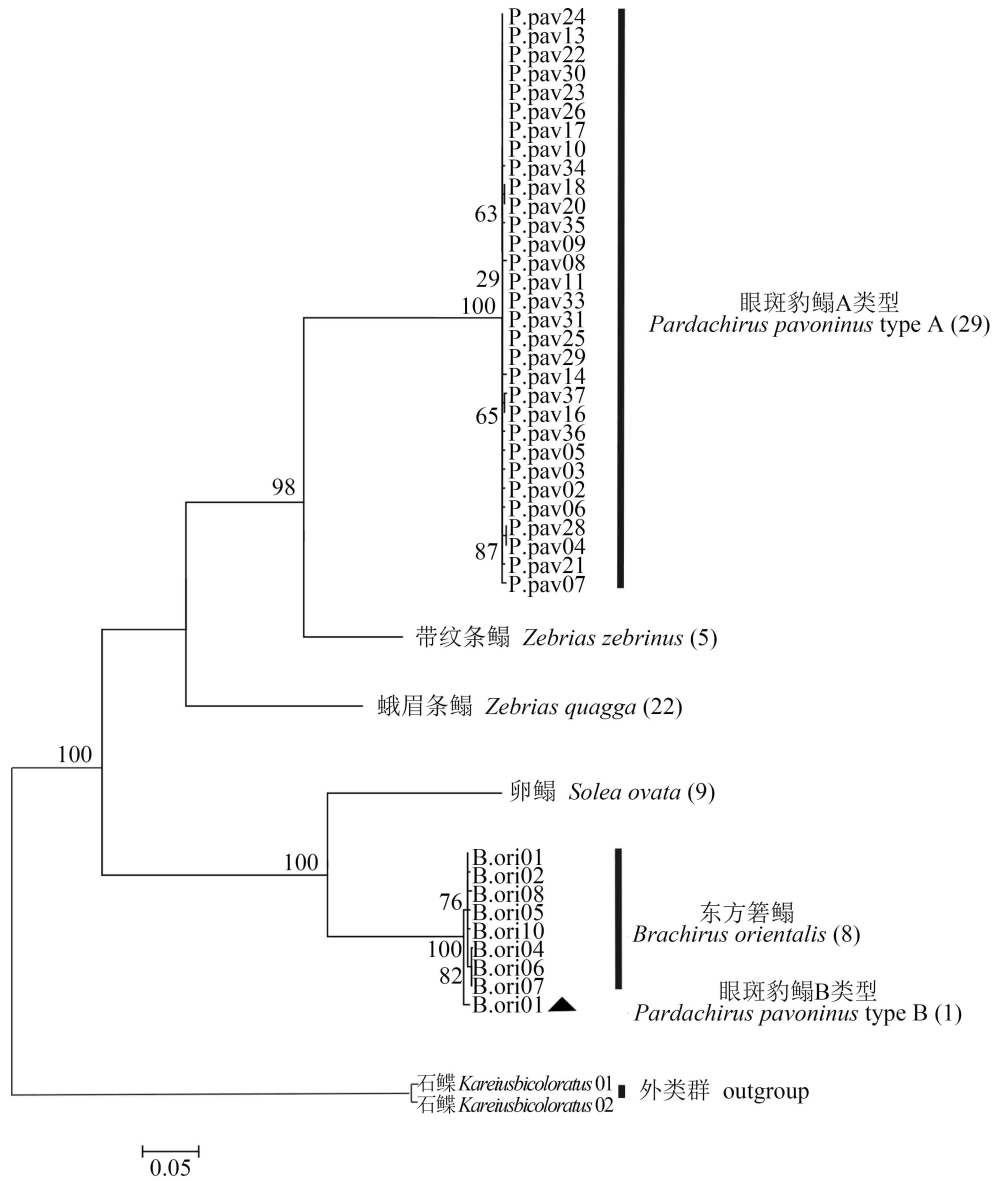


图 2 基于ITS1序列构建的5种鳎科鱼类最大似然树

括号内数字代表序列条数

Fig. 2 Maximum-likelihood (ML) tree based on ITS1 dataset of five Soleidae species

Numbers in the bracket indicate the amount of clones

眼斑豹鳎基因组中发现了A、B两种差异显著的类型,表明其遵循非协同进化方式,在其余4种鳎类基因组中都仅发现一种序列类型,表明其为协同进化。ML和NJ均显示除眼斑豹鳎B类型序列(P. pav 01)外,其他同种鳎类的不同基因型序列都单独聚为一支,种类区分度非常明显,证明可以用来作为鉴定这些种类的分子标记。然而,眼斑豹鳎特异的B类型与东方箬鳎聚为一支,而不是与同种类的其他序列聚为一支,这对该种类的鉴定造成一定干扰;但是,眼斑豹鳎这种特异的B类型序列出现的频率极低(1/37),

随机挑选克隆时选择它代表眼斑豹鳎进行物种鉴定的可能性非常低,考虑到本研究中118个单克隆中也只出现了一条特异类型序列,说明这种现象非常少见。根据本团队目前掌握的资料,具有多种不同类型的种类都是按种聚支,本研究中这种分散聚支的现象非常罕见^[21];因此认为本研究中ITS1序列能够作为种类鉴定的分子标记。即便如此,在利用核糖体ITS1序列作为种类鉴定时,首先要了解其进化方式,特别是出现多态类型或者特异的类型时,一定要有大量的单克隆序列进行比较,不能用少量序列来代表该

物种信息, 避免因数据不全得到错误的结论。

Hebert等^[33]发现线粒体CO I序列在种类间的遗传差异通常大于0.02, 且显著大于种内遗传距离(10倍以上), 首次提出将CO I序列作为种类鉴定的分子标记。这一观点在许多种类鉴定中得到了很好的验证^[34-36], 如在鲷科鱼类种类鉴定中, 王淑英等^[34]计算了鲷科14属24个种的CO I序列的遗传距离, 发现种间遗传距离最小为0.037, 其余均在0.1左右, 种间遗传距离约为种内平均遗传距离的60倍左右, 符合上述标准, 由此表明CO I序列可以用于鲷科鱼类的种类鉴定。与线粒体CO I序列相比, 核糖体ITS1序列变异更大, 在种间表现出更大的异质性, 因此作为种类鉴定的分子标记, 它比CO I序列有着更灵敏的区分度。如本研究5种鲷科鱼类中, ITS1序列种间最小遗传距离和最大遗传距离分别为0.218和0.475(不考虑眼斑豹鲷B类型序列), 均远大于CO I种间遗传距离。此外, 与线粒体CO I序列相比, 核糖体ITS1作为种类鉴定的分子标记还有另外一个优势: CO I序列在不同种类中长度差异很小, 而ITS1序列长度在不同种类中差异却非常大, 如本研究中最长和最短的ITS1序列相差170 bp左右, 这样就可以省去测序过程, 直接通过凝胶电泳序列长度来区分不同种类, 既简化了种类鉴定的过程又节省了费用, 对今后的鱼类种类的鉴定具有重要的指导意义。

多种ITS1片段类型在其他种类甚至个体中也有发现, 很多研究明确指出假基因的存在是导致出现多种片段类型最主要的原因^[12, 14, 17-19]。假基因一般都具有较快的变异速率和较明显的序列差异、较低的GC含量、较不稳定的二级结构以及较低的最小自由能等特点^[37-38]。序列比对不难发现本研究中眼斑豹鲷(P. pav 01)代表的B类型序列在多处存在插缺位点, 且长度(433 bp)明显小于A类型序列(585 bp); 此外, A、B两个类型序列在平均GC含量(70.2% vs. 67.0%)和二级结构最小自由能方面(-1757.6 vs. -1412.3 kJ/mol)均符合A类型为真基因、B类型为假基因的特征。

Álvarez等^[39]总结了多种ITS类型序列进化的三种可能去向, 一是两种或多种类型的片段分别进入到不同的后代支系中, 有的支系可能存在其中一种或几种类型, 另一支系可能存在另外一些类型; 第二种是在进化过程中两种或多种类型均存在于后代支系中, 类型间保持比较

稳定的状态, 不发生交换重组; 第三种是两种或两种以上的类型之间发生较高频率的交换重组, 进而在后代支系中产生新的类型。本研究中眼斑豹鲷假基因序列(P. pav 01)和东方箬鲷相似性非常高(约96.2%), 而与眼斑豹鲷真基因序列相似度仅有43.5%, 根据Álvarez等^[39]的观点推测眼斑豹鲷和东方箬鲷有着较近的祖先, 在其祖先的基因组中可能存在两种ITS1片段类型(即眼斑豹鲷中A类型和B类型), 在进化过程中, 分化成两个支系, 其中A和B两种类型都进入眼斑豹鲷基因组中, 而东方箬鲷基因组中则只有一种类型。在自然选择压力下, 眼斑豹鲷B类型序列逐渐丧失功能, 变异成假基因, 目前所获得的和东方箬鲷序列极其相似的眼斑豹鲷假基因正好处于退化的过渡阶段。推测随着变异事件的积累, 眼斑豹鲷假基因类型可能会逐渐从基因组中消失, 这也从另一方面暗示了眼斑豹鲷假基因在现存的基因组中出现频率很低的原因。假基因虽然在复制过程中失去了功能, 但是保留了数百万年或者更早以前祖先功能基因的分子记录, 能够为系统进化研究提供宝贵的分子证据^[18, 35, 40-41]。本研究中通过比较核糖体真假基因来推断种类之间的亲缘及演化关系, 目前在形态分类和其他分子研究方面都没有报道, 这一发现将对今后该类群的系统进化分析有着重要的启示及参考价值。

参考文献:

- [1] Hillis D M, Dixon M T. Ribosomal DNA: molecular evolution and phylogenetic inference[J]. The Quarterly Review of Biology, 1991, 66(4): 411-453.
- [2] Wang Y P, Guo X M. ITS length polymorphism in oysters and its use in species identification[J]. Journal of Shellfish Research, 2008, 27(3): 489-493.
- [3] Li X W, Yang Y, Henry R J, et al. Plant DNA barcoding: from gene to genome[J]. Biological Reviews, 2015, 90(1): 157-166.
- [4] Blaali R, Kumar S, Nilsson R H, et al. ITS1 versus ITS2 as DNA metabarcodes for fungi[J]. Molecular Ecology Resources, 2013, 13(2): 218-224.
- [5] Gulyás G, Sramkó G, Molnár V A, et al. Nuclear ribosomal DNA ITS paralogs as evidence of recent interspecific hybridization in the genus *Ophrys* (Orchidaceae)[J]. Acta Biologica Cracoviensia Series

- Botanica, 2005, 47(2): 61-67.
- [6] Pérez M, Vieites J M, Presa P. ITS1-rDNA-based methodology to identify world-wide hake species of the genus *Merluccius*[J]. Journal of Agricultural and Food Chemistry, 2005, 53(13): 5239-5247.
- [7] Kumar R, Singh M, Kushwaha B, *et al.* Molecular characterization of major and minor rDNA repeats and genetic variability assessment in different species of mahseer found in North India[J]. Gene, 2013, 527(1): 248-258.
- [8] Presa P, Pardo B G, Martínez P, *et al.* Phylogeographic congruence between mtDNA and rDNA ITS markers in brown trout[J]. Molecular Biology and Evolution, 2002, 19(12): 2161-2175.
- [9] Hillis D M, Moritz C, Porter C A, *et al.* Evidence for biased gene conversion in concerted evolution of ribosomal DNA[J]. Science, 1991, 251(4991): 308-310.
- [10] Smith G P. Unequal crossover and the evolution of multigene families[J]. Cold Spring Harbor Symposia on Quantitative Biology, 1974, 38: 507-513.
- [11] Freire R, Arias A, Méndez J, *et al.* Sequence variation of the internal transcribed spacer (ITS) region of ribosomal DNA in *Cerastoderma* species (Bivalvia: Cardiidae)[J]. Journal of Molluscan Studies, 2010, 76(1): 77-86.
- [12] Xu J P, Zhang Q Q, Xu X F, *et al.* Intragenomic variability and pseudogenes of ribosomal DNA in Stone flounder *Kareius bicoloratus*[J]. Molecular Phylogenetics and Evolution, 2009, 52(1): 157-166.
- [13] Harris D J, Crandall K A. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda: Cambaridae): implications for phylogenetic and microsatellite studies[J]. Molecular Biology and Evolution, 2000, 17(2): 284-291.
- [14] Li Y, Jiao L, Yao Y J. Non-concerted ITS evolution in fungi, as revealed from the important medicinal fungus *Ophiocordyceps sinensis*[J]. Molecular Phylogenetics and Evolution, 2013, 68(2): 373-379.
- [15] Redmond N E, McCormack G P. Ribosomal internal transcribed spacer regions are not suitable for intra-or inter-specific phylogeny reconstruction in haplosclerid sponges (Porifera: Demospongiae)[J]. Journal of the Marine Biological Association of the United Kingdom, 2009, 89(6): 1251-1256.
- [16] Bayly M J, Ladiges P Y. Divergent paralogues of ribosomal DNA in eucalypts (Myrtaceae)[J]. Molecular Phylogenetics and Evolution, 2007, 44(1): 346-356.
- [17] Keller I, Chintauan-Marquier I C, Veltsos P, *et al.* Ribosomal DNA in the grasshopper *Podisma pedestris*: escape from concerted evolution[J]. Genetics, 2006, 174(2): 863-874.
- [18] Zheng X Y, Cai D Y, Yao L H, *et al.* Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in *Pyrus*[J]. Molecular Phylogenetics and Evolution, 2008, 48(3): 892-903.
- [19] Song H X, Gao S P, Jiang M Y, *et al.* The evolution and utility of ribosomal ITS sequences in Bambusinae and related species: divergence, pseudogenes, and implications for phylogeny[J]. Journal of Genetics, 2012, 91(2): 129-139.
- [20] Wei X X, Wang X Q, Hong D Y. Marked intragenomic heterogeneity and geographical differentiation of nrDNA ITS in *Larix potaninii* (Pinaceae)[J]. Journal of Molecular Evolution, 2003, 57(6): 623-635.
- [21] 龚理, 徐晖, 李军, 等. 褐牙鲆(♀)、夏鲆(♂)及其杂交子一代的ITS1序列特征分析[J]. 中国水产科学, 2015, 22(1): 17-23.
- Gong L, Xu H, Li J, *et al.* Characterization of the first internal transcribed spacer of ribosomal DNA in *Paralichthys olivaceus* (♀) and *P. dentatus* (♂)hybrids[J]. Journal of Fishery Sciences of China, 2015, 22(1): 17-23(in Chinese).
- [22] 徐晖, 李军, 孔晓瑜, 等. 6种舌鳎亚科鱼类ITS1序列长度多态性及系统分析[J]. 海洋与湖沼, 2008, 39(1): 35-41.
- Xu H, Li J, Kong X Y, *et al.* Phylogenetic relationship and length variation in the first ribosomal internal transcribed spacer of Cynoglossinae species[J]. Oceanologia et Limnologia Sinica, 2008, 39(1): 35-41(in Chinese).
- [23] Buckler IV E S, Holtsford T P. *Zea* systematics: ribosomal ITS evidence[J]. Molecular Biology and Evolution, 1996, 13(4): 612-622.
- [24] Buckler IV E S, Ippolito A, Holtsford T P. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications[J]. Genetics, 1997, 145(3): 821-832.
- [25] Gong L, Shi W, Yang M, *et al.* Non-concerted evolution in ribosomal ITS2 sequence in *Cynoglossus*

- zanzibarensis* (Pleuronectiformes: Cynoglossidae)[J]. Biochemical Systematics and Ecology, 2016, 66: 181-187.
- [26] Larkin M A, Blackshields G, Brown N P, *et al.* Clustal W and clustal X version 2. 0[J]. Bioinformatics, 2007, 23(21): 2947-2948.
- [27] Hall T A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT[J]. Nucleic Acids Symposium Series, 1999, 41(2): 95-98.
- [28] Tamura K, Peterson D, Peterson N, *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. Molecular Biology and Evolution, 2011, 28(10): 2731-2739.
- [29] Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data[J]. Bioinformatics, 2009, 25(11): 1451-1452.
- [30] 李强, 万建民. SSRHunter, 一个本地化的SSR位点搜索软件的开发[J]. 遗传, 2005, 27(5): 808-810.
Li Q, Wan J M. SSRHunter: development of a local searching software for SSR sites[J]. Hereditas (Beijing), 2005, 27(5): 808-810(in Chinese).
- [31] Chen C A, Chen C P, Fan T Y, *et al.* Nucleotide sequences of ribosomal internal transcribed spacers and their utility in distinguishing closely related *Perinereis* polychaets (Annelida; Polychaeta; Nereididae)[J]. Marine Biotechnology, 2002, 4(1): 17-29.
- [32] Booton G C, Kaufman L, Chandler M, *et al.* Evolution of the ribosomal RNA internal transcribed spacer one (ITS-1) in cichlid fishes of the Lake Victoria region[J]. Molecular Phylogenetics and Evolution, 1999, 11(2): 273-282.
- [33] Hebert P D N, Cywinska A, Ball S L, *et al.* Biological identifications through DNA barcodes[J]. Proceedings of the Royal Society of London- Series B: Biological Sciences, 2003, 270(1512): 313-321.
- [34] 王淑英, 时伟, 江金霞, 等. CO I 和16S rRNA基因序列在鰺科(Soleidae)鱼类种类鉴定中的适用性研究[J]. 热带海洋学报, 2014, 33(3): 57-65.
Wang S Y, Shi W, Jiang J X, *et al.* Applicability of mitochondrial CO I and 16S rRNA gene sequences in species identification of sole fish (Pleuronectiformes: Soleidae)[J]. Journal of Tropical Oceanography, 2014, 33(3): 57-65(in Chinese).
- [35] Steinke D, Zemplak T S, Boutillier J A, *et al.* DNA barcoding of pacific Canada's fishes[J]. Marine Biology, 2009, 156(12): 2641-2647.
- [36] Zhang J B, Hanner R. Molecular approach to the identification of fish in the South China Sea[J]. PLoS One, 2012, 7(2): e30621.
- [37] Bailey C D, Carr T G, Harris S A, *et al.* Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes[J]. Molecular Phylogenetics and Evolution, 2003, 29(3): 435-455.
- [38] Harpke D, Peterson A. 5. 8S motifs for the identification of pseudogenic ITS regions[J]. Botany, 2008, 86(3): 300-305.
- [39] Álvarez I, Wendel J F. Ribosomal ITS sequences and plant phylogenetic inference[J]. Molecular Phylogenetics and Evolution, 2003, 29(3): 417-434.
- [40] Bayly M J, Udovicic F, Gibbs A K, *et al.* Ribosomal DNA pseudogenes are widespread in the eucalypt group (Myrtaceae): implications for phylogenetic analysis[J]. Cladistics, 2008, 24(2): 131-146.
- [41] Brownell E, Krystal M, Arnheim N. Structure and evolution of human and African ape rDNA pseudogenes[J]. Molecular Biology and Evolution, 1983, 1(1): 29-37.

Comparative analysis of the first internal transcribed spacer of ribosomal DNA in five Soleidae species

GONG Li^{1,2,3}, SHI Wei², YANG Min^{2,3}, SI Lizhen^{2,3}, KONG Xiaoyu^{2*}

(1. Guangxi Key Lab of Mangrove Conservation and Utilization, Guangxi Mangrove Research Center, Guangxi Academy of Sciences, Beihai 536000, China;

2. Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China;

3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: For a long time, the ribosomal RNA gene was thought to conform to the paradigm of strict concerted evolution pattern. In fact, an increasing number of intraindividual and intraspecies variations have been discovered, indicating non-concerted evolution. In order to explore the polymorphism of the first internal transcribed spacer (ITS1) in the species of Soleidae in Pleuronectiformes and whether it is suitable as a marker for species identification, a total of 118 ITS1 sequences were determined from five soles. The results showed that two distinct types were found in *Pardachirus pavoninus* genome, suggesting a non-concerted evolution pattern, while concerted evolution was confirmed in other four soles because of absence of intraindividual polymorphism. Sequence analysis demonstrated that ITS1 had significant length heterogeneity, ranging from 412 bp in *Zebrias quagga* to 585 bp in *P. pavoninus*. The base content of sequences in all five soles was in the same trend with C>G>A>T and GC content (69.5%) was far higher than AT content. Except for one sequence of *P. pavoninus*, each of five soles clustered into one clade, respectively, suggesting ITS1 sequence was suitable for species identification in the present study. It is worth noting that the sequence of *P. pavoninus* clustered with *Brachiru orientalis*, suggesting that ITS1 sequence polymorphism interfered species identification, as a result, it is essential to obtain enough clones before using ITS1 sequence as the molecular marker for species identification in order to avoid incorrect conclusions.

Key words: Soleidae; internal transcribed spacer; species identification; molecular marker; pseudogene; non-concerted evolution; cluster analysis

Corresponding author: KONG Xiaoyu. E-mail: xykong@scsio.ac.cn

Funding projects: National Natural Science Foundation of China (31272273); Open Research Fund Program of Guangxi Key Lab of Mangrove Conservation and Utilization (GKLMC-201502)