

DOI: 10.12264/JFSC2021-0084

七种骨舌鱼核糖体序列特征及遗传发育分析

杨叶欣^{1,2}, 刘奕¹, 刘超¹, 宋红梅^{1,3}, 汪学杰¹, 牟希东¹

1. 中国水产科学研究院珠江水产研究所, 农业农村部休闲渔业重点实验室, 广东省现代休闲渔业工程技术研究中心, 广东 广州 510380;
2. 农业农村部都市农业重点实验室, 上海 200240;
3. 广东省水产动物免疫技术重点实验室, 广东 广州 510380

摘要: 利用核糖体 DNA 序列探讨了骨舌鱼科(Osteoglossidae)高阶元的分子系统发育关系。采用分子标记技术, 分析测定了现存的 7 种骨舌鱼核糖体基因, 获得了巨骨舌鱼(*Arapaima gigas*)、尼罗异耳骨舌鱼(*Heterotis niloticus*)、双须骨舌鱼(*Osteoglossum bicirrhosum*)、费氏骨舌鱼(*O. ferreirai*)、美丽硬仆骨舌鱼(*Scleropages formosus*)、乔氏硬仆骨舌鱼(*S. jardinii*)和硬仆骨舌鱼(*S. leichardti*)的 rDNA 全序列, 长度分别为 11714 bp、8957 bp、12057 bp、11556 bp、10377 bp、10724 bp 和 10725 bp, GC 碱基含量为 63.78%~66.13%。采用邻接法和贝叶斯推论法分别构建了系统进化树, 结果显示, 除 ITS1 外, 18S、ITS2、28S 和 18S+ITS1+5.8S+ITS2+28S 联合序列建树均与传统的分类相符合, 研究结果阐明了骨舌鱼科亚科、属和种间的分类关系, 核糖体 DNA 可作为骨舌鱼科的分子标记。

关键词: 骨舌鱼科; 核糖体基因; 18S rDNA; 5.8S rDNA; 28S rDNA; ITS; IGS; 系统发育分析

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2022)01-0013-15

核糖体是细胞内蛋白质合成的“加工厂”, 呈不规则的颗粒状, 在生物的生命活动中具有重要作用^[1]。真核生物细胞核中核糖体 RNA 基因(ribosomal DNA, rDNA)是高度串联的多拷贝重复序列, 位于染色体的核仁组织区。每个重复序列包括 3 个编码基因 18S rDNA、5.8S rDNA 和 28S rDNA, 以及位于 3 个编码基因之间的内转录间隔区(internal transcribed spacer, ITS1 和 ITS2)^[2]。在每个重复序列之间, 被不转录的 DNA 序列(non-transcribed spacer, NTS)以及在 18S 和 28S 两头的外转录间隔区(external transcribed spacers, 5'-ETS and 3'-ETS)分割, 这些序列被称为基因转录间隔区(intergenic spacer, IGS)^[3-4]。核糖体编码区保守性强, 而非编码区进化速度快, 变异率高, ITS1 和 ITS2 已经广泛应用于种质鉴定、遗传多

样性分析和遗传发育进化研究等领域^[5-10]。

骨舌鱼为骨舌鱼科(Osteoglossidae)鱼类的统称, 隶属硬骨鱼纲(Osteichthyes), 辐鳍鱼亚纲(Actinopterygii), 骨舌鱼目(Osteoglossiformes), 俗称“龙鱼”, 是一类古老的大型淡水鱼, 具有较高的经济价值、文化价值和科研价值^[11-12]。其祖先出现在中生代三叠纪(中生代第一纪), 距今 2.42 亿~2.37 亿年, 素有鱼类“活化石”之称。骨舌鱼科共有 2 亚科 4 属 11 种, 现存的鱼类只有 7 种, 分别是巨骨舌鱼(*Arapaima gigas*)、尼罗异耳骨舌鱼(*Heterotis niloticus*)、双须骨舌鱼(*Osteoglossum bicirrhosum*)、费氏骨舌鱼(*O. ferreirai*)、美丽硬仆骨舌鱼(*Scleropages formosus*)、乔氏硬仆骨舌鱼(*S. jardinii*)和硬仆骨舌鱼(*S. leichardti*)^[13-14]。其中美丽硬仆骨舌鱼在自然界中分为 3 个自然群体: 金

收稿日期: 2021-02-26; 修订日期: 2021-04-19。

基金项目: 农业农村部都市农业重点实验室开放基金项目(UA201910); 广州市科技计划项目(201904010409); 广东省现代农业产业技术体系创新团队建设项目(2021KJ150); 国家淡水水产种质资源库项目(FGRC18537)。

作者简介: 杨叶欣(1980-), 女, 助理研究员, 主要从事种质资源与生态安全研究. E-mail: yangyexin@163.com

通信作者: 牟希东, 副研究员, 主要从事种质资源与城市渔业研究. E-mail: muxd@prfri.ac.cn

龙鱼(gold arowana)、红龙鱼(red arowana)和青龙鱼(green arowana)^[15-16]。

目前,利用 AFLP、微卫星标记、转录组、线粒体和全基因组测序等技术对不同群体、不同种类的骨舌鱼遗传结构和系统发育关系进行了大量的研究^[17-23],但是对于骨舌鱼科鱼类核糖体 DNA 基因的序列结构和内部变异知之甚少。相对于线粒体基因主要通过卵细胞传递、只包含单亲的遗传特征而言,核糖体基因具有双亲的遗传特征^[24],其构建的系统发育树能更全面地反映出物种间的真实关系。且 ITS 和 IGS 区与编码区相比,在物种进化过程中受到的选择压力相对小,进化方向更自由,在属或种等低阶元水平上具有进化更快、信息量更丰富和多态性更高等优点,所以在动植物分类及多态性研究中已经有广泛应用^[5-9,25-28]。在本研究中,利用二代测序和三代测序相结合的方法,测定了现存的所有骨舌鱼类完整的核糖体基因全序列,并对其进行了遗传结构和系统发育分析,旨在为骨舌鱼类的系统进化研究提供了数据,同时丰富鱼类核糖体基因的多样性及进化方式等方面的研究。

1 材料与方法

1.1 样品采集和 DNA 提取

现存所有骨舌鱼科鱼类(表 1)采自国家淡水水产种质资源库保存的活体或基因资源。其中巨骨舌鱼样品为 2012 年保存的样本^[13];美丽硬仆骨舌鱼(金龙鱼、红龙鱼和青龙鱼)为 2009 年珠江水产研究所从马来西亚引进的个体。每种鱼各选取 1 尾,剪取适量鳍条,用 95%乙醇固定,保存于 -20 °C。参照 EZNA (OMEGA)试剂盒的方法提取总 DNA, -20 °C 保存备用。

1.2 文库构建和高通量测序

二代 Illumina 测序文库构建与测序:检测合格的 DNA 样品参照 Illumina DNA 文库构建标准流程(Illumina, California, USA),构建插入片段大小为 350 bp 的双末端测序文库。文库构建完成后以 qPCR 方法和 Agilent 2100 Bioanalyzer (Agilent Technologies, California, USA)进行质控。对质检合格的 DNA 文库采用 Illumina HiSeq 4000 (Illu-

mina, California, USA)高通量测序平台进行测序,测序策略为 PE150 (Pair-End 150),每例样品测序数据量不少于 1 Gb。

三代 Pacbio 测序文库构建和测序:等量混合每份骨舌鱼 DNA,对混合后的 DNA 进行质检。将质检合格的基因组 DNA 参照 Pacbio sequel 平台(Pacific Bioscience, California, USA)标准实验流程和方法进行建库和测序,流程简述如下:基因组 DNA 经 26 G Needle 片段化,使用 Blue Pippin 选择 10 kb 以上的片段,经末端修复和加 A 尾后,再在片段两端分别连接接头,制备 DNA 文库。库检合格,根据文库的有效浓度及数据产出需求运用 Pacbio sequel 平台(Pacific Bioscience, California, USA)进行测序。

1.3 核糖体基因组序列组装

利用 PacBio Sequel 第三代高通量技术测序,使用 Canu v1.8 (<https://github.com/marbl/canu>)组装骨舌鱼类的核糖体 DNA 序列^[29]。采用 BWA^[30]和 Samtools v1.8^[31]软件将二代测序数据比对到三代组装结果,并采用 Pilon v1.22^[32]对核糖体 DNA 序列作进一步修正,从而获得完整的骨舌鱼核糖体基因组。将拼接好的序列用 NCBI 的在线 Blast 软件,根据不同部分序列同源性确定边界。

1.4 序列组成分析及遗传距离分析

用软件 MEGA X 计算序列的碱基含量、保守位点、变异位点和简约位点,采用双参数模型(Kitumra-2-parameter-distance)进行遗传距离的测算^[33]。用 DNAMEN 分析序列的一致性。使用 SSRHunter3.1 软件查找 2~6 bp 的微卫星序列^[34]。通过 Tandem Repeats Finder 在线软件(<http://tandem.bu.edu/trf/trf.html>)查找串联重复序列^[35],不考虑长度少于 10 bp 的序列。

1.5 系统发育树的构建

利用 rDNA 不同分区序列和联合序列,分别应用邻接法(neighbor-joining, NJ)和贝叶斯推论法(Bayesian inference, BI)构建进化树,其中 NJ 利用 MEGA X 分析,置信度用 Bootstrap 进行 1000 次重复检验。运用 MrBayes3.2.1^[36]构建贝叶斯树,同时运行两个独立的分析,每个分析包含 4 条马尔科夫链(3 条热链,1 条冷链),分析从随机树起

表 1 现存骨舌鱼科鱼类数据信息
Tab. 1 Taxonomic status and ribosomal information of Osteoglossidae

项目 item	异耳鱼亚科 Heterotidinae			骨舌鱼亚科 Osteoglossinae			
	巨骨舌鱼属 <i>Arapaima</i>	异耳骨舌鱼属 <i>Heterotis</i>	骨舌鱼属 <i>Osteoglossum</i>		硬骨舌鱼属 <i>Scleropages</i>		
	巨骨舌鱼(海象) <i>A. gigas</i>	尼罗异耳骨舌鱼 (非洲龙鱼) <i>H. niloticus</i>	双须骨舌鱼 (银龙鱼) <i>O. bicirrhosum</i>	费氏骨舌鱼 (黑龙鱼) <i>O. ferreirai</i>	美丽硬仆骨舌鱼[亚洲龙鱼 (金龙鱼、红龙鱼和青龙鱼)] <i>S. formosus</i>	乔氏硬仆骨舌鱼 (珍珠龙鱼) <i>S. jardinii</i>	硬仆骨舌鱼 (星点珍珠龙鱼) <i>S. leichardti</i>
GenBank 登录号	MW829488	MW829489	MW829490	MW829491	MW829492	MW829493	MW829494
GenBank accession no.							
核糖体基因全长/bp total length of rDNA	11714	8957	12057	11556	10377	10724	10725
GC 含量/% GC content	65.10	63.78	64.56	65.03	65.83	66.13	66.00
串联重复序列/个 number of tandem repeats	16	0	12	11	6	5	5
SSR 数量/个 number of SSR	40	42	79	86	68	63	63
18S 序列长度/bp length of 18S	1835	1835	1835	1835	1835	1835	1835
GC 含量/% GC content	54.11	54.06	54.33	54.28	54.39	54.50	54.50
ITS1 序列长度/bp length of ITS1	718	657	871	825	710	599	598
GC 含量/% GC content	73.26	72.60	77.84	77.94	76.48	78.30	78.26
5.8S 序列长度/bp length of 5.8S	159	159	159	159	159	159	159
GC 含量/% GC content	57.23	57.23	57.86	57.86	57.86	57.86	57.86
ITS2 序列长度/bp length of ITS2	327	462	544	518	612	407	407
GC 含量/% GC content	70.34	73.59	77.02	77.61	80.07	79.36	79.36
28S 序列长度/bp length of 28S	3996	4029	4419	4408	4362	4336	4336
GC 含量/% GC content	62.76	62.42	65.69	65.63	65.36	65.38	65.38
IGS 序列长度/bp length of IGS	4679	1815	4229	3811	2699	3388	3390
GC 含量/% GC content	70.06	71.52	63.73	65.31	68.80	70.04	69.62

注: 括号中为俗称。
Note: Common name of the species is in bracket.

始,共运行 5000000 代,每 500 代取一次样,将前 25%样本作为老化样本舍去,余下的构建进化树和统计后验概率(posterior probability, PP),当 5000000 代运行完成,分裂频率标准差(average standard deviation of the split frequency)低于 0.01 时,认为两个独立分析的结果非常接近,建树结果用 TreeView 查看。

2 结果与分析

2.1 基因组测序数据统计

利用三代高通量技术测序,获得 10.58 G 数据,组装发现 3 种体色的美丽硬仆骨舌鱼(金龙鱼、红龙鱼和青龙鱼)的核糖体序列完全一致。在二代测序中,以金龙鱼为美丽硬仆骨舌鱼样本,与其他 6 种骨舌鱼一起进行二代 Illumina 测序,测序结果的 Clean Reads 总数据产出及质量统计见表 2。结合三代测序和二代测序的组装结果,本研究获得了 7 条完整的核糖体全序列(18S-ITS1-5.8S-ITS2-28S-IGS),包括巨骨舌鱼和尼罗异耳骨舌鱼、南美的双须骨舌鱼和费氏骨舌鱼、美丽硬仆骨舌鱼、大洋洲的乔氏硬仆骨舌鱼和硬仆骨舌鱼。

2.2 核糖体基因序列特征

本研究获得了 7 种现存骨舌鱼科的完整核糖体基因转录单元,序列均包括编码区(18S、5.8S 和 28S)和间隔区(ITS1、ITS2 和 IGS)。转录单元序列长度分别为 11714 bp (巨骨舌鱼)、8957 bp (尼罗异耳骨舌鱼)、12057 bp (双须骨舌鱼)、11556 bp (费氏骨舌鱼)、10377 bp (美丽硬仆骨舌鱼)、10724 bp (乔氏硬仆骨舌鱼)和 10725 bp (硬仆骨舌鱼)(表 1)。整个核糖体序列存在 4935 个变异位点,2336 个简约信息位点,GC 碱基含量在 63.78%~66.13% 之间(图 1,表 1)。

2.2.1 18S 7 种骨舌鱼类 18S 序列长度均为 1835 bp,GC 含量变化小,为 54.11%~54.50%。有 41 个简约信息位点,保守位点比例为 97.33%,变异位点比例为 2.67%。巨骨舌鱼和尼罗异耳骨舌鱼的 18S 序列只有 6 个碱基的差异,双须骨舌鱼和费氏骨舌鱼只有 1 个碱基差别(1836nt: G/T)。乔氏硬仆骨舌鱼和硬仆骨舌鱼的 18S 序列一致,同属的美丽硬仆骨舌鱼与两者有 3 个碱基差异(133nt:

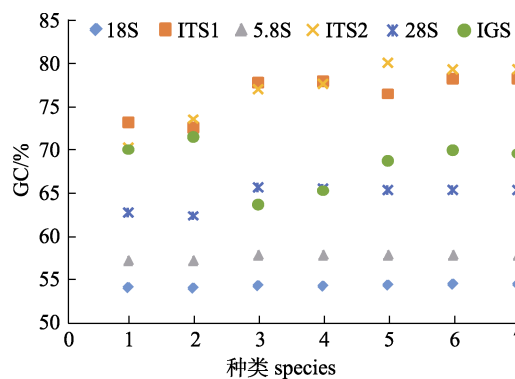


图 1 骨舌鱼核糖体基因中 18S、ITS1、5.8S、ITS2、28S 和 IGS 的 GC 含量

1: 巨骨舌鱼; 2: 尼罗异耳骨舌鱼; 3: 双须骨舌鱼;
4: 费氏骨舌鱼; 5: 美丽硬仆骨舌鱼; 6: 乔氏硬仆骨舌鱼;
7: 硬仆骨舌鱼。

Fig. 1 GC content in 18S, ITS1, 5.8S, ITS2, 28S and IGS sequences in the ribosomal gene of Osteoglossidae
1: *Arapaima gigas*; 2: *Heterotis niloticus*; 3: *Osteoglossum bicirrhosum*; 4: *O. ferreirai*; 5: *Scleropages formosus*; 6: *S. jardinii*; 7: *S. leichardti*.

T/A, 1424nt: C/A, 1425nt: G/A), 表明 18S 是高度保守的。

2.2.2 ITS1 7 种骨舌鱼类 ITS1 序列长度为 598~871 bp, 双须骨舌鱼最长, 硬仆骨舌鱼最短。GC 含量(72.6%~78.3%)远高于 AT 含量(21.7%~27.4%)。简约信息位点 323 个, 保守位点比例为 33.56%, 变异位点比例为 63.21%。序列比对发现, 在 540~692 bp 有 3 段 30 bp 以上的保守序列(图 2)。其中乔氏硬仆骨舌鱼和硬仆骨舌鱼序列差异较小, 序列一致性为 99.51%。

2.2.3 5.8S 7 种骨舌鱼类 5.8S 长度均为 159 bp, 简约信息位点 3 个, 保守位点比例为 98.11%, 变异位点比例 1.89%。3 个变异位点(11nt: G/T; 42nt: A/C; 128nt: C/A)处在亚科之间, 即巨骨舌鱼和尼罗异耳骨舌鱼序列一致, 其他 5 种骨舌鱼间序列一致。GC 含量变化很小, 在 57.2%~57.9% 之间, 是一个高度保守的序列。

2.2.4 ITS2 7 种骨舌鱼类 ITS2 的序列长度为 327~612 bp, 差异比较大, 美丽硬仆骨舌鱼最长, 巨骨舌鱼最短。GC 含量(70.34%~80.07%)远高于 AT 含量(19.93%~29.66%)。简约信息位点 161 个, 保守位点比例为 30.87%, 变异位点比例 54.52%。序列比对发现, 在序列的 5'端和中部有 2 段保守

表 2 7 种骨舌鱼总 DNA 二代测序数据统计表
Tab. 2 Total DNA sequencing data of Osteoglossidae

物种 speices	读长 read	读数 read number	碱基数 base number	总碱基数 total base number	GC 含量/% GC content	质量值 20 /% Q20	质量值 30 /% Q30
巨骨舌鱼 <i>Arapaima gigas</i>	Read1	4465257	669788550	1339577100	42.55	99.62	98.73
	Read2	4465257	669788550		42.64	96.86	90.47
尼罗异耳骨舌鱼 <i>Heterotis niloticus</i>	Read1	4897612	734641800	1469283600	43.20	99.50	98.31
	Read2	4897612	734641800		43.26	97.81	93.15
双须骨舌鱼 <i>Osteoglossum bicirrhosum</i>	Read1	4690420	703563000	1407126000	45.19	99.38	98.14
	Read2	4690420	703563000		45.31	97.06	91.21
费氏骨舌鱼 <i>O. ferreirai</i>	Read1	3846166	576924900	1153849800	46.32	99.35	98.06
	Read2	3846166	576924900		46.37	97.24	91.70
美丽硬仆骨舌鱼 <i>Scleropages formosus</i>	Read1	197463943	29619591450	59239182900	45.15	97.00	92.64
	Read2	197463943	29619591450		45.28	92.71	85.17
乔氏硬仆骨舌鱼 <i>S. jardinii</i>	Read1	6244737	936710550	1873421100	45.13	99.48	98.21
	Read2	6244737	936710550		45.21	97.82	93.17
硬仆骨舌鱼 <i>S. leichardti</i>	Read1	5078369	761755350	1523510700	42.91	99.49	98.27
	Read2	5078369	761755350		43.00	97.79	93.13

序列, 结果如图 3 所示。其中乔氏硬仆骨舌鱼和硬仆骨舌鱼序列差异较小, 序列一致性为 99.53%。

2.2.5 28S 7 种骨舌鱼类 28S 的全序列长度为 3996~4419 bp, GC 含量为 62.4%~65.9%。简约信息位点 249 个, 保守位点比例为 91.96%, 变异位点比例 7.77%。亚科内的序列差异不大, 一致性在 96%以上。异耳鱼亚科仅有 54 个变异位点, 骨舌鱼亚科仅有 103 个变异位点。2 个亚科之间序列差异较大的主要原因是异耳鱼亚科有 3 段 50 bp 以上的碱基序列缺失。

2.2.6 IGS 7 种骨舌鱼类 IGS 序列为 1815~4679 bp, 长度变异大, 巨骨舌鱼最长, 尼罗异耳骨舌鱼最短。GC 含量为 63.73%~71.52%, 简约信息位点 1808 个, 保守位点比例为 14.56%, 变异位点比例 79.68%。乔氏硬仆骨舌鱼和硬仆骨舌鱼序列差异较小, 序列一致性为 98.31%。

综合以上数据可以看出, 骨舌鱼科的编码基因 18S、5.8S 和 28S 的序列和 GC 含量相对较保守, 变化范围较小, 而间隔区 ITS1、ITS2 和 IGS 的变异较大, 符合真核生物核糖体基因的序列特征。

2.3 重复序列

众多研究表明, 核糖体中存在大量重复序列。骨舌鱼科鱼类也存在这种情况, 导致不同骨

舌鱼类基因的长度不同。本研究分析了序列长度超过 25 bp, 且重复次数超过 2 次的串联重复序列, 结果显示全部集中在 IGS 区域(表 3)。除了尼罗异耳骨舌鱼外, 其他 6 种骨舌鱼均发现大片段串联重复序列。其中巨骨舌鱼发现 10 个串联重复序列, 6 个是 100 bp 以上的串联重复序列; 双须骨舌鱼发现 7 个串联重复序列, 最大串联重复序列为 160 bp; 费氏骨舌鱼发现 6 个串联重复序列, 最大串联重复序列为 455 bp; 美丽硬仆骨舌鱼、乔氏硬仆骨舌鱼和硬仆骨舌鱼分别发现 2、3 和 3 个串联重复序列。

另外, 在骨舌鱼核糖体中也发现大量微卫星序列, 除了 5.8S 外, 其他区段均存在微卫星序列(表 4), 主要分布在 IGS 和 28S, 费氏骨舌鱼的微卫星数量达到 86 个。

2.4 遗传距离和聚类分析

利用 MEGA X 和 DNAMEN 分别对骨舌鱼类各个种间 18S、ITS1、5.8S、ITS2、28S 和 IGS 对进行非校正 P 距离的计算和序列的一致性分析(表 5), 结果显示, 遗传距离分别为 0~0.024、0~0.852、0~0.019、0~0.799、0~0.066 和 0.001~1.877, 序列的一致性分别为 97.52%~100%、38.34%~99.5%、98.11%~100%、31.63%~99.53%、84.98%~

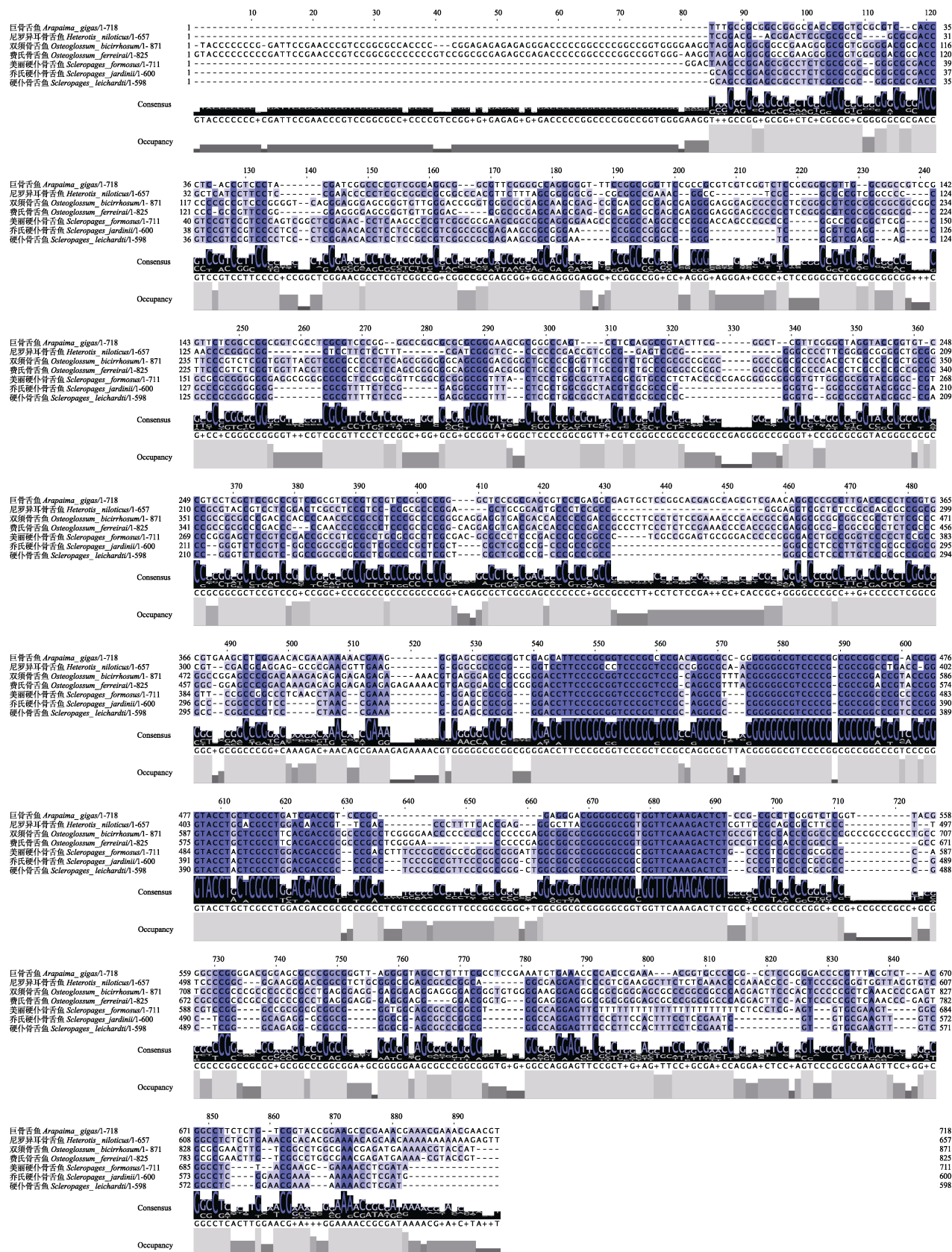


图 2 骨舌鱼类 ITS1 序列比对图

Fig. 2 Multiple alignment of ITS1 sequence of Osteoglossidae

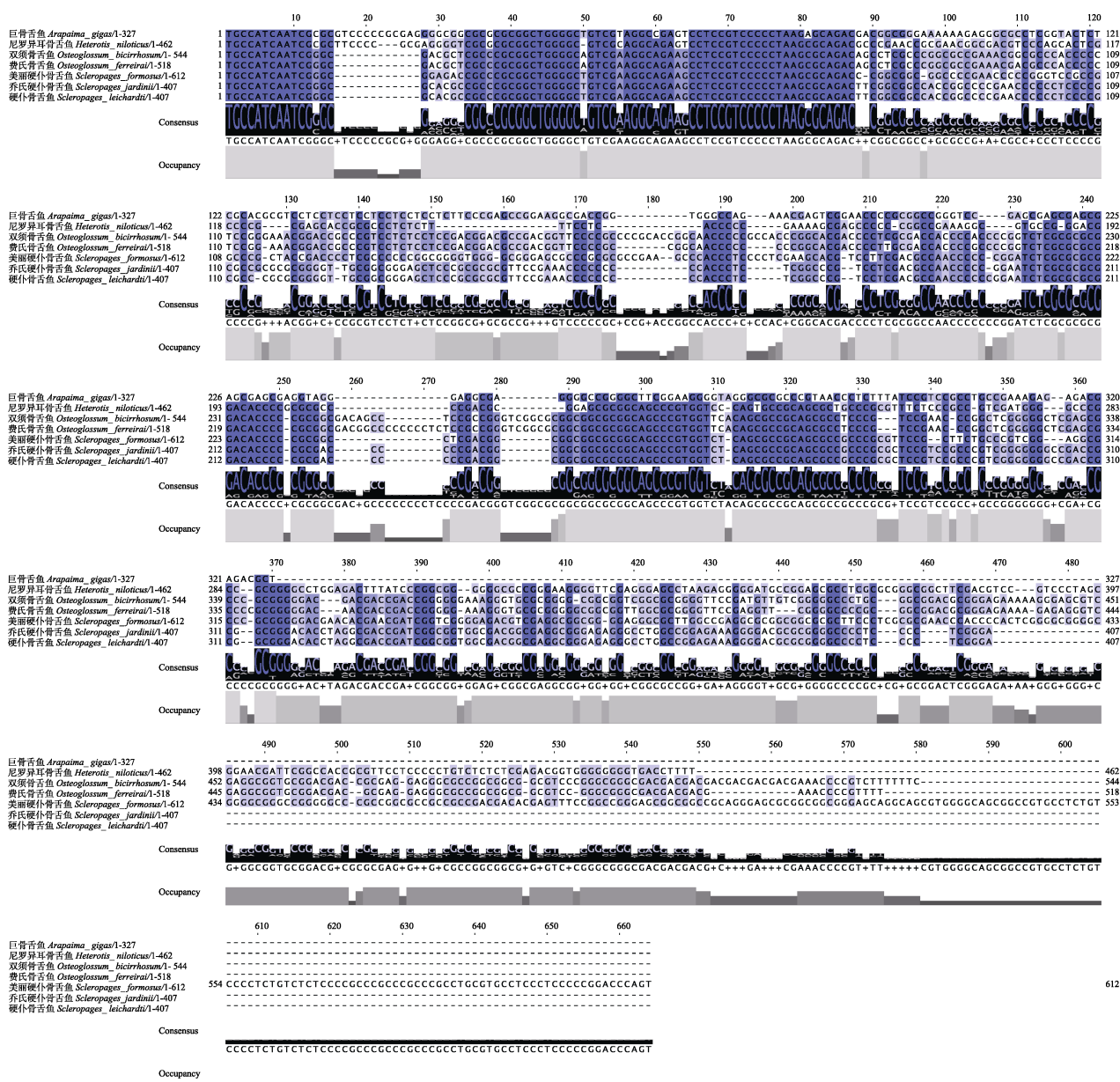


图 3 骨舌鱼类 ITS2 序列比对图

Fig. 3 Multiple alignment of ITS2 sequence of Osteoglossidae

表 3 骨舌鱼科 IGS 区域的串联重复序列

Tab. 3 Tandem repeats in IGS region of Osteoglossidae

物种 speices	重复次数 copy number	长度/bp consensus size	重复序列(5'-3') tandem repeat sequence (5'-3')
巨骨舌鱼 <i>Arapaima gigas</i>	4.2	26	GGCCGAGAGGGCCGGGGTACCAGGAC
			CGGGGTACCAGGCGGCCGAGAGGGCCGGGGTACCAGGACGGCGGGCCTTAA
	5.1	120	GGGAGCAGGTAGGGCCGGAGGAAGGAAGGGCGAGGGGGTCCGGGACTTAA
			GACCGAAGGCCGAGAGGGC
美丽硬仆骨舌鱼 <i>Scleropages formosus</i>			CGGGGTACCAGGCGGCCGAGAGGGCCGGGGTACCAGGACGGCGGGCCTTAA
			GGGAGCAGGTAGGGCCGGAGGAAGGACGGGCGAGGAGGTCCGGGACTTAA
	2.5	238	ACCGAAGGCCGAGGGCCGGGGTACCAGGCGGCCGAGAGGGCCGGGGTACGA
			GGAAAGAGGGCCTCAGAGGCGGGGTAGGGCCCGACGAAGGAAGGGCGAG
			GGGGTACGGGACTTAAGACCGAAGGCCGAGAGGGC

(待续 to be continued)

(续表 3 Tab. 3 continued)

物种 speices	重复次数 copy number	长度/bp consensus size	重复序列(5'-3') tandem repeat sequence (5'-3')
巨骨舌鱼 <i>Arapaima gigas</i>	2.1	146	GGCCAGGGGTACCAGACGGCCGAGAGGGCCGGGGTACGAGGACGGCGGGGCC TTAAGGGAGCAGGTAAGGCCGGAGGAAGGAAGGGCGAGGAGGTCCGAGACT TAAGACCGAAGGCCGAGGGCCGGGGTACCAGACGTCCAGTAA
	14.3	29	GGGGTACCAGACGTCCGCTTTAAGGCCCA
	16.8	58	CTCGGGGGGCGGGGGCCGACTCGGGGAGACGCGAGGCCGAGGCCAGCGCA CCGACAG
	5.3	176	CTCGGGGGGCGGGGGCGACCCAGGGAGGACGCGGGCCGAAGCCCAGCGCA CCGACAGCTCGGGGAAGCCGGGGTCTGACTCGTGGAGGTCTCGGGGGCGAGGC CCAGCGCACCGACAGCTCGGGGGGGCCGGGGCCGACCAGGGGAGACGGGAG GCCGAAGCCCAGCGCACCGACAG
	2.5	354	CCACCACCTCGGAGGGGCGGGGGCCGACTCGGGGAGGTACGGGCCGAGGC CCAGCGCACCGACAGCTCGGGGGGGCCGGGGCCGACTCGGGGAGGTCTCGCG GCCGAGGCCCAACGCTCTTACAACCTCGGGGGAGCCCGAGGGACCCGGGGA GACGGGAGGCCGAGGCCAGCGCACCGACAGCTCGGGGGAGCCCGGAGGGA CCCAGGGAGACGGGAGGCAGAAGCCCAGCGCACCGACAGCTCGGGGAAGCC GGGGTGCACCCGTGGAGACACGAGGCCGAGACCCAGCGCACCGACAGCTCG GGGGGGCCGAGCCGACCAGGGGAGACGGGAGGCCGAAGCCCAGCGCA
	8.4	119	CTCGGGGGGCGGGGGCCGACTCGGGGAGGTCTCGGGGCCGAGGCCAGCGC ACCGACAGCTCGGGGGGGCCGGGGCCGACTCAGGGGAGACGGGAGGCCGAA GCCAGCGCACCGACAG
	10.4	51	GAGAGCCTCTGCCGGGGCCCTGACCCTGCGCCGATCGTGGGTCCCCCTAC
	15.6	32	CACCCTTAAGCCTCCGGCGAGCGGAGGCCCGC
	3.1	160	CACCCTTAAGCCTCCGGCGAGCGGAGGCCCGCCACCCTTAAGCCTCCGGCGACC GGAGGCCCGCCACCCTTAAGCCCCGTCGAGCGCAGGCCACCACCCTTAAGCC TCCTGCGAGCGGAGGGCCGCCACCCTTAAGCCTCTGCGAGCGGAGGGCCGC CGCTCAGAGACTTTGCGCCCCAAGCAAGTGCACATTTCTGGGCCGTGCGCTCA GAGACCTGGACCCCCCCCCACGTCTG
	4.9	76	TGCACATTTCTGGGCTTCGCTCGGAGACCTGAGCCCCGAGCGCC
双须骨舌鱼 <i>Osteoglossum bicirrhosum</i>	10.6	44	TGCACATTTCTGGGCTTCGCTCGGAGACCTGAGCCCCGAGCGCC
	5.3	92	ACTTTAGGTCTGCGCTGGGAGACCTGAGCCCCGAGCGCC
	2.2	92	CACACTTCAGGCCCTGCGCTGAGAGACACGAGCCCCAAGCGCCTGCACATTT CGGGCCCTGCGCGCAGACGCTTAGCCCCCTTCGCGCGCC
	2.3	71	CACATTTCTGGGCCCTGCGCGCAGAGACTTTAGCCCCCTTCGCGCGCAAGCAC GCTTCAGGCCCTGCGTGCT
	7.6	32	CACCCTTAAGCCTCCGGCGAGCGGAGGCCCGC
费氏骨舌鱼 <i>Osteoglossum ferreirai</i>	3.8	64	CACCCTTAAGCCTCCGGCGAGCGGAGGCCCGCCACCCTTAAGCCTCCGGCGA GCGCAGGCCCGC
	2.5	95	CACCCTTAAGCCTCCGGCGAGCGGAGGCCCGCCACCCTTAAGCCCCGGCGAC CGGAGGGCCGCCACCCTTAAGACTCCGGCGAGCGCAGGCCCGC
	14.3	76	CGGAGACCTGTACCCCCCTGTGCGCGCACAGACACTTTGCGCCCCAAGCA GCTGCACATTTCTGGGCCGTGAGCT
	4.7	228	CGGAGACTTGGACCCCCCGTCTGCGCGCGCAGACACTTTGCGCCCCAAGCAA GCTTACGTTTCGGGCCGTGAGCTCGGAGACCTGTACCCCCCTGTGCGCGCGC ACAGACACTTTGCGCCCCAAGCAACTGCACATTTCTGGGCCCTTGTGCTCGGAG ACCTGTACCCCCCGTCTGCGCGCACAGACACTTTGCGCCCCAAGCAAGCTG CACATTTCTGGGCCGTGAGCT
	2.3	455	ACCCCCCGTCTGCGCGCAGAGACACTTTGCGCCCCAAGCAACTCACGTTTC GGGCCGTGAGCTCGGAGACCTGTACCCCCCTGTGCGCGCGCACAGACACTTT GCGCCCCAAGCAACTGCACATTTCTGGGCCCTTGTGCTCGGAGACCTGTACCCC CCCCGTCTGCGCGCACAGACACTTTGCGCCCCAAGCAAGCTCACATTTCTGGGC CGTGAAGCTTACGTTTCGGGCCGTGAGCTCGGAGACCTGTACCCCCCTGT TCGCGCGCACAGACACTTTGCGCCCCAAGCAAGCGCACATTTCTGGGCCCTTGC GCTCAGAGACTGTGACCCCCCGTCTGCGCGCACAGACACTTTGCGCCCCA AGCAAACGCACATTTCTGGGCCGTGAGCTCGGAGACTTAC
	2.2	49	TCAAAGTCCCCCTGGGGGACCAGGCGCCCGGAGGCCGCTCGGAAACGC
	2.5	49	AAGTCCCTTCGGGTACCAGGCGTCCGCGGAGGAAAGCCAGAAATTTTCT
乔氏硬仆骨舌鱼/ 硬仆骨舌鱼 <i>Scleropages jardinii/leichardti</i>	2	153	CGGCGGCGGGGACTTAAGCCCCGGCGGGCCGAGGGGAGCGGGGGGGCCAAAG CCCCCTGCCCCGACACTCAAGCTTCAGACTTAAGCCCCGGGGGACGCGCA GCTAGAGGCGGCGACGCCGCGGCCGAGGGGCGACGGCCCTGGTACCC
	2.8	50	CCGAAAAATTTTCAAAGTCCCGCGGGGGACCAGGCGTCCCGAGCCACC
	2.9	48	CCCCCGGAAATATTTTCAAAGTCCCGCGGGGGACCAGGCGTCCCGAT

表 4 骨舌鱼科鱼类核糖体基因中微卫星序列数量分布
Tab. 4 Number of microsatellite sequences in rDNA of Osteoglossidae

区域 region	巨骨舌鱼 <i>Arapaima gigas</i>	尼罗异耳骨舌鱼 <i>Heterotis niloticus</i>	双须骨舌鱼 <i>Osteoglossum bicirrhosum</i>	费氏骨舌鱼 <i>Osteoglossum ferreirai</i>	美丽硬仆骨舌鱼 <i>Scleropages formosus</i>	乔氏硬仆骨舌鱼 <i>Scleropages jardinii</i>	硬仆骨舌鱼 <i>Scleropages leichardti</i>
18S	7	7	6	6	6	6	6
ITS1	3	5	8	7	5	5	5
ITS2	6	5	4	4	10	5	5
28S	15	14	24	23	23	22	22
IGS	9	11	37	46	24	25	25
合计 totle	40	42	79	86	68	63	63

表 5 骨舌鱼科在 18S/ITS1/5.8S/ITS2/28S/IGS 核糖体序列的遗传距离及序列一致性
Tab. 5 Genetic distances and sequence identity of 18S/ITS1/5.8S/ITS2/28S/IGS rDNA in Osteoglossidae

物种 species	巨骨舌鱼 <i>Arapaima gigas</i>	尼罗异耳骨舌鱼 <i>Heterotis niloticus</i>	双须骨舌鱼 <i>Osteoglossum bicirrhosum</i>	费氏骨舌鱼 <i>O. ferreirai</i>	美丽硬仆骨舌鱼 <i>Scleropages formosus</i>	乔氏硬仆骨舌鱼 <i>S. jardinii</i>	硬仆骨舌鱼 <i>S. leichardti</i>
巨骨舌鱼 <i>Arapaima gigas</i>	—	99.68/53.25/100/ 43.06/96.3/19.31	97.79/43.87/ 98.11/35.02/ 84.98/37.63	97.73/42.92/ 98.11/35.1/ 85.13/33.53	97.68/50.73/98.11 /31.63/85.65/26.68	97.73/46.03/98.11 /43.05/86.29/32.4	97.73/40.35/ 98.11/42.82/ 86.32/32.24
尼罗异耳骨舌鱼 <i>Heterotis niloticus</i>	0.003/0.679/0/ 0.665/0.014/ 1.007	—	97.57/39.69/98.11 /45.66/86.07/20.8	97.52/40.72/ 98.11/45.7/ 86.13/20.77	97.57/45.09/98.11 /46.41/86.88/31.91	97.63/53.79/98.11 /50.59/87.6/24.67	97.63/51.91/ 98.11/51.37/ 87.6/24.5
双须骨舌鱼 <i>Osteoglossum bicirrhosum</i>	0.022/0.678/ 0.019/0.784/ 0.066/1.877	0.024/0.822/0.019/ 0.571/0.065/1.164	—	99.95/91.3/100/ 89.84/99.73/ 75.34	99.68/51.01/100/ 53.06/96.42/31.67	99.73/38.34/100/ 52.01/96.38/39.25	99.73/38.66/ 100/52.02/ 96.36/40.75
费氏骨舌鱼 <i>O. ferreirai</i>	0.022/0.654/ 0.019/0.799/ 0.066/1.865	0.024/0.815/0.019/ 0.559/0.064/1.084	0.001/0.017/0/ 0.024/0/0.128	—	99.62/53.14/100/ 51.82/96.44/35.58	99.68/43.85/100/ 51.12/96.44/46.1	99.68/43.33/ 100/51.31/ 96.42/49.12
美丽硬仆骨舌鱼 <i>Scleropages formosus</i>	0.023/0.722/ 0.019/0.707/ 0.062/1.373	0.024/0.611/0.019/ 0.556/0.062/0.843	0.003/0.852/0/ 0.439/0.019/ 0.769	0.004/0.832/0/ 0.413/0.019/ 0.756	—	99.84/69.13/100/ 53.21/97.73/42.64	99.84/69.45/ 100/53.59/ 97.7/42.43
乔氏硬仆骨舌鱼 <i>S. jardinii</i>	0.022/0.609/ 0.019/0.706/ 0.063/1.483	0.023/0.478/0.019/ 0.490/0.063/0.872	0.003/0.779/0/ 0.406/0.016/ 0.857	0.003/0.740/0/ 0.394/0.016/ 0.854	0.002/0.198/0/ 0.369/0.010/0.483	—	100/99.51/ 100/99.53/ 99.95/98.31
硬仆骨舌鱼 <i>S. leichardti</i>	0.022/0.608 /0.019/0.702/ 0.063/1.488	0.023/0.478/0.019/ 0.492/0.063/0.872	0.003/0.779/0/ 0.406/0.016/0.862	0.003/0.740//0/ 0.394/0.015/ 0.857	0.002/0.196/0/ 0.370/0.009/0.484	0/0/0/0/0.001	—

注: 左下角为遗传距离, 右上角为一致性。
Note: Uncorrected pairwise distances are below diagonal and sequence identities are above diagonal.

99.95%和 19.31%~98.31%。表明越保守的序列, 种间的遗传距离越小, 序列的一致性越高。

利用 rDNA 不同分区序列(18S、ITS1、ITS2 和 28S)和联合序列(18S+ITS1+5.8S+ITS2+28S), 以新西兰绿唇贻贝(*Perna canaliculus*)(GenBank 登录号为 MK419109.1)为外群, 使用 MEGA X、Bayes 软件分别构建了 NJ 和 BI 系统发育树。因 5.8S 序列较短和 IGS 序列长度差异过大, 未进行系统发育分析。

聚类结果如图 4 所示, 两种树的拓扑结构是一致的, 所有骨舌鱼类聚为一枝(红色框中)。除 ITS1 外, 其他 4 个序列构建的进化树均与目前 7 种骨舌鱼的传统分类相符合。表明骨舌鱼类核糖体 18S、ITS2 和 28S 单区序列以及 18S+ITS1+5.8S+ITS2+28S 多序列串联建树都可以较好地解决骨舌鱼科亚科、属和种的分类关系。说明核糖体 DNA 可作为骨舌鱼科以下阶元的分子标记。

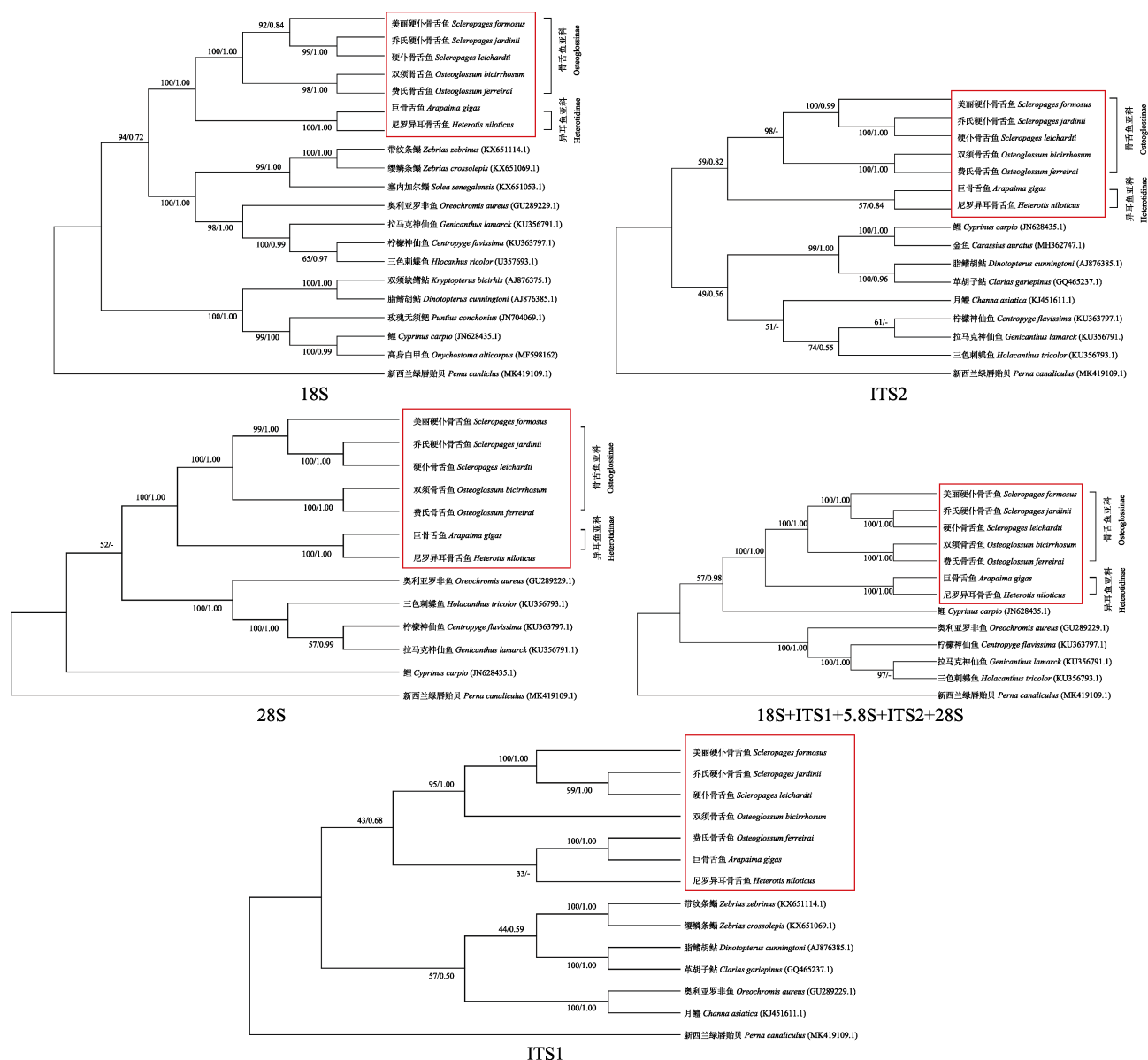


图 4 分别基于 5 个 rDNA 序列构建的 NJ/BI 分子系统进化树

Fig. 4 NJ/BI phylogenetic trees based on 5 rDNA sequences

3 讨论

3.1 编码区序列的遗传进化分析

本研究中 7 种骨舌鱼的 18S rDNA 长度均为 1835 bp, 序列一致性为 97.52%~100%, 保守位点比例为 97.33%, 可见 18S 序列在骨舌鱼科的遗传进化中是高度保守的。GC 含量(54.11%~54.5%)高于 AT 含量(45.5%~45.94%), 与鲇形目(GC 含量 55.77%~56.33%)^[37]、鲈形目(GC 含量 52.6%~57.1%)^[38]、鲤(Cyprinus carpio)(GenBank 登录号

JN628435.1, GC 含量 56.21%)、高身白甲鱼(Onychostoma alticorpus)(MF598162.1, GC 含量 56.62%)、带纹条鲷(Zebrias zebrinus)(KX651114.1, GC 含量 54.29%)、缨鳞条鲷(Z. crossolepis)(KX651069.1, GC 含量 54.27%)和塞内加尔鲷(Solea senegalensis)(KX651053.1, GC 含量 53.82%)的结果是一致的。本研究中 7 种骨舌鱼的 5.8S 长度较短, 为 159 bp, GC 含量为 57.23%~57.86%, 与鲇形目(5.8S 长度 157 bp, GC 含量 57.96%~59.24%)^[37]、鲈形目(5.8S 长度 160 bp, GC

含量 55.6%~58.9%)^[38]的研究结果是一致的。本研究中 7 种骨舌鱼 28S 长度为 3996~4419 bp, 保守位点比例为 91.86%, GC 含量为 62.42%~65.69%, 高于 60%, 与鲤(*C. carpio*) (JN628435.1, GC 含量 63.50%)、鲈形目柠檬神仙鱼(*Centropyge flavisima*) (KU363797.1, GC 含量 62%)、奥利亚罗非鱼(*Oreochromis aureus*) (GU289229.1, GC 含量 61.08%) 是一致的。骨舌鱼同一亚科中, 28S 序列一致性均在 96%以上, 可见, 28S 在骨舌鱼的亚科进化水平是高度保守的。

GC 含量作为核苷酸序列的一项重要指标, 与重复序列的分布、甲基化、基因密度等许多基因组特征指标有密切的关系^[39], GC 含量越高, 说明 DNA 分子结构越稳定^[40], 甚至还与进化地位有一定关系^[41]。本研究中骨舌鱼类的 3 个编码区 GC 含量与其他鱼类相比均处于一个相对稳定的水平, 为编码区序列的保守性提供了基础。骨舌鱼科的 3 个编码区相比较, 序列长度 28S>18S>5.8S, 保守性 5.8S>18S>28S, GC 含量 28S>5.8S>18S。分子系统发育进化分析表明, 基于 18S 和 28S 构建的系统进化树均与传统分类学一致, 可作为骨舌鱼科以下分类阶元的分子标记。

骨舌鱼科核糖体基因的 3 个编码区均高度保守, 其中 5.8S 序列最保守, 在异耳鱼亚科内和骨舌鱼亚科内完全一致。对于 18S, 骨舌鱼属的双须骨舌鱼和费氏骨舌鱼序列一致性和遗传距离分别为 99.95%和 0.001; 硬骨舌鱼属的乔氏硬仆骨舌鱼和硬仆骨舌鱼序列一致性和遗传距离分别为 100%和 0, 与同属的美丽硬仆骨舌鱼一致性和遗传距离分别为 99.84%和 0.002。对于 28S, 双须骨舌鱼和费氏骨舌鱼序列一致性和遗传距离分别为 99.73%和 0.002; 乔氏硬仆骨舌鱼和硬仆骨舌鱼序列一致性和遗传距离分别为 99.95%和 0, 与同属的美丽硬仆骨舌鱼序列一致性和遗传距离分别为 97.7%和 0.01。可见在骨舌鱼亚科中, 18S 和 28S 序列的保守性在进化中与地理分布可能有一定关系。双须骨舌鱼和费氏骨舌鱼分布于南美洲, 地理位置比较接近, 保守性高遗传变异小; 乔氏硬仆骨舌鱼和硬仆骨舌鱼分布于大洋洲, 地理位置比较近, 保守性高遗传变异小; 而美丽硬仆骨

舌鱼分布于亚洲, 与同属的大洋洲龙鱼产生了一定的遗传分化, 这与笔者前期的线粒体研究结果是一致的^[14]。作为一种古老的鱼类, 推测长期的地理隔离可能是骨舌鱼遗传分化的原因之一。

3.2 间隔区的遗传进化分析

本研究所测骨舌鱼科的 ITS1 序列长度为 598~871 bp, 符合真核生物 ITS1 小于 1100 bp 的范围^[42], 也在硬骨鱼类 ITS1 的长度区间(272~918 bp)^[38]。骨舌鱼 ITS1 的 GC 含量比较高(72.6%~78.3%), 高于一般的硬骨鱼类 GC 含量, 如鲈形目鱼类(51.05%~71.04%)^[37]、石斑鱼属(*Epinephelus*) (58.5%~67.9%)、射水鱼(*Toxotes chatareus*)(71.4%)^[43]、鲷亚目鱼类(62.6%)^[44]、鲷科鱼类(67.0%~72.1%)^[45-46]和金鱼(*Cyprinus auratus*)(66.9%)^[47]等。骨舌鱼的 ITS2 序列长度为 327~612 bp, 在硬骨鱼类 ITS2 范围(128~694 bp)中^[38]。GC 含量较高(70.34%~80.07%), 与其他鱼类研究结果类似, 如鲈形目鱼(52.71%~74.73%)^[48]、鲈形目鱼(62.4%~77.5%)^[44, 28]和鲷科鱼类(63.95%~70.16%)^[49]等。研究发现, 物种变异主要发生在间隔区, 且 ITS1 比 ITS2 更具变异性^[50], 本研究结果也是如此(ITS1 变异率 63.21%>ITS2 变异率 54.52%)。ITS 区域的变异通常与重复序列有关, 包括简单的重复序列和相对较长的重复单元基序^[42, 51-52]。骨舌鱼类的 ITS 区均存在微卫星序列, 却很少发现 10 bp 以上的较长重复序列, 仅在双须骨舌鱼和费氏骨舌鱼的 ITS1 各有 1 段 2*20 bp (CCCGCCCGCCCGCCCGCCTG) 和 2*31 bp (GCCCCGCGCGCGCCCGCGCGCCCCACCCCTC) 的重复序列, 在美丽硬仆骨舌鱼 ITS2 发现 1 段 2*21 bp (GCGGCGGCAGCAGGCAGCGCG) 的重复序列, 在其他骨舌鱼中并未发现。

骨舌鱼科的 IGS 序列长度变异大, 从 1815 bp 到 4679 bp, GC 含量为 63.73%~71.52%。在以往的研究中, IGS 序列中含有大量的重复序列, IGS 序列的长度与这些重复序列的长度和数量密切相关^[53]。骨舌鱼的 IGS 序列中, 除尼罗异耳骨舌鱼外, 其他 6 种鱼有 2~15 个 10 bp 以上的串联重复单元, 这可能是导致骨舌鱼 IGS 序列长度变化较大的原因。目前鱼类的 IGS 全序列还未见相关报道, 在其他物种中, IGS 的长度在几千到几万不等, 如人

29685 bp (GenBank 登录号为 U13369.1)、新西兰绿唇贻贝(*P. canaliculus*) 1400 bp (GenBank 登录号为 MK419109.1)、矮小拟镖剑水蚤(*Paracyclops nana*) 1936 bp (GenBank 登录号为 FJ214952.1)、中华绒螯蟹(*Eriocheir sinensis*) 4240 bp^[54]、皱纹盘鲍(*Haliotis discus hannai*) 4642~4654 bp^[55]、黑木耳(*Auricularia auricula-judae*) 5757 bp^[56]、紫菜(*Pyropia yezoensis*) 5984 bp^[57]及吸虫 1822~2419 bp^[58]。骨舌鱼 ITS1、ITS2 和 IGS 的变异位点比例分别为 63.21%、54.52%和 79.68%，可见非编码区的进化速率快，需要较高的 GC 含量来保证其 DNA 分子的结构稳定性。另外 Bernardi 等^[59]认为 GC 含量越高越容易适应热带气候，推测骨舌鱼核糖体间隔区高 GC 含量也与骨舌鱼类生活在热带气候环境有关。

骨舌鱼科 ITS1、ITS2 和 IGS 序列一致性分别为 38.34%~99.5%、31.63%~99.53%和 19.31%~98.31%，在不同属之间的一致性低于 55%，可见，间隔区在不用属间的遗传变异大。分布于南美洲骨舌鱼属的双须骨舌鱼和费氏骨舌鱼，ITS1、ITS2 和 IGS 序列遗传距离分别为 0.017、0.024 和 0.128。分布于大洋洲硬骨舌鱼属的乔氏硬仆骨舌鱼和硬仆骨舌鱼 ITS1、ITS2 和 IGS 遗传距离分别为 0、0 和 0.001，与分布在东南亚同属的美丽硬仆骨舌鱼遗传距离分别为 0.2、0.37 和 0.48。可见骨舌鱼核糖体基因的间隔区与编码区一样，在进化中也出现了与地理分布相关的遗传变异现象，即地理分布近，遗传变异小。分子系统发育分析表明，由于 ITS1 变异太大，构建的进化树与传统分类不相符，不能作为骨舌鱼科的分子标记。ITS2 和 18S+ITS1+5.8S+ITS2+28S 联合序列构建的系统进化树均与传统分类学一致，可作为骨舌鱼科的分子标记。

4 结论

本研究通过二代测序和三代测序相结合的技术方法，组装获得了现存 7 种骨舌鱼完整的 rDNA 全序列，包含 18S、ITS1、5.8S、ITS2、28S 和 IGS 区。分析发现非编码区的遗传变异速率比编码区快，符合真核生物核糖体基因的序列特征。通过

NJ 和 BI 方法进行系统发育分析，结果表明 18S、ITS2、28S 以及 18S+ITS1+5.8S+ITS2+28S 联合序列构建的进化树均与传统的分类相符合，为骨舌鱼类的系统进化研究提供了数据参考。

参考文献:

- [1] Weider L J, Elser J J, Crease T J, et al. The functional significance of ribosomal (r)DNA variation: Impacts on the evolutionary ecology of organisms[J]. Annual Review of Ecology, Evolution, and Systematics, 2005, 36: 219-242.
- [2] Hillis D M, Dixon M T. Ribosomal DNA: Molecular evolution and phylogenetic inference[J]. The Quarterly Review of Biology, 1991, 66(4): 411-453.
- [3] Appels R, Honeycutt R L. rDNA: Evolution over a billion years[M]//DNA Systematics. Boca Raton: CRC Press, 1986: 81-135.
- [4] Krawczyk K, Nobis M, Nowak A, et al. Phylogenetic implications of nuclear rRNA IGS variation in *Stipa* L. (Poaceae)[J]. Scientific Reports, 2017, 7: 11506.
- [5] Shapoval N A, Lukhtanov V A. Intragenomic variations of multicopy ITS2 marker in *Agrodiaetus* blue butterflies (Lepidoptera, Lycaenidae)[J]. Comparative Cytogenetics, 2015, 9(4): 483-497.
- [6] Bélanger-Lépine F, Leung C, Glémet H, et al. Balancing selection on the number of repeats in the ribosomal intergenic spacer present in naturally occurring yellow perch (*Perca flavescens*) populations[J]. Genome, 2018, 61(1): 1-6.
- [7] Normand A C, Packeu A, Cassagne C, et al. Nucleotide sequence database comparison for routine dermatophyte identification by internal transcribed spacer 2 genetic region DNA barcoding[J]. Journal of Clinical Microbiology, 2018, 56(5): e00046-18.
- [8] Wu B S, Si L Z, Kong X Y, et al. The analysis of ITS1 characteristics of 11 species from 5 families and its application in phylogenetic research[J]. Journal of Fisheries of China, 2018, 42(4): 465-475. [武宝生, 司李真, 孔晓瑜, 等. 5 科 11 种鱼类 ITS1 特征分析及其在系统分类研究中的适用性[J]. 水产学报, 2018, 42(4): 465-475.]
- [9] West C, James S A, Davey R P, et al. Ribosomal DNA sequence heterogeneity reflects intraspecies phylogenies and predicts genome structure in two contrasting yeast species[J]. Systematic Biology, 2014, 63(4): 543-554.
- [10] Si L Z, Wu B S, Kong X Y, et al. Analysis of the GC content of ribosomal genes of 11 species of Perci-formes and comparison with other teleostean fishes[J]. Journal of Fishery Sciences of China, 2017, 24(4): 657-668. [司李真, 武宝生, 孔晓瑜, 等. 11 种鲈形目鱼类的核糖体基因 GC 含量及其

- 与硬骨鱼类的特征比较[J]. 中国水产科学, 2017, 24(4): 657-668.]
- [11] Greenwood P, Rosen D, Weitzman S, et al. Phyletic studies of teleostean fishes, with a provisional classification of living forms[J]. Bulletin of the American Museum of Natural History, 1966, 131(4): 341-455.
- [12] Mu X D, Wang P X, Hu Y C, et al. Morphological structure and growth characteristics of juvenile Asian arowana (*Scleropages formosus*)[J]. Journal of Fisheries of China, 2010, 34(9): 1379-1386. [牟希东, 王培欣, 胡隐昌, 等. 亚洲龙鱼形态结构及幼鱼生长特性[J]. 水产学报, 2010, 34(9): 1379-1386.]
- [13] Moreau M A, Coomes O T. Potential threat of the international aquarium fish trade to silver arowana *Osteoglossum bicirrhosum* in the Peruvian Amazon[J]. Oryx, 2006, 40(2): 152-160.[LinkOut]
- [14] Mu X D, Wang X J, Song H M, et al. Mitochondrial DNA as effective molecular markers for the genetic variation and phylogeny of the family Osteoglossidae[J]. Gene, 2012, 511(2): 320-325.
- [15] Rowley J J L, Emmett D A, Voen S. Harvest, trade and conservation of the Asian arowana *Scleropages formosus* in Cambodia[J]. Aquatic Conservation: Marine and Freshwater Ecosystems, 2008, 18(7): 1255-1262.
- [16] Mu X D, Hu Y C, Wang X J, et al. Genetic variability in cultured stocks of *Scleropages formosus* in mainland China revealed by microsatellite markers[J]. Journal of Animal and Veterinary Advances, 2011, 10(5): 555-561.
- [17] Mu X D, Liu Y, Wang X J, et al. Characterization of the mitochondrial genome and phylogeny of the black arowana (*Osteoglossum ferreirai*)[J]. Biologia, 2014, 69(9): 1222-1230.
- [18] Mu X D, Yang Y X, Liu Y, et al. Complete mitochondrial genome of northern spotted barramundi, *Scleropages jardinii*[J]. Mitochondrial DNA, 2015, 26(5): 698-699.
- [19] Gao B J, Zhang J. The complete mitochondrial genome of *Scleropages formosus* (Osteoglossiformes: Osteoglossidae) and phylogenetic studies of Osteoglossiformes[J]. Conservation Genetics Resources, 2017, 9(2): 265-268.
- [20] Austin C M, Tan M H, Croft L J, et al. Whole genome sequencing of the Asian arowana (*Scleropages formosus*) provides insights into the evolution of ray-finned fishes[J]. Genome Biology and Evolution, 2015, 7(10): 2885-2895.
- [21] Mu X D, Gu D G, Yang Y X, et al. Genetic diversity and phylogeny of the family Osteoglossidae by the nuclear 18S ribosomal RNA and implications for its conservation[J]. Biochemical Systematics and Ecology, 2013, 51: 280-287.
- [22] Duan Y N, Liu Y, Hu Y C, et al. Distribution regularity of microsatellites in *Scleropages formosus* genome[J]. Chinese Agricultural Science Bulletin, 2019, 35(23): 152-158. [段永楠, 刘奕, 胡隐昌, 等. 美丽硬仆骨舌鱼全基因组微卫星分布规律特征[J]. 中国农学通报, 2019, 35(23): 152-158.]
- [23] Wang Q L, Liu Y, Song H M, et al. Development and primer selection of EST-SSR molecular markers based on transcriptome sequencing of *Osteoglossum bicirrhosum*[J]. Freshwater Fisheries, 2016, 46(6): 8-13. [王且鲁, 刘奕, 宋红梅, 等. 双须骨舌鱼转录组 EST-SSR 标记开发与引物筛选[J]. 淡水渔业, 2016, 46(6): 8-13.]
- [24] Chase M W, Salamin N, Wilkinson M, et al. Land plants and DNA barcodes: Short-term and long-term goals[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2005, 360(1462): 1889-1895.
- [25] Zhou Y, Du X L, Zheng X, et al. ITS2 barcode for identifying the officinal rhubarb source plants from its adulterants[J]. Biochemical Systematics and Ecology, 2017, 70: 177-185.
- [26] 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. [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.]
- [27] 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.
- [28] Wu B S, Si L Z, Kong X Y, et al. Study on feature of ITS2 in 11 Perciformes species and the application on phylogenetic relationship[J]. Journal of Fishery Sciences of China, 2018, 25(6): 1151-1160. [武宝生, 司李真, 孔晓瑜, 等. 11 种鲈形目鱼类 ITS2 特征及系统应用[J]. 中国水产科学, 2018, 25(6): 1151-1160.]
- [29] Koren S, Walenz B P, Berlin K, et al. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation[J]. Genome Research, 2017, 27(5): 722-736.
- [30] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. Bioinformatics, 2009, 25(14): 1754-1760.
- [31] Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data[J]. Bioinformatics, 2011, 27(21): 2987-2993.
- [32] Walker B J, Abeel T, Shea T, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement[J]. PLoS One, 2014, 9(11): e112963.
- [33] 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.
- [34] Li Q, Wan J M. SSRHunter: development of a local searching software for SSR sites[J]. *Hereditas (Beijing)*, 2005, 27(5): 808-810. [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.]
- [35] Benson G. Tandem repeats finder: A program to analyze DNA sequences[J]. *Nucleic Acids Research*, 1999, 27(2): 573-580.
- [36] Ronquist F, Huelsenbeck J P. MrBayes 3: Bayesian phylogenetic inference under mixed models[J]. *Bioinformatics*, 2003, 19(12): 1572-1574.
- [37] Yang G, Yu J H, Xu P, et al. Phylogenetic relationship among 8 common species of catfish based on ribosome 18S and ITS sequences[J]. *Chinese Journal of Zoology*, 2010, 45(4): 110-117. [杨光, 俞菊华, 徐跑, 等. 利用 18S 和 ITS 序列揭示 8 种鲇形目鱼类的系统发育[J]. *动物学杂志*, 2010, 45(4): 110-117.]
- [38] Si L Z, Shi W, Yang M, et al. Feature analysis of the internal transcribed spacers in teleost ribosomal DNA[J]. *Journal of Tropical Oceanography*, 2016, 35(6): 74-81. [司李真, 时伟, 杨敏, 等. 硬骨鱼类核糖体基因间隔区的序列特征分析[J]. *热带海洋学报*, 2016, 35(6): 74-81.]
- [39] Galtier N, Piganeau G, Mouchiroud D, et al. GC-content evolution in mammalian genomes: The biased gene conversion hypothesis[J]. *Genetics*, 2001, 159(2): 907-911.
- [40] Yakovchuk P, Protozanova E, Frank-Kamenetskii M D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix[J]. *Nucleic Acids Research*, 2006, 34(2): 564-574.
- [41] Rodríguez-Trelles F, Tarrío R, Ayala F J. Evidence for a high ancestral GC content in *Drosophila*[J]. *Molecular Biology and Evolution*, 2000, 17(11): 1710-1717.
- [42] von der Schulenburg J H G, Hancock J M, Pagnamenta A, et al. Extreme length and length variation in the first ribosomal internal transcribed spacer of ladybird beetles (Coleoptera: Coccinellidae)[J]. *Molecular Biology and Evolution*, 2001, 18(4): 648-660.
- [43] Zhou A G, Xie S L, Chen J T, et al. Molecular phylogenetic relationships of five species of *Epinephelus* based on sequences of rDNA ITS1[J]. *Marine Sciences*, 2015, 39(8): 7-15. [周爱国, 谢少林, 陈金涛, 等. 5 种石斑鱼核糖体 DNA ITS1 序列及其分子系统进化关系[J]. *海洋科学*, 2015, 39(8): 7-15.]
- [44] Zhang Y Z, Wang W, Jiang Z Q. Research on molecular phylogenetic relationship of partial Blennioidei based on ITS-1 sequences[J]. *Modern Agricultural Science and Technology*, 2012(8): 318-320. [张源真, 王伟, 姜志强. 基于 ITS-1 序列的部分鲷亚目鱼类的分子系统进化关系研究[J]. *现代农业科技*, 2012(8): 318-320.]
- [45] 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. [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.]
- [46] Gong L, Shi W, Yang M, et al. Comparative analysis of the first internal transcribed spacer of ribosomal DNA in five Soleidae species[J]. *Journal of Fisheries of China*, 2017, 41(3): 321-329. [龚理, 时伟, 杨敏, 等. 5 种鲷科鱼类核糖体 ITS1 序列比较[J]. *水产学报*, 2017, 41(3): 321-329.]
- [47] Mu X D, Bai J J, Wang X J, et al. Cloning and sequence analysis of ribosomal DNA internal transcribed spacer 1 from goldfish (*Carassius auratus*)[J]. *Freshwater Fisheries*, 2008, 38(1): 74-76, 43. [牟希东, 白俊杰, 汪学杰, 等. 金鱼核糖体转录间隔区(ITS-1)的克隆和序列分析[J]. *淡水渔业*, 2008, 38(1): 74-76, 43.]
- [48] Yuan W A. Identification of fish species based on ribosomal DNA ITS2 locus[J]. *Hereditas (Beijing)*, 2010, 32(4): 369-374. [袁万安. 核糖体转录间隔区 2 应用于鱼类种属的鉴别[J]. *遗传*, 2010, 32(4): 369-374.]
- [49] Yang M, Kong X Y, Shi W, et al. Comparative analysis of ribosomal ITS2 sequences in 10 species of Pleuronectidae[J]. *Chinese Journal of Zoology*, 2018, 53(6): 938-950. [杨敏, 孔晓瑜, 时伟, 等. 鲆科 10 种鱼类核糖体 ITS2 序列比较分析[J]. *动物学杂志*, 2018, 53(6): 938-950.]
- [50] Schlötterer C, Hauser M T, von Haeseler A, et al. Comparative evolutionary analysis of rDNA ITS regions in *Drosophila*[J]. *Molecular Biology and Evolution*, 1994, 11(3): 513-522.
- [51] Tang J, Toè L, Back C, et al. Intra-specific heterogeneity of the rDNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex[J]. *Molecular Biology and Evolution*, 1996, 13(1): 244-252.
- [52] Fritz G N, Conn J, Cockburn A, et al. Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae)[J]. *Molecular Biology and Evolution*, 1994, 11(3): 406-416.
- [53] Mateos M, Markow T A. Ribosomal intergenic spacer (IGS) length variation across the Drosophilinae (Diptera: Drosophilidae)[J]. *BMC Evolutionary Biology*, 2005, 5: 46.
- [54] Yu J H, Li H X, Ge X P, et al. Characteristics analysis of complete ribosomal DNA sequence in Chinese mitten crab (*Eriocheir sinensis*)[J]. *Journal of Fisheries of China*, 2010, 34(5): 696-703. [俞菊华, 李红霞, 戈贤平, 等. 中华绒螯蟹核糖体 DNA 全序列分析[J]. *水产学报*, 2010, 34(5): 696-703.]
- [55] Guo Z S, Ding Y, Zhang X H, et al. Complete nuclear ribosomal DNA sequence analysis of Pacific abalone *Haliotis*

- discus Hannai*[J]. Fisheries Science, 2017, 83(5): 777-784.
- [56] Li L, Zhong C H, Bian Y B. The molecular diversity analysis of *Auricularia auricula-judae* in China by nuclear ribosomal DNA intergenic spacer[J]. Electronic Journal of Biotechnology, 2014, 17(1): 27-33.
- [57] He Y. Cloning and application of the nuclear ribosomal DNA cistron (nrDNA) sequence of two species of *Pyropia* in China[D]. Suzhou: Soochow University, 2016. [何渊. 我国两种栽培紫菜核糖体 RNA 基因的克隆与应用[D]. 苏州: 苏州大学, 2016.]
- [58] Su X. The complete ribosome sequence amplification and characteristic analysis of five species of trematode[D]. Daqing: Heilongjiang Bayi Agricultural University, 2017. [宿欣. 五种吸虫核糖体全序列的扩增与特征分析[D]. 大庆: 黑龙江八一农垦大学, 2017.]
- [59] Bernardi G, Olofsson B, Filipski J, et al. The mosaic genome of warm-blooded vertebrates[J]. Science, 1985, 228(4702): 953-958.

Characteristics of ribosomal gene of seven Osteoglossidae species and molecular phylogeny

YANG Yexin^{1,2}, LIU Yi¹, LIU Chao¹, SONG Hongmei^{1,3}, WANG Xuejie¹, MU Xidong¹

1. Key Laboratory of Leisure Fisheries, Ministry of Agriculture and Rural Affairs; Guangdong Engineering Technology Research Center for Advanced Recreational Fisheries; Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China;
2. Key Laboratory of Urban Agriculture, Ministry of Agriculture and Rural Affairs, Shanghai 200240, China;
3. Key Laboratory of Aquatic Animal Immune Technology of Guangdong Province, Guangzhou 510380, China

Abstract: Osteoglossidae, a family of ancient fishes known as “living fossils,” has high economic, cultural, and scientific value. In this study, we analyzed the molecular phylogeny of higher-order elements in the Osteoglossidae family using ribosomal DNA sequences. The complete sequences of the ribosomal gene of *Arapaima gigas*, *Heterotis niloticus*, *Osteoglossum bicirrhosum*, *O. ferreirai*, *Scleropages formosus*, *S. jardinii*, and *S. leichardti*, obtained using second-generation and third-generation sequencing methods, were as follows: 11714 bp, 8957 bp, 12057 bp, 11556 bp, 10377 bp, 10724 bp, and 10725 bp, respectively. The base content of GC ranged from 63.78% to 66.13%. The lengths of 18S rDNA were 1835 bp; GC content ranged from 54.06% to 54.5%. The genetic distance ranged from 0.000 to 0.024, and the sequence identity varied from 97.52% to 100%. The lengths of ITS1 rDNA were between 598 bp and 871 bp; GC content ranged from 72.6% to 78.3%. The genetic distance ranged from 0.000 to 0.852, and the sequence identity varied from 38.34% to 99.5%. The lengths of 5.8S rDNA were 159 bp; GC content ranged from 57.23% to 57.86%. The genetic distance ranged from 0.000 to 0.019, and the sequence identity varied from 98.11% to 100%. The lengths of ITS2 rDNA were between 327 bp and 612 bp; GC content ranged from 70.34% to 80.07%. The genetic distance ranged from 0.000 to 0.799, and the sequence identity varied from 31.63% to 99.53%. The lengths of 28S rDNA were between 3996 bp and 4419 bp; GC content ranged from 62.42% to 65.69%. The genetic distance ranged from 0.000 to 0.066, and the sequence identity varied from 84.98% to 99.95%. The lengths of IGS rDNA were between 1815 bp and 4679 bp; GC content ranged from 63.73% to 71.52%. The genetic distance ranged from 0.001 to 1.877, and the sequence identity varied from 19.31% to 98.31%. The phylogenetic tree was constructed using the neighbor-joining method and Bayesian inference method on the 18S, ITS1, ITS2, and 28S single sequences and the 18S+ITS1+5.8S+ITS2+28S combined sequence. The results showed that the topological structure of the two trees was consistent. Moreover, this result is concordant with the traditional classification of seven species of bony tongue fish, except for ITS1, which can be used as a molecular marker of Osteoglossidae. Our findings provide data for the phylogenetic study of ancient fishes and to enrich the diversity and evolution mode research of fish using ribosomal RNA genes.

Key words: Osteoglossidae; ribosomal gene; 18S rDNA; 5.8S rDNA; 28S rDNA; ITS; IGS; phylogenetic analysis

Corresponding author: MU Xidong. E-mail: muxd@prfri.ac.cn