

海南万泉河口尖鳍鲤线粒体基因组测序与结构特征分析

蒙雨¹, 宋一清^{1,2}, 谢青颖¹, 蔡杏伟^{3,4}, 张清凤^{3,4}, 董杨^{3,4}, 李芳远^{3,4},
谢松光^{1,2}, 申志新^{3,4}

(1. 海南大学海洋生物与水产学院, 海南 海口 570228; 2. 海南大学南海海洋资源利用国家重点实验室, 海南 海口 570228; 3. 海南省海洋与渔业科学院, 海南 海口 571126; 4. 海南淡水生物资源及生态环境保护研究中心, 海南 海口 571126)

摘要: 为了解海南尖鳍鲤(*Cyprinus acutidorsalis*)线粒体基因组结构特征及其与鲤科(Cyprinidae)鱼类的进化关系, 本文通过二代测序获得海南万泉河口尖鳍鲤线粒体基因组全序列, 并对结构特征进行分析, 结果显示具有典型的环状闭环双链分子, 全长 16 581 bp, 碱基组成为 A (31.96%)、G(15.69%)、C(27.53%)和 T(24.82%), 包含 13 个蛋白质编码基因(protein-coding genes, PCGs)、2 个 rRNA 基因、22 个 tRNA 基因和 1 个可变控制区 D 环。对碱基含量分析发现尖鳍鲤的碱基组成中具有较高 AT 含量的偏向性, 13 个 PCGs 中有不少的偏好性密码子, 如 CGA(RSCU>3.506)、CCA(RSCU>2.208)等。除 *tRNA^{Gln}*、*tRNA^{Ala}*、*tRNA^{Asn}*、*tRNA^{Cys}*、*tRNA^{Leu}*(CUN)、*tRNA^{Tyr}*、*tRNA^{Ser}*(UCN)、*tRNA^{Pro}* 和 *ND6* 外, 其余基因的排列于 H 链上。比较海南与广西海域尖鳍鲤线粒体基因组发现, 两者 PCGs 相似性在 96%~100%, *ND4*、*CO II*、*Cytb* 基因在起始位置、长度以及起始密码子存在不同程度的差异。另外, 将与尖鳍鲤同属的鲤(*C. carpio*)、鲫(*C. auratus*)的 13 个 PCGs 进行两两比较表明, 3 种鲤属鱼类的 PCGs 相似性在 85%~99%。通过与其他鲤科鱼类的线粒体基因组序列构建系统发育树, 发现尖鳍鲤和鲤亲缘关系最近。本研究结果有助于理解不同鲤科鱼类的亲缘关系, 为尖鳍鲤线粒体进化分析和种质资源研究提供参考依据。

关键词: 尖鳍鲤(*Cyprinus acutidorsalis*); 线粒体基因组; 系统发育; 序列分析

中图分类号: Q959 文献标识码: A 文章编号: 1000-3096(2023)11-0066-13

DOI: 10.11759/hyxx20230720001

线粒体基因组(mtDNA)是独立存在于高等动物细胞核以外的唯一遗传物质, 具有结构简单、基因排列紧凑、容易扩增、以母系遗传为主、进化率高等特点^[1]。鱼类线粒体基因组是一个双链闭环环状 DNA 分子, 大小占整个 DNA 约为 1%^[2], 长度通常在 16 kb 左右, 且排列结构一般非常保守, 十分紧凑。尽管鱼类线粒体基因组整体上是高度保守的, 但某些区域仍然可以表现出高度变异性^[3], 这些区域通常是非编码区域或控制区域, 例如重复序列、转座子和调节元件。除了少数物种外, 大部分动物线粒体基因组都具有相同的 37 个基因, 包括典型的基因互补编码 13 个氧化磷酸化酶的蛋白质亚单位、线粒体核糖体的两个 rRNA, 以及翻译线粒体基因组编码的蛋白质所需的 22 个 tRNAs^[4]。相邻基因没有间隔或存在两种基因前端和后端碱基重叠的现象。线粒体基因组具有自己的 DNA 复制、转录、mRNA 处理和

蛋白质翻译系统, 可以作为基因组进化的模型^[5]。

随着 Sanger 测序方法的成熟和第二代和第三代测序技术的不断改进, 线粒体基因组的测序变得简单快捷, 已成为种类鉴定、系统发育以及群体遗传分化等领域广泛应用的分子标记之一^[6], 为鱼类种质资源的保护提供了有效手段^[7], 进化最快的 D-loop

收稿日期: 2023-07-20; 修回日期: 2023-10-07

基金项目: 海南省科技专项项目(ZDYF2023SHFZ101, ZDYF2021XDNY299); 海南省教育厅项目(Hnkyzc2022-4); 海南大学科研启动基金项目(KYQD(ZR)21130)

[Foundation: Hainan Province Science and Technology Special Fund, No.ZDYF2023SHFZ101, No. ZDYF2021XDNY299; Project Supported by the Education Department of Hainan Province, No. Hnkyzc2022-4; Scientific Research Fund of Hainan University for R&D, No.KYQD(ZR)21130]

作者简介: 蒙雨(2001—), 女, 海南三亚人, 硕士, 主要从事鱼类生态学, E-mail: 1554588619@qq.com; 宋一清(1988—), 女, 通信作者, 吉林梨树人, 博士, 副研究员, 主要从事鱼类生态学研究, E-mail: songyq@hainanu.edu.cn; 申志新(1964—), 男, 通信作者, 青海西宁人, 博士, 研究员, 主要从事渔业资源与保护研究, E-mail: shen_266@msn.com

区可以用于研究某区域的种群遗传多样性水平^[8]; *CO I* 基因也是许多鱼类分类和鉴定的理想 DNA 条形码^[9]。通过对不同物种的线粒体基因组比较分析不仅可以研究适应性进化,也可以追踪分析特异物种的起源和进化^[10]。结合形态学分析和线粒体基因组母系遗传和序列变异迅速的特征重建物种的系统发育关系,为解决物种起源和分化问题提供了更好的遗传学证据^[11]。

尖鳍鲤(*Cyprinus acutidorsalis*)隶属于鲤形目(Cypriniformes)、鲤科(Cyprinidae)、鲤亚科(Cyprininae)、鲤属(*Cyprinus*)。鲤亚属,又名海鲤,广西俗称“海鲫”,其种群仅分布在海南万泉河、南渡江河口以及广西的钦江河口,分布范围狭窄^[12],目前是中国特有的生活于河口生态系统的鲤科鱼类^[13]。由于人类活动干扰等原因,尖鳍鲤的适宜栖息地大面积缩小,种质极度退化,种群数量锐减,仅有的几处河口已较难发现其踪迹,处于极危状态^[14]。

有关尖鳍鲤的研究主要集中于早期发育和养殖技术,研究区域仅分布在广西钦江河口^[15-16]。前期有少量研究基于线粒体基因的片段揭示了尖鳍鲤在鲤形目鱼类的发育地位和遗传多样性^[17-18],关于线粒体基因组的结构特征还鲜有报道。本研究利用高通量第二代测序技术对进行测序,组装与注释尖鳍鲤线粒体全基因组序列,分析海南和广西水域尖鳍鲤线粒体基因的相似性,并比较了尖鳍鲤与鲤、鲫的 13 个 PCGs 碱基组成差异等线粒体基因组特征,构建尖鳍鲤与其他 16 种鲤科鱼类的系统发育树。研究结果将为海南尖鳍鲤种类鉴定和系统进化提供科学依据。

1 材料与方法

1.1 实验材料及 DNA 的提取

尖鳍鲤样本采集自海南省琼海市博鳌镇万泉河口沙美内海。剪取背部肌肉样本放入无水乙醇中保存。使用动物组织基因组试剂盒(OMEGA),根据说明书的操作步骤提取总 DNA,通过琼脂糖凝胶电泳和核酸蛋白测定仪对其 DNA 纯度和浓度进行测定。

1.2 基因组测序

委托上海凌恩生物科技有限公司进行尖鳍鲤的线粒体基因组测序,具体实验步骤如下:将提取的 DNA 样品以 1 μg 起始量构建文库,使用 Covaris M220 超声打断分割 DNA 长度为 300~500 bp 的片段

并在两端加接头(Illumina),采用 TBS380(Invitrogen)定量,将文库回收(Bio-Rad),再根据数据比例混合后上机,在 cBot 固相载体上进行桥式 PCR 扩增生成 clusters,在 Illumina NovaSeq 测序平台(Truseq SBS Kit),进行 2×150 bp 上机测序(300 cycles)。

1.3 基因组组装与分析

用 SPAdes v3.14.1(<http://bioinf.spbau.ru/spades>)软件对 Clean Data 进行拼接。挑选覆盖深度足够高且组装长度较长的序列作为候选序列,比对 NT 库确认线粒体 Scaffold 序列,并根据 overlap 连接序列。根据参考基因组确定线粒体组装序列的起始位置和方向,获得尖鳍鲤的线粒体基因组序列。

使用 MITOS(<http://mitos.bioinf.uni-leipzig.de/index.py>)软件对线粒体基因组进行编码蛋白、tRNA 和 rRNA 基因的预测。然后对 MITOS 预测的初始基因去冗余,并人工校正基因的起始、终止密码子位置,获得高准确性的保守基因集。使用 cusp (EMBOSS v6.6.0.0)软件获得密码子的偏好性值。利用软件 CGView(<http://stothard.afns.ualberta.ca/cgviewserver/>)对样品基因组进行圈图展示。尖鳍鲤线粒体的基因蛋白序列与已知的蛋白数据库进行 blastp($evalue < 1e-5$)比对,进行 NR、GO、Swiss-Prot 注释分析。

从 NCBI (<https://www.ncbi.nlm.nih.gov/>)上获取 16 种鲤科鱼类的线粒体基因组序列(表 1),以花鳉(序列号:NC_006540.1)为外类群,通过 Clustal W 进行序列多重比对,使用 MEGA11.0 软件采用邻接法构建 Neighbor-joining 系统进化树,选择 Bootstrap 法检验(1 000 次循环)各分支的置信度,同时基于 Kimura2-parameter 计算 16 种鲤科鱼类间的遗传距离。

2 结果与分析

2.1 线粒体基因组结构特征

尖鳍鲤线粒体全基因组序列为典型的环状结构,线粒体基因组全长 16 581 bp, PCGs 总长度 11 409 bp,平均长度 878 bp,其长度占基因组的 68.81%。碱基组成为 A (31.96%)、G(15.69%)、C(27.53%)和 T(24.82%); A+T 含量为 56.78%,显著大于 G+C 含量(43.22%)(表 2),表现为明显的 AT 偏好性。尖鳍鲤线粒体基因组的排列顺序和结果组成与其他鱼类一致,共 38 个基因(图 1):包括 13 个 PCGs, 22 个 tRNA 基因, 2 个 rRNA 基因(12S-rRNA 和 16S-rRNA)和一个非编码控制区(D-loop 区)。

表 1 用于系统发育分析的鲤科鱼类及花鳗鲡物种列表

Tab. 1 List of the Cyprinidae and *Anguilla marmorata* species for the phylogenetic analysis

物种名	NCBI 序列登录号	总长(bp)
细尾白甲鱼(<i>Onychostoma lepturum</i>)	NC_054158.1	16 601
虹彩光唇鱼(<i>Acrossocheilus iridescens</i>)	NC_031551.1	16 596
越南鱖(<i>Acheilognathus tonkinensis</i>)	NC_042407.1	16 767
东方墨头鱼(<i>Garra orientalis</i> Nichols)	NC_021935.1	17 288
麦穗鱼(<i>Pseudorasbora parva</i>)	NC_015614.1	16 600
高体鳊鲂(<i>Rhodeus ocellatus</i>)	NC_011211.1	16 680
条纹小鲃(<i>Puntius semifasciolatus</i>)	NC_020096.1	16 954
光倒刺鲃(<i>Spinibarbus hollandi</i> Oshima)	NC_026129.1	16 521
点纹银鲃(<i>Squalidus wolterstorffi</i>)	NC_022190.1	16 602
南方白甲鱼(<i>Onychostoma gerlachi</i>)	NC_026549.1	16 601
露丝塔野鲮(<i>Labeo rohita</i>)	NC_017608.1	16 626
麦瑞加拉鲮(<i>Cirrhinus cirrhosus</i>)	NC_033964.1	16 588
鲤(<i>Cyprinus carpio</i>)	NC_001606.1	16 575
鲫(<i>Carassius auratus</i>)	CM010491.1	16 579
间鲮(<i>Hemibarbus medius</i> Yue)	NC_024527.1	16 614
花鳗鲡(<i>Anguilla marmorata</i>)	NC_006540.1	16 745

表 2 尖鳍鲤线粒体基因组信息

Tab. 2 Mitochondrial genome information of *C. acutidorsalis*

区域	长度/bp	T/U/%	C/%	A/%	G/%	A+T/%	G+C/%
线粒体基因组	16 581	24.82	27.53	31.96	15.69	56.78	43.22
蛋白质编码基因	11 409	26.68	28.28	29.93	15.11	56.61	43.39
第一位点	3 803	26.14	27.03	30.84	15.99	56.98	43.02
第二位点	3 803	27.22	30.5	27.93	14.36	55.14	44.86
第三位点	3 803	26.69	27.29	31.03	14.99	57.72	42.28
tRNA	1 562	27.08	20.87	28.94	23.11	56.02	43.98

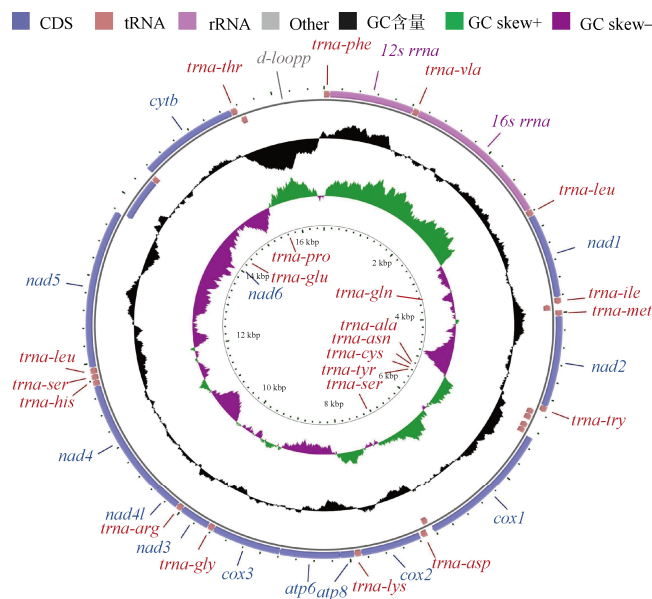


图 1 尖鳍鲤线粒体基因组图谱

Fig. 1 Mitochondrial genome map of *C. acutidorsalis*

2.2 尖鳍鲤线粒体基因组 rRNA、tRNA 及 D-loop 区结构分析

22 个 tRNA 的序列总长度为 1 562 bp, 平均长度为 71 bp。2 个 rRNA 的序列总长度为 2 632 bp。尖鳍鲤的线粒体基因组、PCGs、第一位点、第二位点、第三位点、tRNA 的碱基 AT 含量均比 GC 含量高(表 2)。线粒体控制区(D-loop 区)结构位于基因 *tRNA^{Phe}* 和 *tRNA^{Pro}* 之间, 长度为 927 bp, A+T 含量为 69.6%。尖鳍鲤 12S rRNA 位于 *tRNA^{Phe}* 和 *tRNA^{Val}* 之间, 由

952 个碱基组成; 16S rRNA 位于 *tRNA-Val* 和 *tRNA-Leu* 之间, 由 1 679 个碱基组成, 12S rRNA 比 16S rRNA 更加保守。基因 *tRNA^{Gln}*、*tRNA^{Trp}*、*ATP6*、*CO III*、*tRNA^{Gly}*、*tRNA^{Arg}*、*ND4*、*ND6*、*tRNA^{Pro}* 的存在着碱基重叠的现象, 重叠碱基数 1~7bp, 最长的重叠区位于 *ATP8* 和 *ATP6* 基因、*ND4L* 和 *ND4* 基因之间。除 *tRNA^{Gln}*、*tRNA^{Ala}*、*tRNA^{Asn}*、*ND6*、*tRNA^{Cys}*、*tRNA^{Leu}*(CUN)、*tRNA^{Tyr}*、*tRNA^{Ser}*(UCN)、*tRNA^{Pro}* 外, 其余基因的排列于 H 链上(表 3)。

表 3 海南万泉河口尖鳍鲤线粒体基因组结构特征

Tab. 3 Mitochondrial genome structure of *C. acutidorsalis* from the Wanquan River Estuary in Hainan

基因	起始位置	长度 /bp	编码链	间隔片段 长度/bp	起始密 码子	终止密 码子	基因	起始位置	长度 /bp	编码 链	间隔片段 长度/bp	起始密 码子	终止密 码子
<i>tRNA^{Phe}</i>	1~69	69	H	—	—	—	<i>ATP8</i>	7 954~8 118	165	H	1	ATG	TAG
12S rRNA	70~1 022	952	H	0	—	—	<i>ATP6</i>	8 112~8 795	684	H	-7	ATG	TAA
<i>tRNA^{Val}</i>	1 025~1 096	72	H	2	—	—	<i>CO III</i>	8 795~9 580	786	H	0	ATG	TAA
16S rRNA	1 097~2 775	1679	H	0	—	—	<i>tRNA^{Gly}</i>	9 580~9 651	72	H	-1	—	—
<i>tRNA^{Leu}</i> (UUR)	2 776~2 851	76	H	0	—	—	<i>ND3</i>	9 652~10 002	351	H	0	ATG	TAG
<i>ND1</i>	2 853~3 827	975	H	1	ATG	TAA	<i>tRNA^{Arg}</i>	10 001~10 070	70	H	-2	—	—
<i>tRNA^{Ile}</i>	3 832~3 903	72	H	4	—	—	<i>ND4L</i>	10 071~10 367	297	H	0	ATG	TAA
<i>tRNA^{Gln}</i>	3 902~3 972	71	L	-2	—	—	<i>ND4</i>	10 361~11 741	1381	H	-7	ATG	T
<i>tRNA^{Met}</i>	3 975~4 043	69	H	2	—	—	<i>tRNA^{His}</i>	11 742~11 810	69	H	0	—	—
<i>ND2</i>	4 044~5 090	1047	H	0	ATG	TAG	<i>tRNA^{Ser}</i> (AGY)	11 811~11 879	69	H	0	—	—
<i>tRNA^{Trp}</i>	5 089~5 159	71	H	-2	—	—	<i>tRNA^{Leu}</i>	11 881~11 953	73	H	1	—	—
<i>tRNA^{Ala}</i>	5 162~5 230	69	L	2	—	—	<i>ND5</i>	11 957~13 780	1824	H	3	ATG	TAA
<i>tRNA^{Asn}</i>	5 232~5 304	73	L	1	—	—	<i>ND6</i>	13 777~14 298	522	L	-4	ATG	TAA
<i>tRNA^{Cys}</i>	5 338~5 404	67	L	33	—	—	<i>tRNA^{Leu}</i> (CUN)	14 299~14 367	69	L	0	—	—
<i>tRNA^{Tyr}</i>	5 405~5 474	70	L	0	—	—	<i>Cytb</i>	14 373~15 513	1141	H	5	ATG	T
<i>CO I</i>	5 482~7 026	1545	H	7	ATC	TAA	<i>tRNA^{Thr}</i>	15 514~15 585	72	H	0	—	—
<i>tRNA^{Ser}</i> (UCN)	7 027~7 097	71	L	0	—	—	<i>tRNA^{Pro}</i>	15 585~15 654	70	L	-1	—	—
<i>tRNA^{Asp}</i>	7 101~7 172	72	H	3	—	—	<i>tRNA^{Lys}</i>	7 877~7 952	76	H	0	—	—
<i>CO II</i>	7 186~>787	691	H	13	ATG	T	<i>D-loop</i>	15 655~16 581	927	H	0	—	—

2.3 尖鳍鲤密码子使用的相对频率(RSCU)

对 13 个 PCGs 的 RSCU 值统计显示(图 2), CGA(Arg)、CTA(Leu)、TCA(Ser)和 CCA(Pro)是最常用的 4 种密码子, 其中, CGA 的含量最高(3.506%), CTA 含量为 2.87%, 其他 2 种密码子的含量均在 2%以上。另外, 尖鳍鲤线粒体基因组中编码同一氨基酸的密码

子的 CGA、CGC、AAC、TGC、CAA、CAC、ATT、CTA、AAA、ATG、TTC、CCA、AGC、TCA、TCC、ACA、ACC、TGG、TAC 的 RSCU 均大于 1, 共 19 个, 均为偏好性(Codon Usage Bias)密码子, 其余的密码子有较弱的使用偏性。尖鳍鲤的高频密码子中有 11 个是以 C/G 碱基结尾, 其余 8 个以 A/T 碱基结尾(表 4)。

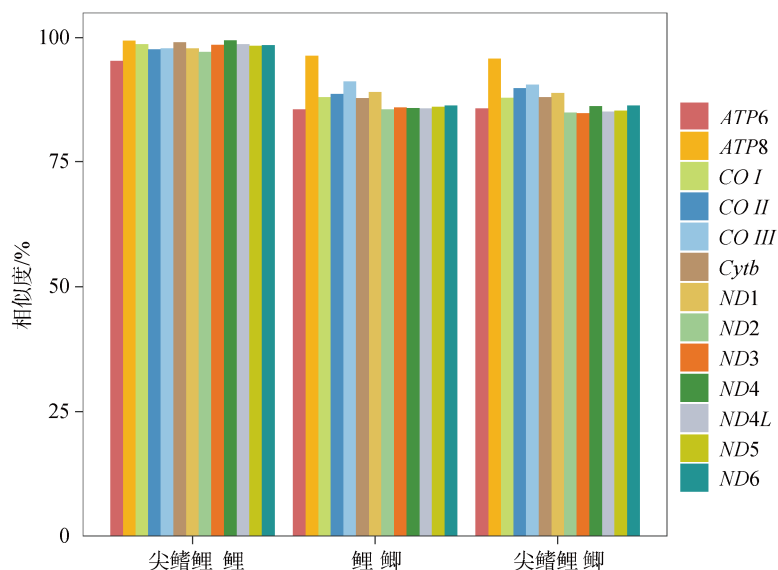


图 2 尖鳍鲤线粒体基因组密码子的偏好性

Fig. 2 Codon bias in the mitochondrial genome of *C. acutidorsalis*

表 4 尖鳍鲤线粒体基因组密码子的使用

Tab. 4 Use of codons in the mitochondrial genome of *C. acutidorsalis*

密码子	氨基酸	RSCU	tRNA	密码子	氨基酸	RSCU	tRNA
CGA	Arg	3.506	<i>tRNA^{Arg}</i>	ATA	Met	0.904	
CGC	Arg	1.091		ATG	Met	1.000	<i>tRNA^{Met}</i>
CGG	Arg	0.623		TTC	Phe	1.300	<i>tRNA^{Phe}</i>
CGT	Arg	0.701		TTT	Phe	0.700	
AAC	Asn	1.267	<i>tRNA^{Asn}</i>	CCA	Pro	2.208	<i>tRNA^{Pro}</i>
AAT	Asn	0.733		CCC	Pro	0.981	
TGC	Cys	1.600	<i>tRNA^{Cys}</i>	CCG	Pro	0.189	
TGT	Cys	0.400		CCT	Pro	0.623	
CAA	Gln	1.960	<i>tRNA^{Gln}</i>	AGC	Ser	1.114	<i>tRNA^{Ser}</i>
CAG	Gln	0.040		AGT	Ser	0.203	
CAC	His	1.577	<i>tRNA^{His}</i>	TCA	Ser	2.203	<i>tRNA^{Ser}</i>
CAT	His	0.423		TCC	Ser	1.519	
ATC	Ile	0.988	<i>tRNA^{Ile}</i>	TCG	Ser	0.228	
ATT	Ile	1.108		TCT	Ser	0.734	
CTA	Leu	2.870	<i>tRNA^{Leu}</i>	ACA	Thr	1.967	<i>tRNA^{Thr}</i>
CTC	Leu	0.970		ACC	Thr	1.365	
CTG	Leu	0.346		ACG	Thr	0.107	
CTT	Leu	0.758		ACT	Thr	0.562	
TTA	Leu	0.941	<i>tRNA^{Leu}</i>	TGG	Trp	1.000	
TTG	Leu	0.115		TAC	Tyr	1.096	<i>tRNA^{Tyr}</i>
AAA	Lys	1.870	<i>tRNA^{Lys}</i>	TAT	Tyr	0.904	
AAG	Lys	0.130					

2.4 海南与广西尖鳍鲤 PCGs 差异与相似性

海南和广西尖鳍鲤 13 个 PCGs 序列相似性都在 95%

以上(表 5), 其中 *ATP8* 和 *Cytb* 相似性超过 99%, 主要差异体现在碱基重叠的长度大小。海南岛尖鳍鲤的起始密

码子除了 *CO I* 是 ATC, 其他线粒体 PCGs 的起始密码子为 ATG, 而广西尖鳍鲤的 *CO I* 基因的起始密码子为 GTG, 其余同为 ATC。广西尖鳍鲤蛋白编码基因存在 TA 这类

不完全终止密码子。海南尖鳍鲤的 *CO II*、*ND4*、和 *Cytb* 基因的不完全终止密码子为 “T”, 广西尖鳍鲤 PCGs 的不完全终止密码子(T、TA)出现的频率比海南尖鳍鲤高。

表 5 海南与广西尖鳍鲤的 PCGs 相似性比较

Tab. 5 Comparison of the protein-coding genes between Hainan and Guangxi *C. acutidorsalis*

基因	海南尖鳍鲤				广西尖鳍鲤				PCGs 相似性/%
	起止范围/bp	长度/bp	起始密码子	终止密码子	起止范围/bp	长度/bp	起始密码子	终止密码子	
<i>ND1</i>	2 853~3 827	975	ATG	TAA	2 853~3 827	975	ATG	TAA	98.36
<i>ND2</i>	4 044~5 090	1 047	ATG	TAG	4 044~5 088	1 045	ATG	T	97.71
<i>CO I</i>	5 482~7 026	1 545	ATC	TAA	5 476~7 026	1 551	GTG	TAA	98.71
<i>CO II</i>	7 186~7 876	691	ATG	T	7 186~7 876	691	ATG	T	98.99
<i>ATP8</i>	7 954~8 118	165	ATG	TAG	7 954~8 118	165	ATG	TAG	99.39
<i>ATP6</i>	8 112~8 795	684	ATG	TAA	8 112~8 795	684	ATG	TAA	96.63
<i>CO III</i>	8 795~9 580	786	ATG	TAA	8 795~9 579	785	ATG	TA	97.96
<i>ND3</i>	9 652~10 002	351	ATG	TAG	9 652~10 000	349	ATG	T	98.58
<i>ND4L</i>	10 071~10 367	297	ATG	TAA	10 071~1 0367	297	ATG	TAA	98.65
<i>ND4</i>	10 361~11 741	1 381	ATG	T	10 361~11 741	1 381	ATG	T	98.41
<i>ND5</i>	11 956~13 780	1 824	ATG	TAA	11 957~13 780	1 823	ATG	TAA	98.63
<i>ND6</i>	13 777~14 298	522	ATG	TAA	13 777~14 298	522	ATG	TAA	98.28
<i>Cytb</i>	14 373~15 513	1 141	ATG	T	14 373~15 513	1 141	ATG	T	99.91

2.5 尖鳍鲤、鲤和鲫线粒体基因组比较分析

尖鳍鲤、鲤、鲫的线粒体基因排列顺序与其他大多数鱼类线粒体基因组一样均含有 13 个 PCGs、2 种 rRNA 基因和 22 种 tRNA 基因及非编码区的轻链复制起始区和复制和转录序列(D-loop)。3 种鲤科鱼类的线粒体基因长度十分相似, 并且 13 个 PCGs 中的部分基因的起始与终止密码子相同(表 5)。3 种鲤科鱼类的 PCGs 的 AT 密码子的偏好性都在 50%以上, 均具有 AT 偏好性。3 种鲤亚科

鱼类的起始密码子 ATG 和终止密码子 TAA 出现频率较高, 存在着完整和不完整两类终止密码子表 6。对 3 种鲤属鱼类的线粒体全基因组 13 个 PCGs 的编码序列进行两两比较, 结果显示 3 种鲤属鱼类的 PCGs 相似性在 85%~99%(图 3), 尖鳍鲤和鲤的亲缘关系较鲫更近。尖鳍鲤和鲤的基因 *ATP8*、*ND4*、*Cytb* 两两比对结果之间的相似性都在 99%以上(表 7), 其中 *ATP8* 是 13 个 PCGs 中相似性最高的基因。

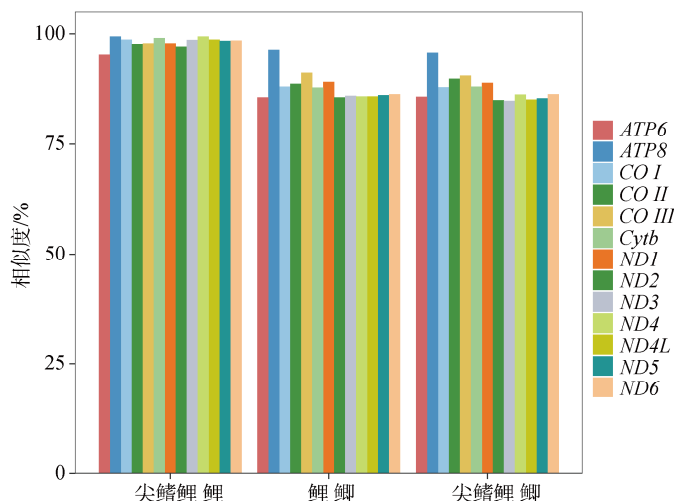


图 3 3 种鲤亚科鱼类的 13 个蛋白编码基因之间的相似性比较

Fig. 3 Similarity of 13 protein-coding genes among the three Cyprininae species

表 6 3 种鲤科鱼类线粒体基因组结构特点

Tab. 6 Mitochondrial genome structure of three species of Cyprinidae

基因	尖鳍鲤			鲫			鲤			起始密码子	终止密码子
	起始	终止	长度/bp	起始	终止	长度/bp	起始	终止	长度/bp		
<i>tRNA^{Phe}</i>	1	69	69	924	992	69	929	997	69		
<i>12S rRNA</i>	70	1 022	952	993	1 946	954	929	1 948	1 019		
<i>tRNA^{Val}</i>	1 025	1 096	72	1 947	2 018	72	1 49	2 020	72		
<i>16S rRNA</i>	1 097	2 775	1 679	2 019	3 699	1 681	2 21	3 701	1 681		
<i>tRNA^{Leu}</i>	2 776	2 851	76	3 700	3 777	78	3 02	3 778	77		
<i>ND1</i>	2 853	3 827	975	3 778	4 752	975	3 79	4 753	975	ATG	TAA
<i>tRNA^{Ile}</i>	3 832	3 903	72	4 756	4 829	74	4 757	4 829	73		
<i>tRNA^{Gln}</i>	3 902	3 973	71	4 827	4 897	69	4 827	4 897	71		
<i>tRNA^{Met}</i>	3 975	4 043	69	4 899	4 967	69	4 900	4 968	69		
<i>ND2</i>	4 044	5 090	1 047	4 968	6 014	1 047	4 969	6 015	1 047	ATG	TAG
<i>tRNA^{Trp}</i>	5 089	5 159	71	6 013	6 083	71	6 014	6 084	71		
<i>tRNA^{Ala}</i>	5 162	5 230	69	6 086	6 154	69	6 087	6 154	68		
<i>tRNA^{Asn}</i>	5 232	5 304	73	6 156	6 228	73	6 156	6 228	73		
<i>tRNA^{Cys}</i>	5 338	5 404	67	6 261	6 331	71	6 261	6 329	69		
<i>tRNA^{Tyr}</i>	5 405	5 474	70	6 330	6 400	71	6 328	6 397	70		
<i>CO I</i>	5 482	7 026	1 545	6 402	7 952	1 551	6 399	7 964	1 566	ATC/GTG/GTG	TAA
<i>tRNA^{Ser}</i>	7 027	7 097	71	7 953	8 032	80	7 950	8 020	71		
<i>tRNA^{Asp}</i>	7 101	7 172	72	8 27	8 098	72	8 024	8 095	72		
<i>CO II</i>	7 186	7 876	691	8 11	8 801	691	8 109	8 799	691	ATG	T
<i>tRNA^{Lys}</i>	7 877	7 952	76	8 02	8 877	76	8 800	8 875	76		
<i>ATP8</i>	7 954	8 118	165	8 79	9 043	165	8 877	9 041	165	ATG	TAG
<i>ATP6</i>	8 112	8 795	684	9 37	9 720	694	9 035	9 718	684	ATG	TAA
<i>CO III</i>	8 795	9 580	786	9 20	10 505	786	9 718	10 503	786	ATG	TAA
<i>tRNA^{Gly}</i>	9 580	9 651	72	10 05	10 576	72	10 503	10 574	72		
<i>ND3</i>	9 652	10 002	351	10 77	10 927	351	10 575	10 925	351	ATG	TAG
<i>tRNA^{Arg}</i>	10 001	10 070	70	10 26	10 995	70	10 924	10 993	70		
<i>ND4L</i>	10 071	10 367	297	10 96	11 292	297	10 994	11 290	297	ATG	TAA
<i>ND4</i>	10 361	11 741	1 381	11 86	12 666	1 381	11 284	12 664	1 381	ATG	T
<i>tRNA^{His}</i>	11 742	11 810	69	12 667	12 735	69	12 665	12 733	69		
<i>tRNA^{Ser}</i>	11 811	11 879	69	12 736	12 804	69	12 734	12 802	69		
<i>tRNA^{Leu}</i>	11 881	11 953	73	12 806	12 878	73	12 734	12 802	69		
<i>ND5</i>	11 957	13 780	1 824	12 882	14 705	1824	12 880	14 703	1 824	ATG	TAA
<i>ND6</i>	13 777	14 298	522	14 702	15 223	522	14 700	15 218	519	ATG	TAA/TAG/TAA
<i>tRNA^{Leu}</i>	14 299	14 367	69	15 224	15 292	69	15 219	15 287	69		
<i>Cytb</i>	14 373	15 513	1 141	15 298	16 438	1141	15 293	16 433	1 141	ATG	T
<i>tRNA^{Thr}</i>	15 514	15 585	72	16 439	16 510	72	16 434	16 505	72		
<i>tRNA^{Pro}</i>	15 585	15 654	70	16 509	16 580	72	16 504	16 575	72		
<i>D-loop</i>	15 655	16 581	927	1	923	923	1	928	928		

2.6 系统进化分析

为进一步了解鲤科鱼类的系统进化关系,以花鳉为外类群,基于邻接法构建尖鳍鲤和其他鲤科鱼类的系统进化树(图 4)。该进化树由两大进化分支

构成,尖鳍鲤和鲤聚为一支,亲缘关系最近。由进化树可知尖鳍鲤和其他鲤鱼之间有较相近的亲缘关系,虹彩光唇鱼、条纹小鲃、光刺倒鲃、鲫、东方墨头鱼和麦穗鱼均为单独的一支。由 Kimura2-parameter

模型计算出的各鲤科鱼类间的遗传距离显示(表 8), 平均遗传距离为 0.184, 其中遗传距离最小为尖鳍鲤和鲤(0.046), 进一步表明二者亲缘关系十分接近。

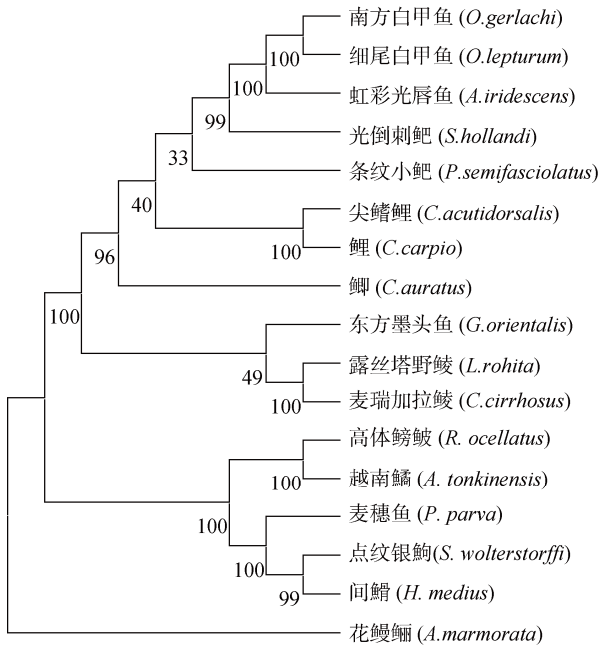


图 4 基于线粒体基因组全序列构建的 NJ 系统进化树
Fig. 4 NJ phylogenetic tree based on the complete mitochondrial genome sequence

3 讨论

3.1 尖鳍鲤的线粒体基因组的特征

尖鳍鲤线粒体基因组的基因分布组成和排列顺序与现有已知的鲤科鱼类相似, 各基因长度基本一致, 线粒体基因组中 PCGs 的密码子使用与其他脊椎动物相似, 可知线粒体基因组在进化上是高度保守的^[19]。海南尖鳍鲤线粒体基因由 37 个基因组成, 排列紧凑, 存在基因重叠的情况。尖鳍鲤中基因重叠区有 8 处, 重叠碱基数在 1~7 bp 间, 共 26 bp, 13 个蛋白编码重复区有 3 处, 这种基因重叠特征同样出现在其他鱼类, 如麦穗鱼^[20]、南方白甲鱼^[21], 其中个别基因重叠的碱基数略有差异, 如玻璃缺鳍鲶(*Kryptopterus vitreolus*)^[22]的 *ATP8* 和 *ATP6* 基因之间重叠为 10 bp。研究表明, 由于 DNA 双链之间的自然突变和选择压力的差异, 碱基在整个基因组上的分布往往是不均匀的^[23]。在线粒体基因组全序列分析中, 密码子 AT 碱基偏好性和反 G 现象, 且存在 RSCU 大于 1 的高频密码子, 如 TCC、TCA 等^[24]。在 19 个偏好性密码子中以 G 或 C 碱基结尾的占了大约 58%, 说明以 G 或 C 碱基结尾比以 A 或 T 碱基结尾出现的频率要高一些, 且尖鳍鲤线粒体基因的 D-loop 区 A 和 T 碱基频率较高(69.6%)。线粒体基因

表 7 3 种鲤科鱼类线粒体基因组蛋白编码基因碱基组成与相似性

Tab. 7 Composition and similarity of the mitochondrial genome protein-coding gene base in the three Cyprinidae species

种	碱基	ND1	ND2	CO I	CO II	ATP8	ATP6	CO III	ND3	ND4L	ND4	ND5	ND6	Cytb
尖 鳍 鲤	A+T	54.18	54.29	53.1	58.85	64.29	56.46	56.66	60.43	56.33	53.43	56.12	63.57	58.8
	T	23.27	20.36	24.05	28.23	41.43	22.97	27.14	23.74	28.74	23.43	23.47	40	27.57
	C	26.18	34.64	27.62	25.84	22.86	30.14	32.38	20.86	31.03	30	31.43	10.71	29.24
	A	30.91	33.93	29.05	30.62	25.71	33.49	29.52	36.69	27.59	30	32.65	13.57	31.23
	G	19.64	11.07	19.29	15.31	10	13.4	10.95	18.71	12.64	16.57	12.45	35.71	11.96
鲤	A+T	54.91	53.21	53.1	58.41	67.14	56.18	57.14	43.66	52.83	53.43	56.53	51.15	58.74
	T	24.36	20	23.81	27.86	40	23.71	27.62	18.31	28.57	23.43	23.06	39.29	28.37
	C	25.82	34.29	27.62	26.87	22.85	31.44	32.38	25.35	32.86	29.42	31.22	10	28.37
	A	30.55	33.21	29.29	30.85	27.14	32.47	29.52	38.03	24.26	30	33.47	12.86	30.37
	G	19.27	11.07	19.29	14.43	10	12.37	10.48	18.31	14.29	17.14	12.24	37.86	12.89
鲫	A+T	55.2	55.71	51.43	61.69	61.4	61.86	54.29	52.11	60.92	59.72	60.61	50.71	60.47
	T	25.09	20.71	23.57	29.35	17.14	29.9	25.24	16.9	34.48	28.29	26.94	37.14	30.9
	C	24.73	24.64	28.57	23.88	30	26.8	34.76	28.16	27.59	24.86	27.55	13.57	26.25
	A	30.11	35	27.86	32.34	44.26	31.96	29.05	35.21	26.44	31.43	33.67	13.57	29.57
	G	20.07	9.64	20	14.43	8.57	11.34	10.95	19.72	11.49	15.43	11.84	35.71	13.29
相似性		ND1	ND2	CO I	CO II	ATP8	ATP6	CO III	ND3	ND4L	ND4	ND5	ND6	Cytb
尖鳍鲤和鲤/%		97.85	97.13	98.71	97.68	99.39	95.32	97.84	98.58	98.68	99.42	98.36	98.47	99.04
鲤和鲫/%		89.13	85.67	88.07	88.71	96.36	85.67	91.22	86.04	85.86	85.88	86.13	86.40	87.91
尖鳍鲤和鲫/%		88.92	85.00	87.94	89.87	95.76	85.82	90.59	84.90	85.19	86.31	85.42	86.40	88.08

表 8 基于线粒体基因组的 16 种鲤科鱼类间的遗传距离

Tab. 8 Genetic distance between 16 Cyprinidae species based on the mitochondrial genome

种名	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. 鲫鱼																
2. 尖鳍鲤	0.148															
3. 鲤	0.113	0.046														
4. 高体鳊	0.224	0.264	0.229													
5. 麦穗鱼	0.197	0.238	0.203	0.231												
6. 露丝塔野鲮	0.136	0.151	0.118	0.228	0.198											
7. 条纹小鲃	0.128	0.156	0.121	0.236	0.208	0.132										
8. 东方墨头鱼	0.150	0.180	0.146	0.249	0.219	0.143	0.166									
9. 点纹银鲃	0.175	0.213	0.177	0.212	0.168	0.169	0.184	0.199								
10. 间鳍	0.181	0.220	0.184	0.224	0.172	0.180	0.193	0.213	0.139							
11. 光倒刺鲃	0.139	0.168	0.133	0.252	0.222	0.145	0.151	0.176	0.201	0.210						
12. 南方白甲鱼	0.148	0.172	0.136	0.249	0.232	0.146	0.148	0.179	0.200	0.212	0.157					
13. 虹彩光唇鱼	0.142	0.168	0.133	0.247	0.226	0.143	0.144	0.175	0.196	0.209	0.148	0.108				
14. 麦瑞加拉鲮	0.127	0.150	0.114	0.237	0.211	0.091	0.133	0.152	0.176	0.188	0.150	0.149	0.145			
15. 越南鲮	0.214	0.257	0.221	0.197	0.217	0.216	0.231	0.242	0.208	0.214	0.244	0.240	0.242	0.231		
16. 细尾白甲鱼	0.148	0.172	0.137	0.248	0.234	0.146	0.151	0.176	0.201	0.210	0.155	0.067	0.108	0.150	0.241	

组中碱基偏好性的原因可能是选择压力、自然突变等因素^[25]。这些特征与其他鲤科鱼类一样,如鲤科鳊属中华鳊(*Sarcocheilichthys sinensis*)、小鳊(*S. parvus*)、黑鳍鳊(*S. nigripinnis*)^[26]的 A+T 密码子富集在 D-loop 区,平均含量明显大于 G+C 含量。海南尖鳍鲤线粒体 *ND6* 基因在 L 链上,其余 PCGs 分布于 H 链上。研究表明,分布在 H 链上的线粒体基因倾向于形成单链保护,对水解和氧化更敏感^[27]。在尖鳍鲤 13 个 PCGs 中的 *CO I*、*ND4*、*Cytb* 基因中存在着 T 不完全终止密码子,再经过转录、加工补全成完整的终止密码子,提高了碱基的利用率,更加高效与节约^[28]。

3.2 海南和广西水域尖鳍鲤 PCGs 的差异性

海南尖鳍鲤线粒体基因组长度微小的差异可能是不同群体之间的非编码基因的长度多态性导致的^[2]。鱼类中起始密码子的使用可能存在特定的模式和差异,可能是地理种群隔离的结果^[29]。海南尖鳍鲤的 *ND2*、*CO III* 和 *ND3* 蛋白编码基因略长于广西尖鳍鲤,表现出位点多态性。生活在海南的尖鳍鲤和广西的尖鳍鲤关系十分密切,表现出高度相似性,但由于两者是在不同地理间的同一种群,仍会存在一定的差异性。这与不同地理种群的高体鳊(*Seriola dumerili*)^[30]是和暗色唇鲮(*Semilabeo obscura*)^[31]的研究结果相似,中国和日本海域高体鳊 *CO II* 表现出位点的多态性猜测是由地理种群隔离形成的。在海南和

广西水域尖鳍鲤差异性比较中发现 *CO I* 基因的起始密码子不同,前者为 ATG,后者为 GTC。在暗色唇鲮和瓯江高体鳊(*Rhodeus ocellatus*)^[32]线粒体基因组测序研究中,它们的 *CO I* 基因起始密码子为 GTC,而在孙利元等^[33]研究的 8 种石首鱼科(*Sciaenidae*)*CO I* 基因的起始密码子均为 ATG。

3.3 尖鳍鲤、鲤和鲫线粒基因组特征差异性

根据尖鳍鲤、鲤以及鲫的 13 个 PCGs 分别两两间进行多序列对比表明尖鳍鲤和鲤有着较高的相似度,并且这 3 种鱼都表现为明显的 AT 碱基的偏好性,推测原因是它们均属于鲤亚科,尖鳍鲤和鲤同属鲤亚科鲤属,而鲫属于须鲫属(*Carassioideus*)。3 种鱼 *ATP8* 基因 G 的含量都小于或等于 10%,存在着反 G 现象。尖鳍鲤的 *12S rRNA* 基因比鲤长 2 bp,比鲫长 67 bp,结合 3 种鲤亚科相似性分析说明尖鳍鲤和鲤的相似度更高,两者亲缘关系最近。3 种鱼类的 13 个 PCGs 中 *ATP8* 基因相似性最高,均在 95% 以上。尖鳍鲤和鲤的 *ND4* 基因相对最保守,相似度为高达 99.42%,其次是 *ATP8*(99.39%)和 *Cytb*(99.04%),表明 *ND4* 基因在两种鱼类中有都存在非常保守的功能和序列,受到的选择压力和功能束缚更强。这一结果也可以用来支持分类学上对尖鳍鲤和鲤的归属问题,进一步支持了它们可能属于同一种鱼类或者亲缘关系非常

接近的亚种的假设。而在对包括玻璃缺鳍鲶在内的 5 种鱼类核苷酸序列比对, *CO I* 基因相似度为 78.98%~92.78%, 相对最保守^[22]。由于海南和广西尖鳍鲤线粒体基因组的 PCGs 相似性都在 95% 以上, 差异不显著, 不适用于鉴别不同地理区域的尖鳍鲤, 但 PCGs 可以有效区分尖鳍鲤、鲤和鲫。在 3 种鲤亚科鱼类完整终止密码子有 TAA、TAG, 基因 *CO II*、*ND4*、*Cytb* 的终止密码子都为不完整的 T, 在同为鲤科鱼类麦穗鱼蛋白编码基因的密码子使用情况研究中, 麦穗鱼不完整终止密码子为 TA 或 T^[20], 南方白甲鱼^[21]完全终止密码子为 TAA 和 TAG, 不完全终止密码子与麦穗鱼相同。

3.4 尖鳍鲤系统发育分析

鲤科鱼类是淡水中最大的科, 其形态学鉴定是长期以来在分类学研究中相对困难的工作^[34]。鲤系分为野鲮亚科 (Labeoninae) 和鲤亚科 (鲤亚科 (Barbinae)、鲤亚科、裂腹鱼亚科 (Schizothoracinae)) 两大类, 鲤在分类上隶属于鲤形目、鲤亚目、鲤科、鲤属, 分布区间东起中国和西伯利亚、西至罗马, 是淡水硬骨鱼种最重要的一类^[35], 尖鳍鲤与鲤在分类学都属于鲤属。本研究的系统发育进化树中, 尖鳍鲤与鲤同属于一个进化支, 表明两者的形态分类和分子进化结果可能类似, 种间差异小, 亲缘关系最近。尖鳍鲤和鲤的遗传距离仅为 0.046, 并且两者的线粒体基因组 PCGs 相似性均大于 95%, 猜测可能具有相同或相似的基因组成, 从而使两者具有类似的形态特征和生理行为。在进化树中, 点纹银鲃 (*Squalidus wolterstorffi*) 和间鲃 (*Hemibarbus medius*) 处于系统树根部的位置, 而尖鳍鲤和鲤演化时间较晚。在姚东林等^[17]的研究中, 尖鳍鲤和三角鲤 (*C. multitaeniata*)、锦鲤 (*C. carpio haematopterus*) 亲缘关系也十分接近, 遗传距离分别为 0.003 和 0.013, 三者都归属于鲤亚科, 说明鲤亚科鱼类在遗传和进化方面存在高度的相似背景。尖鳍鲤和鲤两者在遗传上高度相似, 但确定它们确切的进化关系, 还需借助分子标记如 SSR (简单重复序列标记)、SNP (单核苷酸多态性) 等手段。

4 结论

本研究利用二代测序, 测定了海南尖鳍鲤的线粒体基因组全序列、基因含量、碱基组成、密码子偏好性等, 明确了尖鳍鲤线粒体基因基本结构和特

征。基于 DNAMAN 的 13 个 PCGs 序列比较表明, 海南尖鳍鲤与广西种群基因的相似性在 95% 以上。分析比较 3 种鲤科鱼类的线粒体基因组结构特征, 海南尖鳍鲤和鲤的亲缘关系最近。研究结果为海南尖鳍鲤种群遗传结构和系统进化分析提供依据, 为解释尖鳍鲤的起源和进化以及保护生物多样性提供理论基础。

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.
- [2] GUO X H, LIU S J, LIU Q, et al. New progresses on mitochondrial DNA in fish[J]. Acta Genetica Sinica, 2004, 31(9): 983-1000.
- [3] PROSDOCIMI F, DE CARVALHO D C, DE ALMEIDA R N, et al. The complete mitochondrial genome of two recently derived species of the fish genus *Nannoperca* (Perciformes, Percichthyidae)[J]. Molecular Biology Reports, 2012, 39(3): 2767-2772.
- [4] CHANG Y, HUANG F, LO T. The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome[J]. Journal of Molecular Evolution, 1994, 38(2): 138-155.
- [5] BERNT M, BRABAND A, SCHIERWATER B, et al. Genetic aspects of mitochondrial genome evolution[J]. Molecular Phylogenetics and Evolution, 2013, 69(2): 328-338.
- [6] DONG C, XU J, WANG B, et al. Phylogeny and evolution of multiple common carp (*Cyprinus carpio* L.) populations clarified by phylogenetic analysis based on complete mitochondrial genomes[J]. Marine Biotechnology, 2015, 17(5): 565-575.
- [7] BROUGHTON R E, MILAM J E, ROE B A. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA[J]. Genome Research, 2001, 11(11): 1958-1967.
- [8] LIU J, DING X, ZENG Y, et al. Genetic diversity and phylogenetic evolution of Tibetan sheep based on mtDNA D-Loop sequences[J]. PLoS One, 2016, 11(7): e0159308.
- [9] BING P X, HE S L, ZHI L Z, et al. DNA barcoding for identification of fish species in the Taiwan Strait[J]. PLoS One, 2018, 13(6): e0198109.
- [10] KOWALCZYK M, STANISZEWSKI A, KAMIŃSKA K, et al. Advantages, possibilities, and limitations of mitochondrial DNA analysis in molecular identification[J]. Folia Biologica (Kraków), 2021, 69(3): 101-111.
- [11] SCHROETER J C, MALOY A P, REES C B, et al. Fish

- mitochondrial genome sequencing: expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA[J]. Conservation Genetics Resources, 2020, 12: 433-446.
- [12] 姚东林, 李正光, 王超等. 华南地区 6 个鲤亚科鱼类种群的氨基酸价值[J]. 西北农业学报, 2017, 26(7): 983-989.
- YAO Donglin, LI Zhengguang, WANG Chao, et al. Amino acid analysis of six kinds of genes Cyprininae fish population in southern China[J]. Acta Agriculturae Boreali-Occidentalis Sinica, 2017, 26(7): 983-989.
- [13] 王幼槐. 中国鲤亚科鱼类的分类、分布、起源及演化[J]. 水生生物学集刊, 1979, 6(4): 424.
- WANG Youhuai. On the classification, distribution, origin and evolution of the fishes referred to the subfamily cyprininae of China, with description of a new species[J]. Acta Hydrobiologica Sinica, 1979, 6(4): 424.
- [14] 申志新, 李高俊, 蔡杏伟等. 海南省淡水野生鱼类多样性演变及保护建议[J]. 中国水产, 2018, 516(11): 56-60.
- SHEN Zhixin, LI Gaojun, CAI Xingwei, et al. The evolution and protection of freshwater fish species in Hainan Province[J]. China Fisheries, 2018, 516(11): 56-60.
- [15] 姚东林, 王超, 吕子君等. 尖鳍鲤的肌肉基本营养成分和脂肪酸组成[J]. 广东农业科学, 2014, 41(19): 131-134.
- YAO Donglin, WANG Chao, LV Zijun, et al. Study on basic nutrition and fatty acid composition of muscle in *Cyprinus acutidorsalis* Wang[J]. Guangdong Agricultural Sciences, 2014, 41(19): 131-134.
- [16] 易祖盛, 王春, 陈湘彝. 尖鳍鲤的早期发育[J]. 中国水产科学, 2002, 2: 120-124+198-199.
- YI Zusheng, WANG Chun, CHEN Xianglin. Embryonic and larval development of sea carp, *Cyprinus acutidorsalis*[J]. Journal of Fishery Sciences of China, 2002, 2: 120-124+198-199.
- [17] 姚东林, 王超, 周爱国, 等. 基于 *Cyt b* 标记研究尖鳍鲤在鲤形目鱼类中的系统发育地位[J]. 广东农业科学, 2014, 41(12): 114-118.
- YAO Donglin, WANG Chao, ZHOU Aiguo, et al. Phylogenetic analysis of *Cyprinus acutidorsalis* Wang in Cypriniformes based on the *Cyt b* marker[J]. Guangdong Agricultural Sciences, 2014, 41(12): 114-118.
- [18] WANG C, CHEN J T, YAO D L, et al. The complete mitochondrial genome of the sea carp, *Cyprinus acutidorsalis* (Cypriniformes: Cyprinidae)[J]. Mitochondrial DNA, 2015, 26(5): 686-687.
- [19] JACOBSEN M W, DA FONSECA R R, BERNATCHEZ L, et al. Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the *NADH2* gene in whitefish (*Coregonus* spp.)[J]. Molecular Phylogenetics and Evolution, 2016, 95: 161-170.
- [20] 陈涛, 史妍茹, 尤平. 麦穗鱼线粒体基因组序列测定及分析[J]. 动物分类学报, 2012, 37(1): 10-19.
- CHEN Tao, SHI Yanru, YOU Ping. Sequence and analysis of complete mitochondrial genome of *pseudorasbora parva*[J]. Acta Zootaxonomica Sinica, 2012, 37(1): 10-19.
- [21] 程琪. 南方白甲鱼线粒体基因组序列测定分析及白甲鱼属系统发育研究[D]. 武汉: 中南民族大学, 2016.
- CHENG Qi. Complete mitochondrial genome sequencing and analysis of *Onychostoma gerlachi* and phylogenetic research of genus *onychostoma*[D]. Wuhan: South-central University For Nationalities, 2016.
- [22] 孙志鹏, 吕伟华, 匡友谊, 等. 玻璃缺鳍鲶线粒体全基因组序列测定与分析[J]. 水产学杂志, 2018, 31(2): 12-19.
- SUN Zhipeng, LV Huawei, KUANG Youyi, et al. Sequencing and analysis of complete mitochondrial genome of catfish *Kryptopterus vitreolus*[J]. Chinese Journal of Fisheries, 2018, 31(2): 12-19.
- [23] 钟东, 赵贵军, 张振书, 等. 基因组内碱基分布整体均衡与局部不均衡的研究进展[J]. 遗传, 2002, 3: 351-355.
- ZHONG Dong, ZHAO Guijun, ZHANG Zhenshu, et al. Advance in the entire balance and local unbalance of base distribution in genome[J]. Hereditas(Beijing), 2002, 3: 351-355.
- [24] KAMARAJ B, KUMAR A, PUROHIT R. Evolutionary reconstruction and population genetics analysis of aurora kinases[J]. Plos One, 2013, 8(9): e75763.
- [25] WEI L, HE J, JIA X, et al. Analysis of codon usage bias of mitochondrial genome in *Bombyx mori* and its relation to evolution[J]. BMC Evolutionary Biology, 2014, 14: 1-12.
- [26] 王恒. 三种鲷属鱼类线粒体全基因组的测序与比较[D]. 合肥: 安徽农业大学, 2015.
- WANG Heng. Sequence and comparison of the mtDNA of three *Sarcocheilichthys* fishes[D]. Hefei: Anhui Agricultural University, 2015.
- [27] BROWN W M, PRAGER E M, WANG A, et al. Mitochondrial DNA sequences of primates: tempo and mode of evolution[J]. Journal of Molecular Evolution, 1982, 18(4): 225-239.
- [28] MONTROYA J, OJALA D, ATTARDI G. Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs[J]. Nature, 1981, 290(5806): 465-470.
- [29] DUCHAUD E, ROCHAT T, HABIB C, et al. Genomic diversity and evolution of the fish pathogen *Flavobacterium psychrophilum*[J]. Frontiers in Microbiology, 2018, 9: 138.
- [30] 王开杰, 徐永江, 崔爱君, 等. 高体(鱼师)(*Seriola dumerili*)线粒体全基因组测定及结构特征分析[J]. 海洋与湖沼, 2022, 53(1): 120-132.

- WANG Kaijie, XU Yongjiang, CUI Aijun, et al. Complete sequence and gene organization of the mitochondrial genome of *seriola dumerili*[J]. *Oceanologia et Limnologia Sinica*, 2022, 53(1): 120-132.
- [31] 韩耀全, 黄励, 朱琼华, 等. 基于 Illumina Hiseq 4000 高通量测序平台的暗色唇鲷线粒体基因组特征分析[J]. *福建农业学报*, 2020, 35(2): 130-139.
- HAN Yaoquan, HUANG Li, ZHU Qionghua, et al. Characteristics of mitochondrial genomes of *Semilabeo obscures* analyzed using Illumina Hiseq 4000 High-throughput sequencing technique[J]. *Fujian Journal of Agricultural Sciences*, 2020, 35(2): 130-139.
- [32] 刘凯, 冯晓宇, 吴燕琴, 等. 瓯江高体鳊线粒体基因组测序及结构特征分析[J]. *农业生物技术学报*, 2021, 29(6): 1182-1197.
- LIU Kai, FENG Xiaoyu, WU Yanqin, et al. Sequencing and structural characteristics analysis of mitochondrial genome in *Rhodeus ocellatus* from Ou Riverv[J]. *Journal of Agricultural Biotechnology*, 2021, 29(6): 1182-1197.
- [33] 孙利元, 杨天燕, 孟玮, 等. 8 种石首鱼类线粒体基因组特征及分子系统进化分析[J]. *海洋科学*, 2017, 41(3): 48-54.
- SUN Liyuan, YANG Tianyan, MENG Wei, et al. Analysis of the mitochondrial genome characteristics and phylogenetic relationships of eight sciaenid fishes[J]. *Marine Sciences*, 2017, 41(3): 48-54.
- [34] PALANDAČIĆ A, NASEKA A, RAMLER D, et al. Contrasting morphology with molecular data: an approach to revision of species complexes based on the example of European Phoxinus (Cyprinidae)[J]. *BMC Evolutionary Biology*, 2017, 17(1): 1-17.
- [35] HONTELA A, STACEY N. *Cyprinidae*[M]. Boca Raton: CRC Press, 2019: 53-78.

Mitochondrial genome sequencing and structural characteristic analysis of *Cyprinus acutidorsalis* from Wanquan Estuary, Hainan

MENG Yu¹, SONG Yi-qing^{1, 2}, XIE Qing-ying¹, CAI Xing-wei^{3, 4}, ZHANG Qing-feng^{3, 4}, DONG Yang^{3, 4}, LI Fang-yuan^{3, 4}, XIE Song-guang^{1, 2}, SHEN Zhi-xin^{3, 4}

(1. College of Marine Biology and Fisheries in Hainan University, Haikou 570228, China; 2.State Key Laboratory of Marine Resources Utilization in South China Sea, Haikou 570228, China; 3.Hainan Academy of Ocean and Fisheries Sciences, Haikou 571126, China; 4. Hainan Freshwater Biological Resources and Ecological Environment Protection Research Center, Haikou 571226, China)

Received: Jul, 20, 2023

Key words: *Cyprinus acutidorsalis*; mitochondrial genome; phylogeny; sequence analysis

Abstract: In this study, the complete mitochondrial genome sequence of *Cyprinus acutidorsalis* from the Wanquan River Estuary of Hainan Province is obtained using second-generation sequencing. Their structural characteristics are also analyzed. The results show a typical circular closed-loop double-stranded molecule with a total length of 16 581 bp. The base composition is A (31.96%), G (15.69%), C (27.53%), and T (24.82%), which contains 13 protein-coding genes (PCGs), two rRNA genes, 22 tRNA genes, and one variable control region D ring. The base content analysis reveals that the base composition of the carp has a high AT content bias. Furthermore, the 13 PCGs have many favorable codons, such as CGA (RSCU > 3.506) and CCA (RSCU > 2.208). Except for the *tRNA^{Gln}*, *tRNA^{Ala}*, *tRNA^{Asn}*, *tRNA^{Cys}*, *tRNA^{Leu}* (*CUN*), *tRNA^{Tyr}*, *tRNA^{Ser}* (*UCN*), *tRNA^{Pro}*, and *ND6*, the remaining genes are arranged on the H chain. The mitochondrial genome comparison between Hainan and Guangxi depicts that the similarity of the PCGs between them is between 96% and 100%. *ND4*, *CO II*, and *Cytb* differ in their starting positions, lengths, and starting codons in varying degrees. In the pound-to-pair comparison of 13 PCGs from *C. carpio* and *Carassius auratus*, which belong to the same genus, the similarity of the PCGs among the three species of *Cyprinus* fish is between 85% and 99%. Constructing phylogenetic trees with the mitochondrial genome sequences of other cyprinids demonstrates the close relationship between *C. acutidorsalis* and *C. carpio*. The results will help us understand the relatives of different Cyprinidae and provide a reference for mitochondrial evolution analysis and germplasm resource research of *C. acutidorsalis*.

(本文编辑: 谭雪静)