



黄带拟鲈线粒体基因组测序及鲈科鱼类系统发育分析

王开杰^{1,2}, 姜 燕¹, 徐永江^{1*}, 柳学周¹, 崔爱君¹,
王 滨¹, 薛志勇³, 毛成全³

(1. 中国水产科学研究院黄海水产研究所,

青岛海洋科学与技术试点国家实验室深蓝渔业工程联合实验室, 山东 青岛 266071;

2. 浙江海洋大学, 国家海洋设施养殖工程技术研究中心, 浙江 舟山 316022;

3. 海阳市黄海水产有限公司, 山东 烟台 265122)

摘要: 为深入开展黄带拟鲈种质鉴定、分类及遗传进化等研究, 通过二代高通量测序与生物信息学分析, 揭示了黄带拟鲈线粒体基因组全序列基因结构及鲈科鱼类系统发育特征。结果显示, 黄带拟鲈线粒体基因组为典型的环状 DNA 结构, 序列全长为 16 570 bp, 碱基组成分别为 A (27.2%)、G (17.18%)、C (30.24%) 和 T (25.38%), 包括 13 种蛋白编码基因, 22 个 tRNA 基因, 2 个 rRNA 基因, 除 *ND6*、tRNA^{Gln}、tRNA^{Ala}、tRNA^{Asn}、tRNA^{Cys}、tRNA^{Tyr}、tRNA^{Ser}、tRNA^{Glu}、tRNA^{Pro} 外, 其余基因均在 H 链上编码, 且基因组存在多处重叠区域。除 *CO I* 和 *ND5* 的起始密码子分别为 ATC 和 ATA 外, 其余 11 个蛋白编码基因的起始密码子均为 ATG, 以典型的 TAA 和 TAG 为终止密码子, 在 *Cyt b* 中存在不完全终止密码子 T。黄带拟鲈线粒体基因组全序列与蛋白编码基因的 A+T 含量分别为 52.58% 和 51.55%, 非编码控制区 (D-loop) A+T 富含 61.69%, 具有明显的 AT 偏好性。除 tRNA^{Ser} (GCT) 外, 其余 21 个 tRNA 均含典型三叶草二级结构。为进一步研究黄带拟鲈系统进化特性, 通过与 18 种鲈科鱼类线粒体基因组全序列及 16S *rRNA* 基因构建系统进化树显示, 黄带拟鲈与高体若鲈同属一个分支, 亲缘关系最近。本研究结果可为黄带拟鲈物种鉴别、遗传多样性评价与保护提供技术依据。

关键词: 黄带拟鲈; 线粒体基因组; 序列分析; 系统发育

中图分类号: Q 785; S 917.4

文献标志码: A

自 1962 年, 在电子显微镜下观察到线粒体环状结构后^[1], 线粒体基因组 DNA 的结构、特征和性质等方面的研究逐步深入。与核基因相比, 线粒体 DNA 结构较为简单, 专一性和独特性非常高^[2], 且具有长度短、含量丰富、偏母性遗传和进化速率快等特点^[3]。在大西洋鳕 (*Gadus morhua*) 线粒体 DNA 基因组序列首次被测定公布后, 线粒

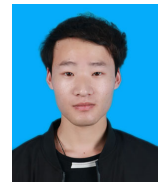
体 DNA 在鱼类群体遗传学、分子系统学和生物地理学等方面的研究广泛开展^[4-7], 极大地促进了生物多样性和遗传进化分析研究的发展。线粒体基因组作为分子标记, 已成为研究物种系统发育进化谱系不可缺少的重要工具, 并取得了诸多科研成果^[8]。在渔业研究领域, 通过比较和分析鱼类线粒体全基因组信息, 为鱼类的系统发育进化、

收稿日期: 2021-07-07 修回日期: 2021-10-24

资助项目: 国家重点研发计划 (2019YFD0900901, 2018YFD0901204); 中国水产科学研究院基本科研业务费专项 (2020TD47); 中国水产科学研究院黄海水产研究所基本科研业务费专项 (20603022021011); 财政部和农业农村部国家现代农业产业技术体系专项 (CARS-47)

第一作者: 王开杰 (照片), 从事鱼类多样性评价与保护技术研究, E-mail: jasons@qq.com

通信作者: 徐永江, 从事海水鱼类繁育理论与技术研究, E-mail: xuyj@ysfri.ac.cn



生物多样性等研究提供了重要的技术手段。如葛倩芸等^[9]通过 PCR 扩增和测序, 获得了鲤 (*Cyprinus carpio*) 和鲫 (*Carassius auratus*) 线粒体基因组序列, 揭示了其遗传变异特征及基因分布规律。李瑶瑶等^[10]通过线粒体全基因组序列实现了鲑科 (*Salmonidae*) 鱼类的物种分类和系统进化分析。

黄带拟鲈 (*Pseudocaranx dentex*) 属鲈形目 (*Perciformes*) 鲈科 (*Carangidae*) 拟鲈属 (*Pseudocaranx*), 又称长缟鲈、岛鲈、纵带鲈、皇帝鲈、大竹筴鱼, 体呈流线型, 因身体两侧中间部位各有一条亮黄色的条带而得名^[11]。黄带拟鲈广泛分布于印度-太平洋及大西洋的温暖水域, 中国主要分布在南海和东海^[3], 非常适宜工厂化循环水、大型深水网箱等新养殖模式, 是适宜我国深远海养殖的优良鱼种。黄带拟鲈肉质鲜嫩、营养丰富, 是制作刺身的上等原料^[11-12], 具有较高的经济价值, 市场前景广阔。近年来, 我国部分水产企业开始引进日本的黄带拟鲈进行养殖实验, 取得了较好的养殖效果。我国黄带拟鲈的人工繁殖研究也已开始, 为推进该鱼种人工繁育与养殖技术研究, 有必要全面认识其遗传背景。目前, 国际上相关报道多为营养品质和养殖相关^[13-14], 我国关于黄带拟鲈种质遗传方面的研究仅见染色体核型特征方面的报道^[15], 有关黄带拟鲈形态种质特征、线粒体基因组等种质遗传特性方面的研究尚未见相关报道。

本研究测定并分析了黄带拟鲈的线粒体全基因组序列及其基因结构特征, 构建鲈科鱼类的分子系统发育树, 探讨了黄带拟鲈与其他鲈科鱼类的亲缘关系, 为黄带拟鲈的种质鉴定、系统发育进化分析以及遗传多样性评价与利用提供依据。

1 材料与方法

1.1 实验材料

实验鱼为大连富谷食品有限公司 2020 年室内工厂化循环水养殖的黄带拟鲈, 为 2019 年自日本德岛县生命科学株式会社运输至公司进行养殖的苗种 (体质量 80~100 g)。从养殖车间内随机捞取 1 尾黄带拟鲈, 体长 26.8 cm, 体质量 490 g。

1.2 样品采集及 DNA 提取

采用 MS-222 (180 mg/L) 麻醉后, 剪取黄带拟鲈胸鳍保存于无水乙醇中。取胸鳍组织 (50~100 mg),

按照 DNA 提取试剂盒 (OMEGA, 美国) 使用说明书的操作步骤提取 DNA。通过琼脂糖凝胶电泳和核酸蛋白测定仪对 DNA 质量进行检测。

1.3 基因测序

将检测合格的 DNA 样品参照试剂盒 (TruSeq™ Nano DNA Sample Prep Kit) 方法构建文库, 用超声波破碎仪 (Covaris M220) 分割成长度为 300~500 bp 的片段, 之后在补平 3'端加 A、连接索引序列 (index) 接头 (TruSeq™ Nano DNA Sample Prep Kit) 进行 PCR 扩增使文库富集, 再用 2% 琼脂糖胶回收文库 (Certified Low Range Ultra Agarose) 后按数据比例混合上机, 在 cBot 固相载体上进行桥式 PCR 扩增生成集群 (clusters), 最后通过 Illumina NovaSeq 测序平台 (青岛欧易生物科技有限公司) 进行双末端测序, 测序深度为 200×。

1.4 序列拼接与分析

由于 Illumina 的原始测序数据会存在一些质量比较低的数据, 为了使后续的组装更加准确, 通过 Trimmomatic v0.39 (<http://www.usadellab.org/cms/index.php?page=trimmomatic>) 软件进行质量剪切 [具体步骤: 去除读长 (reads) 的接头序列 (adapter); 剪切前去除 5'端含有非 AGCT 的碱基; 修剪测序质量较低的 reads 末端 (测序质量值小于 Q20); 去除含 N 的比例达到 10% 的 reads; 舍弃去 adapter 及质量修剪后长度小于 75 bp 的小片段]。将修剪后的片段使用 SPAdes v3.14.1 (<http://bioinf.spbau.ru/spades>) 软件对其进行校对、排列、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, 并根据重叠部分 (overlap) 连接序列, 得到黄带拟鲈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止密码子和终止密码子的位置, 获得高准确性的保守基因集。通过 tRNA scan-SE (<http://www.Genetics.Wustl.Edu/eddy/tRNAscan-SE>) 软件进行 tRNA 定位, RNA 结构 (RNAstructure) 在线预测, 利用软件 CGView (http://stothard.afns.ualberta.ca/cgview_server/) 对样品基因组进行圈图展示, 最后将蛋白编码基因通过 Blast 比对, 进行 NR (NCBI non-redundant protein sequences, 非冗余蛋白序列数据库)、Swiss-Prot (A manually annotated and reviewed protein database, 蛋白质序

列数据库)、eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups, 直系同源蛋白分组比对数据库)、KEGG (Kyoto Encyclopedia of Genes and Genomes, 京都基因和基因组百科全书) 和 GO (Gene Ontology, 基因本体论数据库) 注释分析。

1.5 鲮科鱼类线粒体全基因组系统进化

从 GenBank 中下载 18 种鲮科鱼类线粒体基因组序列 (表 1), 使用 Clustal W 软件进行序列多重比对, 采用 MEGA 7.0 软件中的邻接法 (Neighbor-joining, NJ) 构建系统发育树。

表 1 鲮科鱼类线粒体基因组数据来源

Tab. 1 Origins of mitochondrial genomes of Carangidae fishes

属 genus	物种 species	GenBank 登录号 GenBank accession no.	全长/bp total length
细鲮属 <i>Selaroides</i>	金带细鲮 <i>S. leptolepis</i>	KM522839	16 560
叶鲮属 <i>Atule</i>	游鳍叶鲮 <i>A. mate</i>	KM522838	16 565
副叶鲮属 <i>Alepes</i>	丽叶鲮 <i>A. kleinii</i>	KF728081	16 571
大甲鲮属 <i>Megalaspis</i>	大甲鲮 <i>M. cordyl</i>	KM522836	16 566
若鲮属 <i>Carangoides</i>	甲若鲮 <i>C. armatus</i>	AP004444	16 556
	高体若鲮 <i>C. equula</i>	KM201334	16 588
	马拉巴若鲮 <i>C. malabaricus</i>	KJ174514	16 561
乌鲮属 <i>Parastromateus</i>	乌鲮 <i>P. niger</i>	KJ192332	16 561
竹筴鱼属 <i>Trachurus</i>	竹筴鱼 <i>T. trachurus</i>	AB108498	16 559
尾甲鲮属 <i>Uraspis</i>	白舌尾甲鲮 <i>U. secunda</i>	KT819204	16 554
圆鲮属 <i>Decapterus</i>	长身圆鲮 <i>D. macrosoma</i>	KF841444	16 545
凹肩鲮属 <i>Selar</i>	脂眼凹肩鲮 <i>S. crumenophthalmus</i>	KJ148633	16 610
小条鲮属 <i>Seriolina</i>	小甘鲮 <i>S. nigrofasciata</i>	KT591876	16 531
	日本黄条鲮 <i>S. lalandi</i>	AB517557	16 532
	中国黄条鲮 <i>S. lalandi</i>	MH211123	16 609
	五条鲮 <i>S. quinquerediata</i>	AB517556	16 537
鲮属 <i>Seriola</i>	高体鲮 <i>S. dumerili</i>	AB517558	16 530
	长鳍鲮 <i>S. rivoliana</i>	KP347126	16 530
	纺缍鲮 <i>E. bipinnulata</i>	KT824759	16 542

2 结果

2.1 基因结构组成分析

黄带拟鲮线粒体基因组全序列为典型的双链闭合环状结构 (图 1), 全长为 16 570 bp, 碱基组成分别为 A (27.2%)、G (17.18%)、C (30.24%) 和 T (25.38%); G+C 的含量为 47.42%, 低于 A+T 含量 (52.58%), 表明黄带拟鲮线粒体基因组具有 AT 偏倚性。线粒体基因组序列经注释后提交 GenBank (accession no. MZ398237)。黄带拟鲮与大多数鱼类的线粒体 DNA 结构一样, 共 37 个基因, 相邻基因间存在不同程度的碱基间隔或重叠现象 (表 2)。包括 13 个蛋白质编码基因 (*CO I~III*、*ATP6*、*ATP8*、*ND1~6*、*ND4L* 和 *Cytb*), 22 个

tRNA 基因, 2 个 rRNA 基因 (12S *rRNA* 和 16S *rRNA*) 和 1 个非编码的控制区 (D-loop)。

2.2 黄带拟鲮线粒体 rRNA 与 tRNA 基因结构特征

黄带拟鲮线粒体基因组包含 2 个 rRNA, 均位于 H 链上, 12S *rRNA* 位于 tRNA^{Phe} 和 tRNA^{Val} 之间, 由 951 个碱基组成; 16S *rRNA* 位于 tRNA^{Val} 和 tRNA^{Leu} 之间, 由 1 698 个碱基组成, 且 12S *rRNA* 比 16S *rRNA* 更加保守。

黄带拟鲮线粒体基因组共有 22 个 tRNA, 序列长度 67~76 bp, 长度最短的为 tRNA^{Trp}, 最长的为 tRNA^{Ile}, 总长度为 1 561 bp (表 2)。其中各有 2 个 tRNA^{Ser} 和 tRNA^{Leu}, tRNA^{Ser} 的反密码子分别为

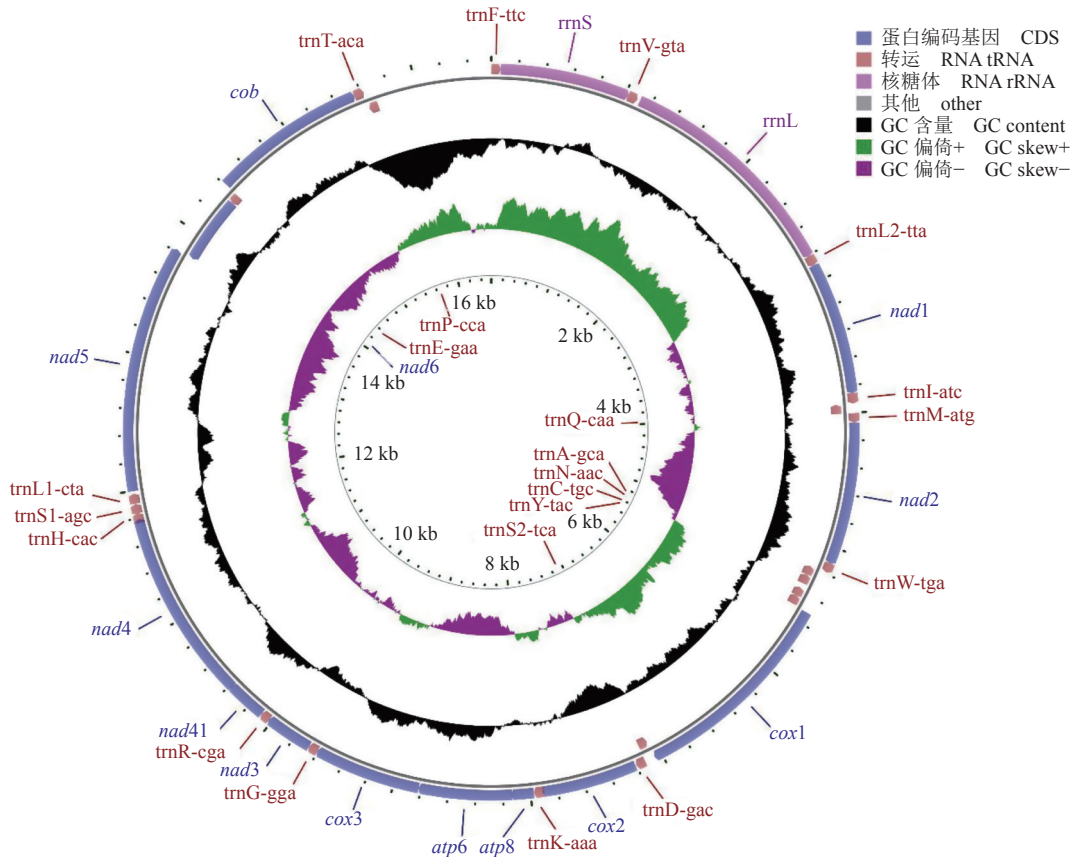


图 1 黄带拟鲮线粒体基因组图谱

Fig. 1 Mitochondrial genome mapping of *P. dentex*

TGA 和 GCT, tRNA^{Leu} 的反密码子分别为 TAA 和 TAG。除了位于 L 链上的 8 个 tRNA (tRNA^{Gln}、tRNA^{Ala}、tRNA^{Asn}、tRNA^{Cys}、tRNA^{Tyr}、tRNA^{Ser}、tRNA^{Glu}、tRNA^{Pro}) 外, 其余 14 个皆位于 H 链上。此外, tRNA^{Ser}-GCT 基因缺失二氢尿嘧啶臂 (DHU 环), 其余 21 个 tRNA 均含氨基酸接受臂、TψC 环、反密码子环和 DHU 环的典型三叶草二级结构。

在 tRNA^{Val}、tRNA^{Leu}、tRNA^{Ile}、tRNA^{Trp}、tRNA^{Arg} 和 tRNA^{His} 氨基酸臂中各有 1 对因 C-T 转换造成的 A-C 不配对, 在 tRNA^{Gln}、tRNA^{Met}、tRNA^{Ala}、tRNA^{Cys}、tRNA^{Glu} 和 tRNA^{Pro} 氨基酸臂上皆有因 C-T 转换造成的 G-T 错配。同时, 在 tRNA^{Ser}(GCT) 氨基酸臂上有 1 对 A-A 不配对, 在 tRNA^{Leu}(TAG) 和 tRNA^{Thr} 上各有 1 对 C-C 不配对。在 tRNA^{Phe} 氨基酸臂中由于颠换导致的 A-G 不配对。在反密码子茎中, 由于 C-T 转换造成的 tRNA^{Trp} 中的 A-C 不配对。在 TψC 茎上, 同样存在较多的 C-T 转换现象, 除此之外, tRNA^{Met} 和 tRNA^{Glu} 各有 1 对 U-U 不配对。DHU 环也存在 G-T 错配现象, 且 tRNA^{Ser}-(GCT) 基因缺失该环。

2.3 蛋白编码基因

黄带拟鲮线粒体基因组共含有 13 个蛋白编码基因, CO I 和 ND5 基因的起始密码子分别为 ATC 和 ATA, 其他蛋白质编码基因起始密码子均为 ATG。Cyt b 的终止密码子为不完全的“T”, ND2、ATP8 和 ND3、ND4 的终止密码子为“TAG”, 剩余的 7 种基因为完全的终止密码子“TAA”。

黄带拟鲮线粒体基因组的蛋白编码基因中除 ND6 基因由 L 链编码外, 其余 12 个蛋白编码基因均位于 H 链上。这 13 个蛋白编码基因共编码 75 个碱性氨基酸、178 个酸性氨基酸、648 个极性氨基酸和 1 218 个非极性氨基酸, 编码基因的总长度为 11 449 bp, 平均长度约为 881 bp, 占基因组总长度的 69.09%; 并且, 基因组中存在多处重叠现象。其中, ATP8 与 ATP6 序列有 7 bp 碱基的重叠, ATP6 与 CO III 有 1 bp 碱基的重叠, ND4L 与 ND4 重叠了 7 bp, ND5 与 ND6 有 4 bp 的重叠。

在这 13 条蛋白编码基因中, 除了 ND1、ND2 和 ND4L 外, 其他 10 条蛋白编码基因的 A+T 含量均高于 50% (表 3), 且 13 个蛋白编码基

表 2 黄带拟鲮线粒体基因组结构特点

Tab. 2 Characteristics of the mitochondrial genome of *P. dentex*

基因名称 gene name	起止位置 beginning and end	长度/bp length	编码链 strand	间隔长度/bp intergenic space	起始密码子 initiation codon	终止密码子 termination codon
tRNA ^{Phe} -GAA	1~68	68	H	—	—	—
12S <i>rRNA</i>	69~1 019	951	H	0	—	—
tRNA ^{Val} -UAC	1 020~1 091	72	H	0	—	—
16S <i>rRNA</i>	1 113~2 810	1 698	H	21	—	—
tRNA ^{Leu} -UAA	2 811~2 884	74	H	0	—	—
<i>ND1</i>	2 885~3 859	975	H	0	ATG	TAA
tRNA ^{Ile} -GAU	3 859~3 934	76	H	-1	—	—
tRNA ^{Gln} -UUG	3 934~4 004	71	L	-1	—	—
tRNA ^{Met} -CAU	4 004~4 072	69	H	-1	—	—
<i>ND2</i>	4 073~5 119	1 047	H	0	ATG	TAG
tRNA ^{Trp} -UCA	5 118~5 188	71	H	-2	—	—
tRNA ^{Ala} -UGC	5 190~5 258	69	L	1	—	—
tRNA ^{Asn} -GUU	5 260~5 332	73	L	1	—	—
tRNA ^{Cys} -GCA	5 369~5 435	67	L	36	—	—
tRNA ^{Tyr} -GUA	5 436~5 505	70	L	0	—	—
<i>CO I</i>	5 513~7 057	1 545	H	7	ATC	TAA
tRNA ^{Ser} -UGA	7 058~7 128	71	L	0	—	—
tRNA ^{Asp} -GUC	7 132~7 202	71	H	3	—	—
<i>CO II</i>	7 210~7 914	705	H	7	ATG	TAA
tRNA ^{Lys} -UUU	7 901~7 975	75	H	-14	—	—
<i>ATP8</i>	7 977~8 141	165	H	1	ATG	TAG
<i>ATP6</i>	8 135~8 818	684	H	-7	ATG	TAA
<i>CO III</i>	8 818~9 603	786	H	-1	ATG	TAA
tRNA ^{Lys} -UCC	9 603~9 672	70	H	-1	—	—
<i>ND3</i>	9 673~10 023	351	H	0	ATG	TAG
tRNA ^{Arg} -UCG	10 022~10 090	69	H	-2	—	—
<i>ND4L</i>	10 092~10 388	297	H	1	ATG	TAA
<i>ND4</i>	10 382~11 791	1 410	H	-7	ATG	TAG
tRNA ^{His} -GUG	11 763~11 834	72	H	-29	—	—
tRNA ^{Ser} -GCU	11 835~11 902	68	H	0	—	—
tRNA ^{Leu} -UAG	11 907~11 979	73	H	4	—	—
<i>ND5</i>	11 998~13 818	1 821	H	18	ATA	TAA
<i>ND6</i>	13 815~14 336	522	L	-4	ATG	TAA
tRNA ^{Glu} -UUC	14 337~14 405	69	L	0	—	—
<i>Cyt b</i>	14 410~15 550	1 141	H	4	ATG	T--
tRNA ^{Thr} -UGU	15 551~15 622	72	H	0	—	—
tRNA ^{Pro} -UGG	15 622~15 692	71	L	-1	—	—
D-loop	15 693~16 570	878	H	0	—	—

注: “—”表示无

Notes: “—” represents none

因的 A+T 总含量为 51.55%，可见其在蛋白编码基因组中也具有 AT 偏好性。由于黄带拟鲮线粒体

基因组全序列的 A+T 含量为 52.58%，可推测其非蛋白编码区基因组也具有 AT 偏好性。

表 3 黄带拟鲮蛋白编码基因碱基组成

Tab. 3 Nucleotide composition of protein-coding genes in *P. dentex*

基因 gene	碱基/bp nucleotide	T	C	A	G	A+T
ND1	975	26.05	34.56	23.08	16.31	49.13
ND2	1 047	23.50	36.96	25.50	14.04	49.00
CO I	1 545	29.84	26.73	24.79	18.64	54.63
CO II	705	26.52	28.65	28.79	16.03	55.31
ATP8	165	25.45	33.33	26.06	15.15	51.51
ATP6	684	27.78	34.50	22.95	14.77	50.73
CO III	786	26.46	31.55	23.54	18.45	50.00
ND3	351	28.21	32.19	24.50	15.10	52.71
ND4L	297	25.25	36.03	21.21	17.51	46.46
ND4	1 410	26.03	33.26	25.11	15.60	51.14
ND5	1 821	26.63	32.62	26.47	14.28	53.1
ND6	522	35.25	17.24	14.75	32.76	50.00
Cyt b	1 141	27.70	32.87	23.05	16.39	50.75

2.4 非编码区的结构特征

非编码区主要包括重链复制起始的非编码控制区 (D-loop) 和轻链复制起始区 (O_L)。 O_L 是一个位于 tRNA^{Asn} 和 tRNA^{Cys} 之间的约 36 bp 的 DNA 片段。D-loop 是一个大小约 878 bp 的片段，其碱基含量为 (T=28.85%，C=23.49%，A=32.84%，G=14.82%)，该区域 A+T 为 61.69%，表现明显的 A+T 偏好性。

2.5 鲮科鱼类线粒体基因组系统发育

将黄带拟鲮线粒体基因组全序列与其他 18 种分属 13 个属的鲮科鱼类线粒体基因组构建分子系统发育树，此外在鱼类中，线粒体 16S *rRNA* 序列也被广泛用于不同分类阶元的系统发育研究，以便进一步探讨黄带拟鲮及鲮科鱼类的系统分化和遗传背景。

线粒体基因组全序列与 16S *rRNA* 构建的进化树结果一致，显示该进化树共由 2 大支构成，分别为由大甲鲮属、副叶鲮属、叶鲮属、若鲮属、细鲮属、拟鲮属、凹肩鲮属、竹筴鱼属、乌鲮属、尾甲鲮属、圆鲮属等构成的鲮科鱼类进化主体分支，以及由纺锤鲮属、鲮属和小条鲮属组成的另一进化分支。黄带拟鲮与高体若鲮在进化树中同属一个分支，亲缘关系最近，与其聚为另一支的竹

筴鱼和长身圆鲮相比分化时间较晚 (图 2, 图 3)。

3 讨论

随着 DNA 测序技术的日渐成熟和生物信息学的快速发展，鱼类线粒体基因组的研究得以广泛应用于鱼类种质资源保护、物种鉴定、群体多态性及系统进化发育研究等领域^[16]。已有研究表明，鱼类线粒体基因组一般为 15~20 kb，多为双链闭合的环状结构，排列紧密、分子量小，不同物种间线粒体基因组差异很大，序列中存在串联重复序列以及碱基的插入和缺失^[17]。本研究测定了黄带拟鲮的线粒体基因组全序列，其全长 16 570 bp，GC 含量约 47.42%，呈现出明显的 AT 偏好性和抗鸟嘌呤 (反 G) 现象。其中 G 碱基含量仅为 17.18%，与其他硬骨鱼类如黄条鲮 (*S. lalandi*) (17.84%)^[8]、三角鲂 (*Megalobrama terminalis*) (16.17%)^[18]、青鳞 (*Oryzias latipes*) (17.60%)^[19] 等鱼类线粒体中的碱基 G 含量相近；C 碱基含量为 30.24%，A 碱基含量为 27.2%，T 碱基含量为 25.38%，与大多数脊椎动物碱基组成相似^[20]。在黄带拟鲮线粒体基因组中，13 个蛋白编码基因除了 ND6 外，其余均在 H 链上，与其他鱼类类似。基因组中有许多的重叠区域，包括在蛋白编码基因中 4 个重叠区 (ATP8~ATP6、ATP6~CO III、

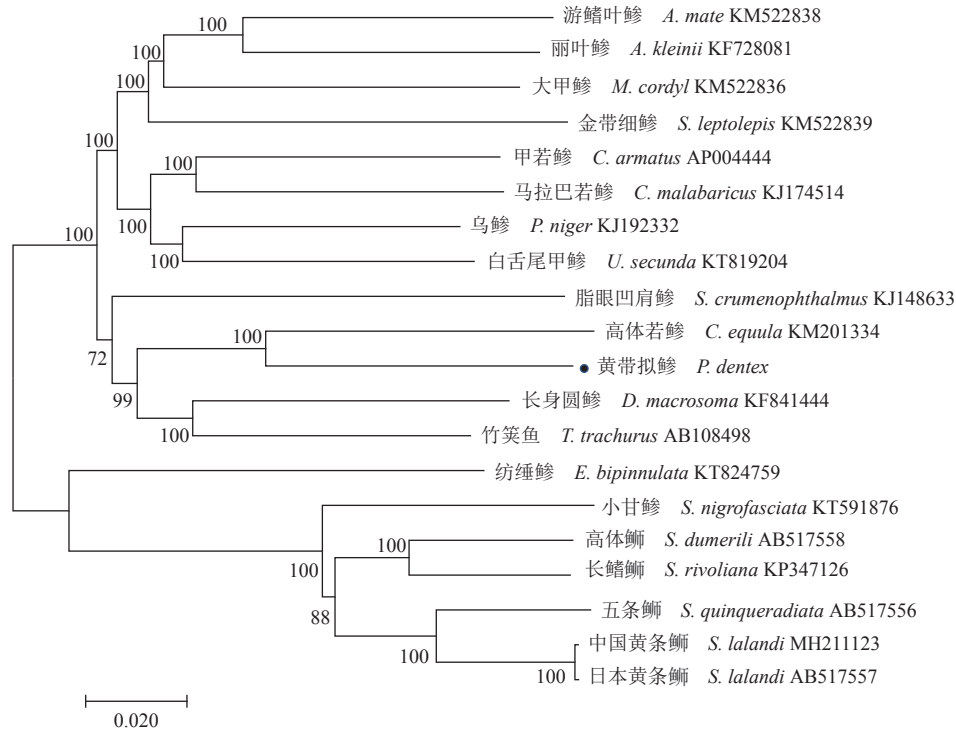


图 2 基于线粒体基因组全序列构建的 NJ 系统进化树

Fig. 2 Molecular phylogenetic tree constructed by NJ method of mitochondrial genome

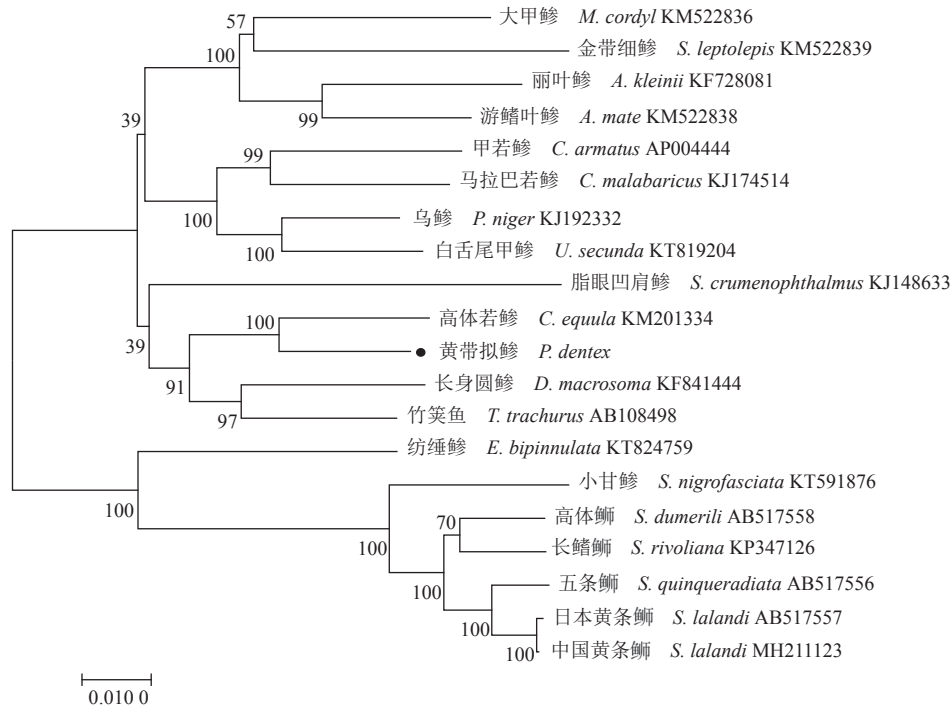


图 3 基于 16S rRNA 基因构建的 NJ 系统进化树

Fig. 3 Neighbor-joining phylogenetic tree based on 16S rRNA gene

ND4L~ND4 和 ND5~ND6), tRNA 之间有 tRNA^{Ile}-tRNA^{Gln}、tRNA^{Gln}-tRNA^{Met} 和 tRNA^{Thr}-tRNA^{Pro} 3 个重叠区, 以及蛋白编码基因与 tRNA 之间存在 6 个重叠区: ND1-tRNA^{Ile}、ND2-tRNA^{Trp}、CO II-

tRNA^{Lys}、CO III-tRNA^{Gly}、ND3-tRNA^{Arg} 和 ND4-tRNA^{His}。在 ATP 酶基因之间出现较多的重叠区, 鱼类中该区域重叠的片段一般只有 7~10 bp, 而在哺乳动物中重叠片段可达 40~46 bp^[21-22]。此外, 除

了 *Cyt b* 编码蛋白的终止密码子为不完全的终止密码子 T, 其余的 12 种编码蛋白都是以常见的 TAA 和 TAG 作为终止密码子。控制区是一个 878 bp 的片段, 该区域调控线粒体的复制和转录^[23]。在 tRNA 结构中, tRNA^{Ser}-GCT 基因缺失二氢尿嘧啶臂 (DHU 臂) 而不能形成三叶草结构, 这在条石鲷 (*Oplegnathus fasciatus*)^[24]、鲱科 (Sisoridae) 鱼类^[25] 中也有类似报道。虽然错配发生频率较高, 但依据 Yokobori 等^[26] 提出的“tRNA 二级结构中的错配现象可以在 RNA 编辑中校正”这一结论, 因此碱基错配并不影响 tRNA 基因的转运功能。

每个蛋白编码基因的进化速率不同, Zar-doya 等^[27] 把 13 个蛋白质的编码基因依次划分为好、中、差 3 个组, 其中, *CO I*、*ND2*、*ND4*、*Cyt b* 和 *ND5* 这 5 个基因为良好, *CO II*、*CO III*、*ND1* 和 *ND6* 为中等, 而 *ATP6*、*ATP8*、*ND3* 和 *ND4L* 为较差。绝大多数的蛋白编码基因进化速率介于控制区与核糖体 RNA 之间, 表现出中等的进化速率。黄带拟鲈进化速率较快的 *CO I*、*Cyt b* 和 *ND* 等基因可作为分子标记, 同其他鲈科鱼类进行鉴别, 为其种质资源保护提供参考。而线粒体基因组中的 16S *rRNA* 序列是非编码蛋白质基因, 不受密码子编码的选择压力影响, 其大部分区域发生的变异为中性突变, 加之 mtDNA 进化速率明显大于核 DNA, 因此采用线粒体 16S *rRNA* 序列同源性比较对研究物种间的系统进化关系具有一定的优势^[28]。在鲈科鱼类中, 多数学者认为全球共分为 4 个亚科: 鲈亚科 (Caranginae)、鲷鲈亚科 (Trachinotinae)、鲷亚科 (Seriolinae) 和 鲷鲈亚科 (Chorineminae), 研究发现鲈亚科与鲷亚科形成姐妹群, 接着与鲷鲈亚科聚为一支, 最后与鲷鲈亚科聚为一大支^[29-30]。本研究所选取鲈科鱼类线粒体基因组范围为 16 530~16 609 bp, 序列大小差异不大, 基于 K2-P 模型, 对鲈科鱼类的系统进化关系有了准确判别。

单个基因所包含的系统信息太少, 不能够反映整个生物分子进化水平, 而分析多个基因组编码的基因序列得到的结果更具说服力^[31]。在鱼类中, 线粒体 16S *rRNA* 序列也被广泛用于研究不同阶元的系统发育关系^[32]。为更好地确认黄带拟鲈系统发育关系, 将其与 18 种鲈科鱼类线粒体基因组及 16S *rRNA* 基因构建的进化树比较, 黄带拟鲈与高体若鲈、长身圆鲈、竹筴鱼同属一个分支, 且分化时间较晚, 与其他鲈科鱼类距离较远, 且

黄带拟鲈与高体若鲈的亲缘关系最近, 二者之间可能有相似的遗传背景。本研究为黄带拟鲈的种质鉴定、系统发育进化分析以及遗传多样性评价与利用提供了依据。

(作者声明本文无实际或潜在的利益冲突)

参考文献 (References):

- [1] Nass M M K, Nass N S. Fibrous structures within the matrix of developing chick embryo mitochondria[J]. *Experimental Cell Research*, 1962, 26(2): 424-427.
- [2] Curole J P, Kocher T D. Mitogenomics: digging deeper with complete mitochondrial genomes[J]. *Trends in Ecology & Evolution*, 1999, 14(10): 394-398.
- [3] 陈大刚, 张美昭. 中国海洋鱼类 [M]. 青岛: 中国海洋大学出版社, 2015: 1096-1097.
Chen D G, Zhang M Z. Marine fishes of China[M]. Qingdao: China Ocean University Press, 2015: 1096-1097 (in Chinese).
- [4] Liu Y, Cui Z X. The complete mitochondrial genome sequence of the cutlassfish *Trichiurus japonicus* (Perciformes: Trichiuridae): genome characterization and phylogenetic considerations[J]. *Marine Genomics*, 2009, 2(2): 133-142.
- [5] 柳淑芳, 吴仁协, 朱玲, 等. 小黄鱼 (*Larimichthys polyactis*) 线粒体基因组结构与特征 [J]. *海洋与湖沼*, 2010, 41(6): 885-894.
Liu S F, Wu R X, Zhu L, et al. Complete sequence and gene organization of mitochondrial DNA of the small yellow croaker *Larimichthys polyactis* [J]. *Oceanologia et Limnologia Sinica*, 2010, 41(6): 885-894 (in Chinese).
- [6] 赫崇波, 高祥刚, 王效敏, 等. 圆斑星鲈线粒体基因组全序列结构及其进化 [J]. *中国水产科学*, 2007, 14(4): 584-592.
He C B, Gao X G, Wang X M, et al. Structure and evolution of complete mitochondrial genome of spotted halibut *Verasper variegatus* [J]. *Journal of Fishery Sciences of China*, 2007, 14(4): 584-592 (in Chinese).
- [7] 毛明光, 顾杰, 刘瑞婷, 等. 太平洋鲑线粒体全基因组测序及结构特征分析 [J]. *水生生物学报*, 2019, 43(1): 17-26.
Mao M G, Gu J, Liu R T, et al. Analysis of complete mitochondrial genome sequences of *Gadus macrocephalus* [J]. *Acta Hydrobiologica Sinica*, 2019, 43(1): 17-26
中国水产学会主办 sponsored by China Society of Fisheries

- (in Chinese).
- [8] 史宝, 柳学周, 刘永山, 等. 黄条鲮线粒体全基因组测序及结构特征分析[J]. *中国水产科学*, 2019, 26(3): 405-415.
- Shi B, Liu X Z, Liu Y S, *et al.* Complete sequence and gene organization of the mitochondrial genome of *Seriola lalandi*[J]. *Journal of Fishery Sciences of China*, 2019, 26(3): 405-415 (in Chinese).
- [9] 葛倩芸, 蔡原, 王建福, 等. 鲤和鲫线粒体(mtDNA)全基因组分析[J]. *基因组学与应用生物学*, 2020, 39(1): 37-43.
- Ge Q Y, Cai Y, Wang J F, *et al.* Complete genome analysis of mtDNA in carp and crucian[J]. *Genomics and Applied Biology*, 2020, 39(1): 37-43 (in Chinese).
- [10] 李瑶瑶, 刘云国, 刘凌霄, 等. 鲮科鱼类线粒体全基因组序列结构特征及其系统发育信息分析[J]. *烟台大学学报(自然科学与工程版)*, 2016, 29(4): 271-279.
- Li Y Y, Liu Y G, Liu L X, *et al.* Mitochondrial genome characteristics and phylogenetic analysis of Salmonidae fishes family[J]. *Journal of Yantai University (Natural Science and Engineering Edition)*, 2016, 29(4): 271-279 (in Chinese).
- [11] Paxton J R, Hoese D F, Allen G R, *et al.* Pisces Petromyzontidae to Carangidae[M]. Canberra: Australian Biological Resources Survey, 1989.
- [12] 庞景贵. 长缟鲮养殖种苗生产现状和存在问题[J]. *现代渔业信息*, 1988(9): 19-22.
- Pang J G. Production status and problems of *Pseudocaranx dentex* cultured seedlings[J]. *Modern Fisheries Information*, 1988(9): 19-22 (in Chinese).
- [13] Nogueira N, Ferreira M, Cordeiro N, *et al.* Quality parameters of wild white trevally (*Pseudocaranx dentex*) natural spawn kept in captivity[J]. *Aquaculture*, 2018, 495: 68-77.
- [14] Takaoka O, Ji S C, Ishimaru K, *et al.* Medicinal herbs promote resistances of larval and juvenile striped jack *Pseudocaranx dentex* for various stresses[J]. *Aquaculture Science*, 2012, 60(2): 199-205.
- [15] 马青, 姜晨, 周丽青, 等. 黄带拟鲮染色体核型特征分析[J]. *中国水产科学*, 2021, 28(5): 561-568.
- Ma Q, Jiang C, Zhou L Q, *et al.* Karyotype characteristics of white trevally (*Pseudocaranx dentex*)[J]. *Journal of Fishery Sciences of China*, 2021, 28(5): 561-568 (in Chinese).
- [16] 肖武汉, 张亚平. 鱼类线粒体DNA的遗传与进化[J]. *水生生物学报*, 2000, 24(4): 384-391.
- Xiao W H, Zhang Y P. Genetics and evolution of mitochondrial DNA in fish[J]. *Acta Hydrobiologica Sinica*, 2000, 24(4): 384-391 (in Chinese).
- [17] Peng Z G, Wang J, He S P. The complete mitochondrial genome of the helmet catfish *Cranoglanis boudierius* (Siluriformes: Cranoglanididae) and the phylogeny of otophysan fishes[J]. *Gene*, 2006, 376(2): 290-297.
- [18] 刘凯, 冯晓宇, 马恒甲, 等. 钱塘江三角鲂线粒体基因组测序及其结构特征分析[J]. *浙江农业学报*, 2020, 32(9): 1591-1608.
- Liu K, Feng X Y, Ma H J, *et al.* Complete sequence and gene organization of mitochondrial genome of *Megalobrama terminalis* from Qiantang River[J]. *Acta Agriculturae Zhejiangensis*, 2020, 32(9): 1591-1608 (in Chinese).
- [19] 马江茹. 西里伯斯青鳞线粒体基因组全序列和中华青鳞群体遗传结构的研究 [D]. 湛江: 广东海洋大学, 2020: 29-30.
- Ma J R. The complete mitochondrial genome sequence of *Oryzias celebensis* and the genetic structure of *Oryzias sinensis*[D]. Zhanjiang: Guangdong Ocean University, 2020: 29-30 (in Chinese).
- [20] 张方, 米志勇. 动物线粒体DNA的分子生物学研究进展[J]. *生物工程进展*, 1998, 18(3): 25-31.
- Zhang F, Mi Z Y. Advance in molecular biology of animal mitochondrial DNA[J]. *Progress in Biotechnology*, 1998, 18(3): 25-31 (in Chinese).
- [21] 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.
- [22] Zhu Y X, Chen Y, Cheng Q Q, *et al.* The complete mitochondrial genome sequence of *Schizothorax macropogon* (Cypriniformes: Cyprinidae)[J]. *Mitochondrial DNA*, 2013, 24(3): 237-239.
- [23] 郭新红, 刘少军, 刘巧, 等. 鱼类线粒体DNA研究新进展[J]. *遗传学报*, 2004, 31(9): 983-1000.
- 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 (in Chinese).
- [24] 孟乾, 张志勇, 张志伟, 等. 斑石鲷和条石鲷线粒体基

- 因组密码子使用分析[J]. 水产科学, 2020, 39(5): 702-709.
- Meng Q, Zhang Z Y, Zhang Z W, *et al.* Codon bias and cluster analysis on coding regions of mtDNA in rock porgy *Oplegnathus punctatus* and rock bream *O. fasciatus*[J]. Fisheries Science, 2020, 39(5): 702-709 (in Chinese).
- [25] 李博. 三种鲷科鱼类线粒体全基因组的测定及鲷科鱼类系统发育分析 [D]. 武汉: 华中农业大学, 2016: 29-30.
- Li B. Complete mitochondrial genome sequence of three sisoridae fishes and phylogenetic analysis of sisoridae[D]. Wuhan: Huazhong Agricultural University, 2016: 29-30 (in Chinese).
- [26] Yokobori S, Pääbo S. Transfer RNA editing in land snail mitochondria[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1995, 92(22): 10432-10435.
- [27] Zardoya R, Meyer A. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates[J]. *Molecular Biology and Evolution*, 1996, 13(7): 933-942.
- [28] 吴仁协, 李超, 刘静. 鲷亚目鱼类线粒体16S rRNA基因序列变异及其分子系统进化关系[J]. *水产学报*, 2013, 37(1): 16-25.
- Wu R X, Li C, Liu J. Mitochondrial 16S rRNA sequence variation and molecular phylogenetic relationships of Stromateoidei fishes (Teleostei: Perciformes)[J]. *Journal of Fisheries of China*, 2013, 37(1): 16-25 (in Chinese).
- [29] Nelson J S. Fishes of the world[M]. 3rd ed. New York: John Wiley & Sons, Inc, 1994: 600.
- [30] 郑文娟, 朱世华, 邹记兴, 等. 基于16S rRNA部分序列探讨12种鲷科鱼类的分子系统进化关系[J]. *水产学报*, 2008, 32(6): 847-854.
- Zheng W J, Zhu S H, Zou J X, *et al.* Molecular phylogenetic relationship of Carangidae based on partial sequence of mitochondrial 16S ribosomal RNA gene[J]. *Journal of Fisheries of China*, 2008, 32(6): 847-854 (in Chinese).
- [31] 赵凯. 鱼类线粒体DNA(mtDNA)及其在分子系统学中的应用[J]. *青海大学学报(自然科学版)*, 2006, 24(2): 49-53.
- Zhao K. Mitochondrial DNA (mtDNA) of fish and application in molecular phylogenetic[J]. *Journal of Qinghai University (Nature Science Edition)*, 2006, 24(2): 49-53 (in Chinese).
- [32] Near T J, Pesavento J J, Cheng C H C. Phylogenetic investigations of Antarctic notothenioid fishes (Perciformes: Notothenioidei) using complete gene sequences of the mitochondrial encoded 16S rRNA[J]. *Molecular Phylogenetics and Evolution*, 2004, 32(3): 881-891.

Complete mitochondrial genome sequence of *Pseudocaranx dentex* and phylogenetic analysis of Carangidae

WANG Kaijie^{1,2}, JIANG Yan¹, XU Yongjiang^{1*}, LIU Xuezhou¹, CUI Aijun¹,
WANG Bin¹, XUE Zhiyong³, MAO Chengquan³

- (1. Joint Laboratory for Deep Blue Fishery Engineering of Pilot National Laboratory for Marine Science and Technology, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China;
2. National Engineering Research Center For Marine Aquaculture, Zhejiang Ocean University, Zhoushan 316022, China;
3. Haiyang Yellow Sea Fishery Company, Yantai 265122, China)

Abstract: *Pseudocaranx dentex* belongs to Actinopterygii, Perciformes, Carangidae, Caranginae, *Pseudocaranx*, and it is widely distributed in the warm waters of Indo-Pacific and Atlantic Ocean. It is mainly distributed in the South China Sea and East China Sea. The research on the breeding of *P. dentex* in China is still in the initial stage, and there are few reports about the genetic background of this species. As a species with high nutritional and economic values, it was studied from the perspective of molecular biology. The complete mitochondrial genome of *P. dentex* was obtained by second-generation sequencing and analyzed through the related software. The results showed that the mitochondrial genome was a typical circular DNA structure, with the total length of the sequence 16 570 bp, and the base composition was A (27.2%), G (17.18%), C (30.24%) and T (25.38%). It contained 13 protein-coding genes, 22 tRNA genes and 2 rRNA genes. Except for *ND6*, tRNA^{Gln}, tRNA^{Ala}, tRNA^{Asn}, tRNA^{Cys}, tRNA^{Tyr}, tRNA^{Ser}, tRNA^{Glu}, tRNA^{Pro}, all the other genes were encoded on the H-strand. Except for ATC and ATA, the starting codon of *CO I* and *ND5* were ATG; the A + T contents of mitochondrial genome and protein encoding genes were 52.58% and 51.55% respectively, and the noncoding control region (D-loop) is rich in 61.69%, which had an obvious A+T bias. Among the 22 tRNA genes, except for tRNA^{Ser}-GCT, the other 21 tRNA genes contained typical secondary structure of clover. The phylogenetic tree was constructed with the complete sequence of mitochondrial genome of 18 species belonging to 13 genera of Carangidae fishes. Each species has an independent branch, and *P. dentex* and *Carangoides equula* belong to the same branch in the evolutionary tree, which indicated that these two species are closely related to each other. Our findings could help to explore the genetic relationship between *P. dentex* and other Carangidae fish, and provide a basis for germplasm identification, phylogenetic evolution analysis, genetic diversity evaluation and utilization of *P. dentex*.

Key words: *Pseudocaranx dentex*; mitochondrial genome; sequence analysis; phylogeny

Corresponding author: XU Yongjiang. E-mail: xuyj@ysfri.ac.cn

Funding projects: National Key R & D Program of China (2019YFD0900901, 2018YFD0901204); Central Public-Interest Scientific Institution Basal Research Fund, CAFS (2020TD47); Central Public-interest Scientific Institution Basal Research Fund, YSFRI, CAFS (20603022021011); China Agriculture Research System of MOF and MARA (CARS-47)