

DOI: 10.3724/SP.J.1118.2019.18365

## 黄条鲷线粒体全基因组测序及结构特征分析

史宝<sup>1</sup>, 柳学周<sup>1,2</sup>, 刘永山<sup>1,2</sup>, 张言祥<sup>3</sup>, 高全义<sup>3</sup>, 徐永江<sup>1</sup>, 王滨<sup>1</sup>, 姜燕<sup>1</sup>,  
宋雪松<sup>1,2</sup>

1. 青岛海洋科学与技术试点国家实验室海洋渔业科学与食物产出过程功能实验室, 中国水产科学研究院黄海水产研究所, 山东 青岛 266071;
2. 上海海洋大学水产与生命学院, 上海 201306;
3. 大连富谷水产有限公司, 辽宁 大连 116400

**摘要:** 通过二代测序、软件拼接获得中国黄海海域黄条鲷(*Seriola aureovittata*)线粒体基因组全序列。其序列全长为16609 bp, 碱基组成分别为 A(26.68%)、G(17.84%)、C(30.12%)和 T(25.36%); 共有 13 条蛋白编码基因, 22 个 tRNA 基因, 2 个 rRNA 基因, 除 *NAD6*、*trnQ*、*trnA*、*trnN*、*trnC*、*trnY*、*trnS*、*trnE*、*trnP* 外, 其余基因均在 H 链上编码。黄条鲷线粒体基因组全序列与蛋白编码基因的 A+T 含量分别为 52.05%和 51.085%, 具有明显的 AT 偏好性。线粒体基因中存在 2 个散在重复序列, 分别位于 *NAD1* 基因序列正义链的中游和 *COX2* 基因序列反义链的上游。在其 22 个 tRNA 基因中, 除了 tRNA<sup>Gly</sup> 外, 均具有典型的三叶草二级结构。黄条鲷线粒体基因组的蛋白编码基因起止位点与密码子除 *COX1*、*NAD5* 外, 均与日本海域出产的黄条鲷(*Seriola lalandi*)完全吻合, 且 *COX1*、*NAD5* 基因皆短于日本黄条鲷; 两者间存在一定的遗传差异。基于 18 种隶属于 13 属的鲷科鱼类线粒体基因组全序列构建系统发生树, 可知小甘鲷(*Seriolina nigrofasciata*)、黄条鲷、五条鲷(*Seriola quinqueradiata*)、高体鲷(*Seriola dumerili*)、长鳍鲷(*Seriola rivoliana*)同属一近支, 且黄条鲷与五条鲷亲缘关系最近, 与小甘鲷进化距离较远。

**关键词:** 黄条鲷; 线粒体基因组; 序列分析; 系统进化

中图分类号: 917

文献标志码: A

文章编号: 1005-8737-(2019)03-0405-11

由于线粒体 DNA (mitochondrial DNA, mtDNA) 独立于核基因组之外, 在自我复制、遗传与表达方面拥有较高的自主性, 且兼有分子小、碱基替代率高、母系遗传等特点, 被越来越多地应用于分子标记方法与系统进化研究等领域<sup>[1-7]</sup>, 以更好地了解物种的进化与多样性<sup>[8]</sup>。基于物种线粒体全基因组序列测定的系统进化分析已经成为研究物种进化谱系不可或缺的重要工具, 并取得了丰富的研究成果。在水产研究领域, 线粒体基因

组的全序列测序工作也已经成为一个热门话题。通过对鱼类线粒体基因组的比对和分析, 研究者们得到了大量的系统数据, 为多种鱼类的系统进化研究提供了新的方向与依据。

黄条鲷(*Seriola aureovittata*)属鲷形目(Perciformes)、鲷亚目(Percoidei)、鲷科(Carangidae)、鲷属, 是暖温带海洋中上层大型掠食性鱼类, 在中国主要分布于黄渤海, 具有体型大、生长快、肉质佳等特点, 有较高的经济价值与市场前景<sup>[9]</sup>。

收稿日期: 2018-11-01; 修订日期: 2019-01-23.

基金项目: 现代农业产业技术体系专项经费资助项目(CARS-47); 中国水产科学研究院中央级公益性科研院所基本科研业务费专项资金(2018GH17); 青岛海洋科学与技术国家实验室海洋渔业科学与食物产出过程功能实验室开放课题(2017-3A01); 青岛市民生科技计划项目(17-3-3-61-nsh); 中国水产科学研究院黄海水产研究所基本科研业务费项目(20603022017016); 国家自然科学基金项目(31772829).

作者简介: 史宝(1979-), 男, 博士, 副研究员, 主要从事鱼类繁育理论与增殖技术研究. E-mail: shibao@ysfri.ac.cn

通信作者: 柳学周, 研究员. E-mail: liuxz@ysfri.ac.cn

多年来,国内外学者大多关注黄条鲷繁育、养殖技术的研究,而对它的基因组或群体组成研究很少。Quéro 等<sup>[10]</sup>曾依据形态学特征将来自全球不同产区的黄条鲷统一命名为“*Seriola lalandi*”。基因测序技术的成熟与广泛应用,为黄条鲷的种质研究提供了新途径。近年,日本学者通过线粒体基因组序列特征探讨了五条鲷(*Seriola quinqueradiata*)、高体鲷(*Seriola dumerili*)、长鳍鲷(*Seriola rivoliana*)、日本海域出产的黄条鲷(*Seriola lalandi*)四者的种间关系<sup>[11]</sup>。澳大利亚学者通过基因测序技术,对黄条鲷种群的单一性提出了不同看法,将中国、日本海域的黄条鲷划归于西北太平洋种群,并命名为“*Seriola aureovittata*”<sup>[12]</sup>。中国学者对从黄海海域捕获的黄条鲷命名为“*Seriola aureovittata*”<sup>[9, 13-14]</sup>,而日本学者对日本海域出产的黄条鲷命名为“*Seriola lalandi*”<sup>[15]</sup>。本研究测定了中国青海海域黄条鲷的线粒体全基因组序列,分析了其主要结构信息,通过与日本海域黄条鲷线粒体基因组特征的比较,证明两者间存在遗传差异;通过与 18 种鲷科鱼类线粒体基因组进行比对,构建鲷科鱼类的分子系统发生树,探讨了黄条鲷与其他鲷属鱼类的亲缘关系。该研究结果为中国黄海海域黄条鲷的种质鉴定、地理种群划分和亲缘关系鉴定提供了重要的基础信息。

## 1 材料与方法

### 1.1 实验材料

本研究以 2017 年在黄海北部大连附近海域捕获的野生黄条鲷为材料,对其线粒体基因组进行测序。

### 1.2 样品采集及 DNA 提取

使用灭菌解剖工具解剖 2017 年捕获的黄条鲷幼鱼,取其体侧的肌肉 20 g,置于冻存管中,液氮保存。使用天根 DNA 提取试剂盒(DP121221),参考试剂盒使用说明,提取肌肉 DNA。

### 1.3 基因测序

委托深圳恒创基因科技有限公司,使用 Illumina 二代测序方法(next-generation sequencing)<sup>[16]</sup>,对黄条鲷肌肉 DNA 样品进行高通量测序。通过质量检测,将合格的样品构建文库,回收

目的扩增片段。用 T4 DNA 聚合酶、Klenow DNA 聚合酶和 T4 PNK 将打断形成的黏性末端修复成平整末端,通过 3'端带有 T 碱基的特殊接头连接;最后用合格的文库进行 Cluster 制备和测序。使用 HisEQ 和 Miseq 测序平台对样品进行测序,测序仪型号为 PE150 Illumina。

### 1.4 序列拼接与分析

使用 DNASTar<sup>[17]</sup>软件进行序列校对、排列、拼接,得到黄条鲷线粒体基因组全序列。通过 Blast 同源序列检索及软件 DNASTar 和 Genetyx 分析蛋白编码基因、rRNA 基因和 D-loop 区。使用 Editseq7.1 统计序列长度、碱基组成、GC 含量及氨基酸偏好性。使用 tRNA scan-SE 软件(<http://www.Genetics.Wustl.Edu./eddy/tRNAscan-SE>)进行 tRNA 定位,RNAstructure 在线预测 tRNA 茎环结构(温度参数设定为 298.15 K)。使用 RepeatMasker 搜索并标记重复序列。

从 GenBank 中下载 18 种鲷科鱼类线粒体基因组序列,包括大甲鲷(*Megalaspis cordyl*, KM522836.1; 16566 bp)、游鳍叶鲷(*Atule mate*, KM522838.1; 16565 bp)、丽叶鲷(*Alepes kleinii*, KF728081.1; 16571bp)、金带细鲷(*Selaroides leptolepis*, KM522839.1; 16560 bp)、马拉巴若鲷(*Carangoides malabaricus*, KJ174514.1; 16561 bp)、甲若鲷(*Carangoides armatus*, AP004444.1; 16556 bp)、乌鲷(*Parastromateus niger*, KJ192332.1; 16561 bp)、白舌尾甲鲷(*Uraspis secunda*, KT819204.1; 16554 bp)、脂眼凹肩鲷(*Selar crumenophthalmus*, KJ148633.1; 16610 bp)、高体若鲷(*Carangoides equula*, KM201334.1; 16588 bp)、竹荚鱼(*Trachurus trachurus*, AB108498.1; 16559 bp)、长身圆鲷(*Decapterus macrosoma*, KF841444.1; 16545 bp)、纺锤鲷(*Elagatis bipinnulata*, KT824759.1; 16542 bp)、小甘鲷(*Seriolina nigrofasciata*, KT591876.1; 16531 bp)、日本黄条鲷(*Seriola lalandi*, AB517557.1; 16532 bp)、五条鲷(*Seriola quinqueradiata*, AB517556.1; 16537 bp)、高体鲷(*Seriola dumerili*, AB517558.1; 16530 bp)、长鳍鲷(*Seriola rivoliana*, KP347126.1/KP733847.1; 16530 bp/16599 bp);使用 Clustal X1.83<sup>[18]</sup>进行多重比对,应用 MEGA 6.0<sup>[19]</sup>软件构建 Neighbor-



表 2 黄条鲷线粒体基因组结构特点  
Tab. 2 Characteristics of the mitogenomes  
of *Seriola aureovittata*

基因名称 gene name	起止位置 begin and end	长度/bp length	编码链 strand	起始密码 initiation condon	终止密码 termination condon
<i>trnF-GAA</i>	1-68	68	H	-	-
<i>12S rRNA</i>	69-1019	951	H	-	-
<i>trnV-UAC</i>	1020-1092	72	H	-	-
<i>16S rRNA</i>	1110-2808	1699	H	-	-
<i>trnL-UAA</i>	2809-2883	75	H	-	-
<i>NAD1</i>	2884-3858	975	H	ATG	TAA
<i>trnI-GAU</i>	3863-3932	70	H	-	-
<i>trnQ-UUG</i>	3932-4002	71	L	-	-
<i>trnM-CAU</i>	4002-4070	69	H	-	-
<i>NAD2</i>	4071-5117	1047	H	ATG	TAA
<i>trnW-UCA</i>	5117-5187	71	H	-	-
<i>trnA-UGC</i>	5189-5257	69	L	-	-
<i>trnN-GUU</i>	5259-5331	73	L	-	-
<i>trnC-GCA</i>	5366-5431	66	L	-	-
<i>trnY-GUA</i>	5432-5501	70	L	-	-
<i>COX1</i>	5584-7053	1470	H	ATG	TAA
<i>trnS-UGA</i>	7054-7124	71	L	-	-
<i>trnD-GUC</i>	7129-7199	71	H	-	-
<i>COX2</i>	7208-7898	691	H	ATG	T-
<i>trnK-UUU</i>	7899-7973	75	H	-	-
<i>ATP8</i>	7975-8142	168	H	ATG	TAA
<i>ATP6</i>	8133-8816	684	H	ATG	TAA
<i>COX3</i>	8816-9601	786	H	ATG	TAA
<i>trnG-UCC</i>	9601-9671	71	H	-	-
<i>NAD3</i>	9672-10022	351	H	ATG	TAG
<i>trnR-UCG</i>	10021-10089	69	H	-	-
<i>NAD4L</i>	10090-10386	297	H	ATG	TAA
<i>NAD4</i>	10380-11760	1381	H	ATG	T-
<i>trnH-GUG</i>	11761-11829	69	H	-	-
<i>trnS-GCU</i>	11830-11896	67	H	-	-
<i>trnL-UAG</i>	11901-11973	73	H	-	-
<i>NAD5</i>	11992-13812	1821	H	ATA	TAA
<i>NAD6</i>	13809-14330	522	L	ATG	TAG
<i>trnE-UUC</i>	14332-14400	69	L	-	-
<i>Cob</i>	14405-15545	1141	H	ATG	T-
<i>trnT-UGU</i>	15546-15617	72	H	-	-
<i>trnP-UGG</i>	15617-15687	71	L	-	-
D-loop	15689-16609	921	H	-	-

黄条鲷线粒体基因组共有 22 个 tRNA, 序列长度范围为 66~75 bp, 总长度为 1552 bp(表 2)。其中各有两个 tRNA<sup>Ser</sup>和 tRNA<sup>Leu</sup>; tRNA<sup>Ser</sup>的反密码子分别为 TGA 和 GCT, tRNA<sup>Leu</sup>的反密码子分别为 TAA 和 TAG。除了位于 L 链上的 8 个 tRNA

(*trnQ*、*trnA*、*trnN*、*trnC*、*trnY*、*trnS*、*trnE*、*trnP*)外, 其余 14 个皆位于 H 链上。

在结构上, 黄条鲷线粒体 tRNA 具有较多变异, C-T 转换现象也十分明显。除 tRNA<sup>Gly</sup>呈短棒状, tRNA<sup>Leu</sup>(TAA)D 茎与反密码子茎间的额外环上多一组茎环结构外, 其余皆为典型的三叶草型二级结构(图 2)。

氨基酸接受臂长度为 4~8 nt, 其中 tRNA<sup>Phe</sup>、tRNA<sup>Arg</sup>为 4 nt, tRNA<sup>Leu</sup>(TAA)、tRNA<sup>Ile</sup>、tRNA<sup>Met</sup>为 5 nt, tRNA<sup>Gly</sup>为 8 nt, 其余皆为 6~7 nt。在 tRNA<sup>Val</sup>、tRNA<sup>Ile</sup>、tRNA<sup>Ser</sup>(TGA)、tRNA<sup>His</sup>接受臂中各有 1 对因 C-T 转换造成的 A-C 不配对, 在 tRNA<sup>Phe</sup>、tRNA<sup>Gln</sup>、tRNA<sup>Ala</sup>、tRNA<sup>Cys</sup>、tRNA<sup>Tyr</sup>、tRNA<sup>Asp</sup>、tRNA<sup>Glu</sup>、tRNA<sup>Pro</sup>接受臂上皆有因 C-T 转换造成的 G-T 错配。同时, 在 tRNA<sup>Ser</sup>(GCT)接受臂上有 1 对 A-A 不配对, 在 tRNA<sup>Leu</sup>(TAG)、tRNA<sup>Thr</sup>上各有 1 对 C-C 不配对。

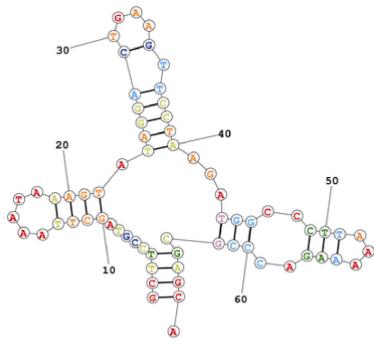
反密码子茎长度除 tRNA<sup>Pro</sup>(9nt)外, 皆在 4~6 nt 之间; 除了 C-T 转换造成的错配与不配外, 在 tRNA<sup>Phe</sup>中多 1 个碱基 T 的插入, 在 tRNA<sup>Met</sup>中多 1 个碱基 C 的插入, 在 tRNA<sup>Leu</sup>(TAG)中多 1 个碱基 A 的插入, 在 tRNA<sup>Ala</sup>中存在 1 对 G-G 不配对。反密码子环长度为 3~9 bp, 除 tRNA<sup>Phe</sup>、tRNA<sup>Met</sup>、tRNA<sup>Leu</sup>(TAG)长度为 4 bp 外, 其余 tRNA 长度皆为奇数。

在 TΨC 茎上, 茎长介于 4~6 nt 之间, 同样存在较多的 C-T 转换现象, 除此之外, tRNA<sup>Phe</sup>有 1 对 C-C 不配对, tRNA<sup>Met</sup>有 1 组 TT-TC 不配对, tRNA<sup>Gln</sup>有 1 个 T 碱基的插入。TΨC 环长度为 3~9 bp, 多为 7 bp。除 tRNA<sup>Met</sup>、tRNA<sup>Ile</sup>长度分别为 4 bp、8 bp 外, 其余 tRNA 长度皆为奇数。

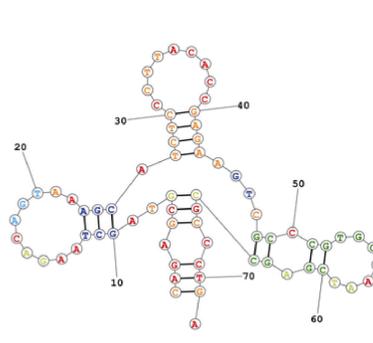
D 茎长度 2~5 nt, 最短为 tRNA<sup>Trp</sup>, 最长为 tRNA<sup>Gln</sup>和 tRNA<sup>Asn</sup>; 同样存在多组 G-T 错配, 且在 tRNA<sup>Leu</sup>(TAA)中多了 1 个碱基 T 的插入。D 环长度为 3~12 bp, tRNA<sup>Leu</sup>(TAA)、tRNA<sup>Arg</sup>、tRNA<sup>Ser</sup>(GCT)最短, tRNA<sup>Lys</sup>最长。

### 2.3 蛋白质编码基因

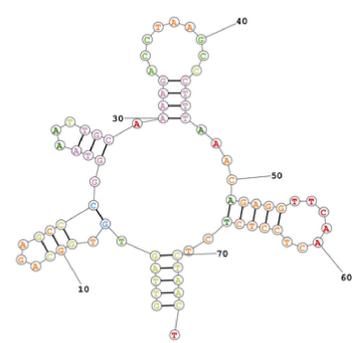
黄条鲷线粒体全基因组中包含 13 个蛋白编码基因, 共编码 154 个碱性氨基酸、177 个酸性氨



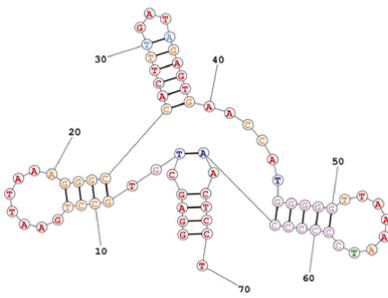
tRNA<sup>Phe</sup>-GAA



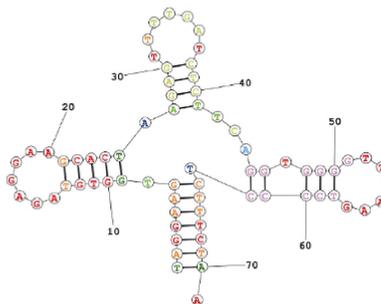
tRNA<sup>Val</sup>-TAC



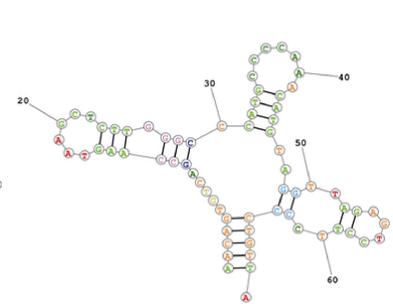
tRNA<sup>Leu</sup>-TAA



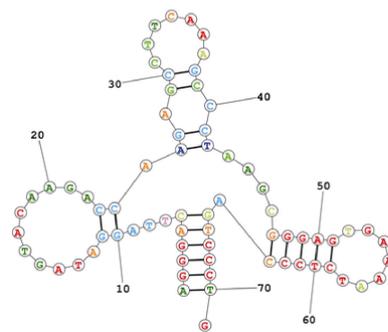
tRNA<sup>Ile</sup>-GAT



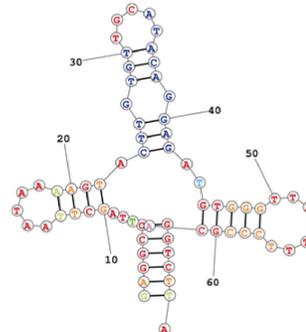
tRNA<sup>Gln</sup>-TTG



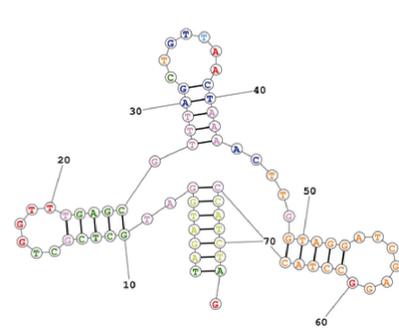
tRNA<sup>Met</sup>-CAT



tRNA<sup>Trp</sup>-TCA



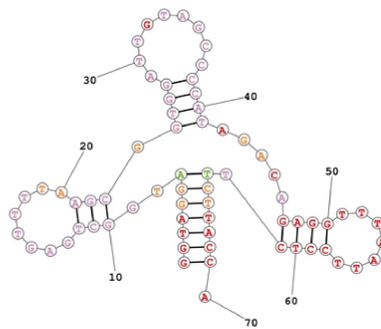
tRNA<sup>Ala</sup>-TGC



tRNA<sup>Asn</sup>-GTT



tRNA<sup>Cys</sup>-GCA



tRNA<sup>Tyr</sup>-GTA

Probability  $\geq$  99%  
99% > Probability  $\geq$  95%  
95% > Probability  $\geq$  90%  
90% > Probability  $\geq$  80%  
80% > Probability  $\geq$  70%  
70% > Probability  $\geq$  60%  
60% > Probability  $\geq$  50%  
50% > Probability

(待续 to be continued)

(续图 2 Fig. 2 continued)

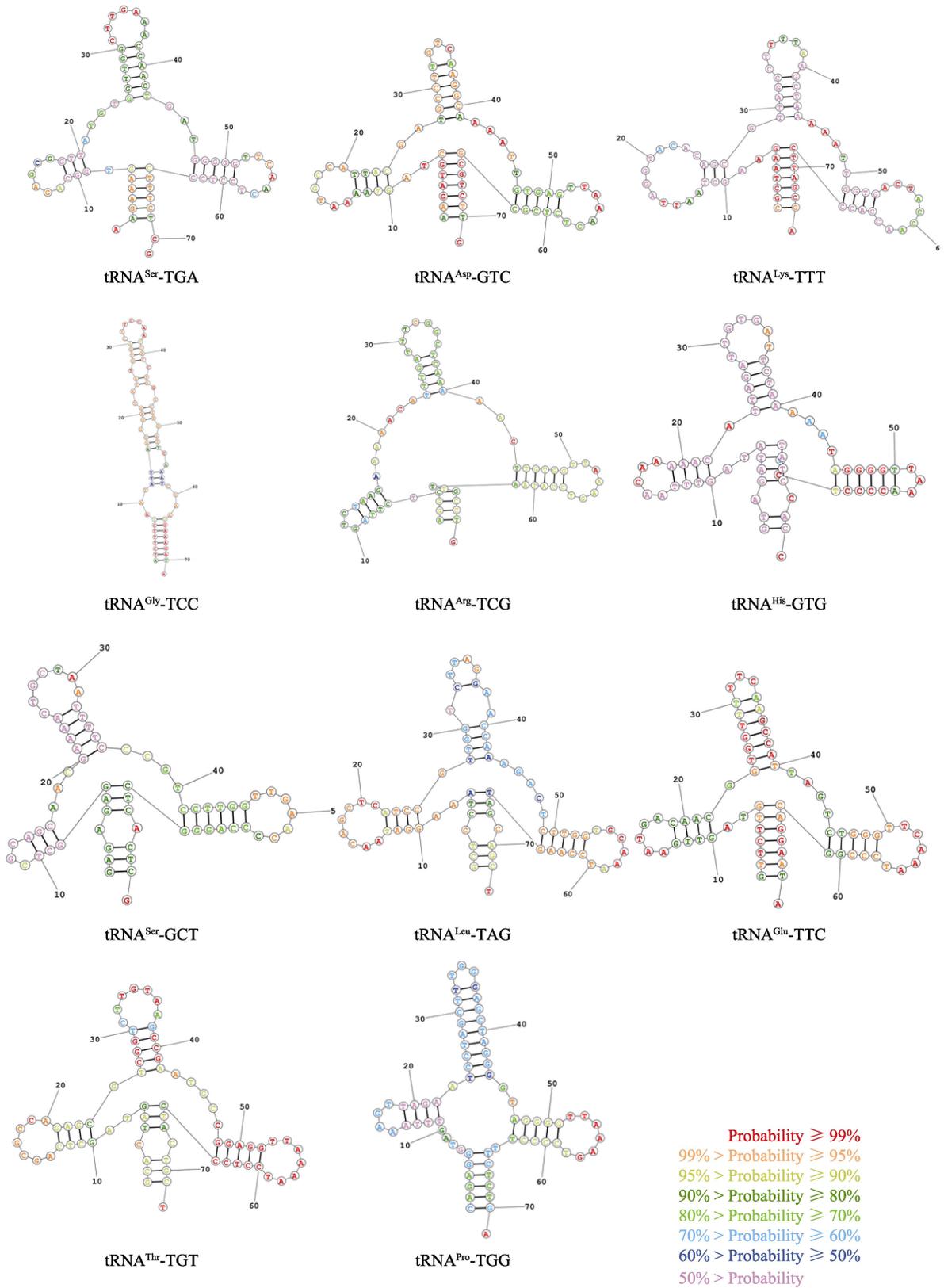


图 2 黄条鲷线粒体 tRNA 二级结构预测图

Fig. 2 Predicted secondary structure of *Seriola aureovittata*'s mitochondrial tRNA

基酸、1845 个极性氨基酸和 857 个非极性氨基酸。13 个基因的总长度为 11838 bp, 平均长度约为 910.6 bp, 占基因组总长度的 71.27%, 除了 *NAD6* 编码基因位于 L 链上之外, 其余 12 个蛋白编码基因都位于 H 链上(表 2)。

除了 *NAD5* 基因的起始密码子是“ATA”外, 其余 12 条基因的起始密码子皆为“ATG”。而终止密码子以“TAA”为主, 共 8 条; *NAD3*、*NAD6* 的终止密码子为“TAG”; 剩余的 3 条基因为不完全的终止密码子“T-”。

基因组中存在多处重叠现象。其中, *ATP8* 与 *ATP6* 序列有 10 bp 碱基的重叠, *ATP6* 与 *COX3* 有 1 bp 碱基的重叠, *NAD4L* 与 *NAD4* 重叠了 7 bp, *NAD5* 与 *NAD6* 有 4 bp 重叠。

在其 13 条蛋白编码基因中, 除了 *NAD1*、*NAD2*、*NAD4L*、*NAD6*、*COB* 外, 其他 8 条蛋白编码基因的 A+T 含量均高于 50%(表 3), 且 13 个蛋白编码基因的 A+T 总含量为 51.085%, 可见其在蛋白编码基因组中也具有 AT 偏好性。由于黄条鰺线粒体全基因组的 A+T 含量为 52.05%, 可推知其非蛋白编码区基因组也具有 AT 偏好性。

#### 2.4 中国黄海海域黄条鰺与日本海域黄条鰺蛋白编码基因差异

表 4 比较了中国黄海海域黄条鰺(*Seriola*

表 3 黄条鰺蛋白编码基因碱基组成  
Tab. 3 Nucleotide composition of protein-coding genes in *Seriola aureovittata*

基因 gene	碱基/bp nucleotide	T	C	A	G	A+T
NAD1	975	26.36	33.44	21.74	18.46	48.10
NAD2	1470	24.83	36.77	23.78	14.61	48.62
COX1	1551	28.30	28.84	23.95	18.91	52.24
COX2	691	28.08	26.63	29.09	16.21	57.16
ATP8	168	23.81	32.74	29.17	14.29	52.98
ATP6	684	27.63	32.46	25.00	14.91	52.63
COX3	786	26.97	31.04	23.92	18.07	50.89
NAD3	351	29.63	32.48	20.51	17.38	50.14
NAD4I	297	30.64	35.02	18.18	16.16	48.82
NAD4	1381	24.55	32.95	26.07	16.44	50.62
NAD5	1821	26.03	32.24	26.36	15.38	52.39
NAD6	522	36.02	16.86	13.41	33.72	49.43
COB	1141	26.73	34.36	23.05	15.86	49.78

*aureovittata*)与日本海域黄条鰺(*Seriola lalandi*)在蛋白编码基因上的异同。二者的蛋白编码基因在起止位置、序列长度以及密码子选择偏好上均高度相似, 所有差异均来自于 *COX1* 蛋白基因与 *NAD5* 蛋白基因。

中国黄海海域黄条鰺的 *COX1* 蛋白基因与 *NAD5* 蛋白基因均略短于日本海域黄条鰺。*COX1* 蛋白编码基因的起始位置后移 81 bp, 终止位置不变, 总长度比日本海域黄条鰺 *COX1* 蛋白基因

表 4 中国黄海海域黄条鰺与日本海域黄条鰺蛋白编码基因对比  
Tab. 4 Protein coding gene comparisons between *Seriola aureovittata* and *Seriola lalandi*

基因 gene	中国黄海海域黄条鰺 <i>S. aureovittata</i>				日本海域黄条鰺 <i>S. lalandi</i>			
	起止位置/bp start-end	长度/bp length	起始密码 start code	终止密码 stop code	起止位置 start-end	长度/bp length	起始密码 start code	终止密码 stop code
NAD1	2884-3858	975	ATG	TAA	2884-3858	975	ATG	TAA
NAD2	4071-5117	1047	ATG	TAA	4071-5117	1047	ATG	TAA
COX1	5584-7053	1470	ATG	TAA	5503-7053	1551	GTG	TAA
COX2	7208-7898	691	ATG	T--	7208-7898	691	ATG	T--
ATP8	7975-8142	168	ATG	TAA	7975-8142	168	ATG	TAA
ATP6	8133-8816	684	ATG	TAA	8133-8816	684	ATG	TAA
COX3	8816-9601	786	ATG	TAA	8816-9601	786	ATG	TAA
NAD3	9672-10022	351	ATG	TAG	9672-10022	351	ATG	TAG
NAD4L	10090-10386	297	ATG	TAA	10090-10386	297	ATG	TAA
NAD4	10380-1760	1381	ATG	T--	10380-1760	1381	ATG	T--
NAD5	11992-3812	1821	ATA	TAA	11974-3812	1839	ATG	TAA
NAD6	13809-14330	522	ATG	TAG	13809-14330	522	ATG	TAG
COB	14405-5545	1141	ATG	T--	4405-5545	11141	ATG	T--

短 81 bp。NAD5 蛋白编码基因起始位置后移 18 bp, 终止位置不变, 总长度比日本海域黄条鲷 NAD5 蛋白基因短 18 bp。中国海域黄条鲷 COX1 蛋白与 NAD5 蛋白起始密码子分别为 ATG、ATA, 而日本海域黄条鲷 COX1 蛋白与 NAD5 蛋白起始密码子分别为 GTG、ATG。

### 2.5 黄条鲷线粒体基因组系统进化分析

将黄条鲷线粒体基因组全序列与其他 18 种、分属 13 个属的鲷科鱼类(包括 4 种鲷属鱼类)线粒体基因组全序列进行多重序列比对和分析, 以研究黄条鲷的分类地位和遗传背景。使用 MEGA 6.0 的邻接法构建分子系统发生树, 分析黄条鲷系统进化关系。该系统进化树由两个大支构成, 分别为由大甲鲷属、叶鲷属、副叶鲷属、细鲷属、若鲷属、乌鲷属、尾甲鲷属、凹肩鲷属、竹荚鱼属、圆鲷属等构成的鲷科鱼类进化主体, 以及由纺锤鲷属、小条鲷属、鲷属所组成的另一条进化枝。

由图 3 可见, 鲷属鱼类所在的进化枝分化较早, 且类群丰富度低, 只有 3 属。其中, 小条鲷属的小甘鲷与鲷属的日本海域黄条鲷、五条鲷、高体鲷、长鳍鲷位于同一近支, 其亲缘关系较近。

而纺锤鲷属虽同处一支, 但分化极早, 是这一支里最早独立分化出来的鱼种, 因此与小条鲷属、鲷属鱼类类群的进化距离较远。在鲷属鱼类中, 黄条鲷与五条鲷亲缘关系最近, 同属于一个细枝。

### 3 讨论

目前, 鱼类线粒体基因组的研究迅速发展。鱼类的线粒体全基因组测序成果广泛应用于鱼类种内种间系统发生关系、物种起源及遗传分化、种内多态及地理分布关系、种质评估和种群遗传结构等研究领域<sup>[20]</sup>。NCBI 数据库公布的鱼类线粒体基因组全序列数据大多相似; 全序列长度一般在 16 kb 左右, 结构呈双链的闭合环状, 通常包括 22 个 tRNA 基因、13 个蛋白编码基因, 2 个 rRNA (16S rRNA 和 12S rRNA) 基因和 1 段 D-loop 控制区。

本研究测定了中国海域黄条鲷的线粒体基因组全序列, 其全长 16609 bp, GC 含量约 40.86%, 呈现出明显的 AT 碱基偏好性和反 G 偏移现象。其中 G 碱基含量仅为 17.84%, 与其他硬骨鱼类如克氏光唇鱼(*Acrossocheilus kreyenbergii*) (16.4%)<sup>[7]</sup>、

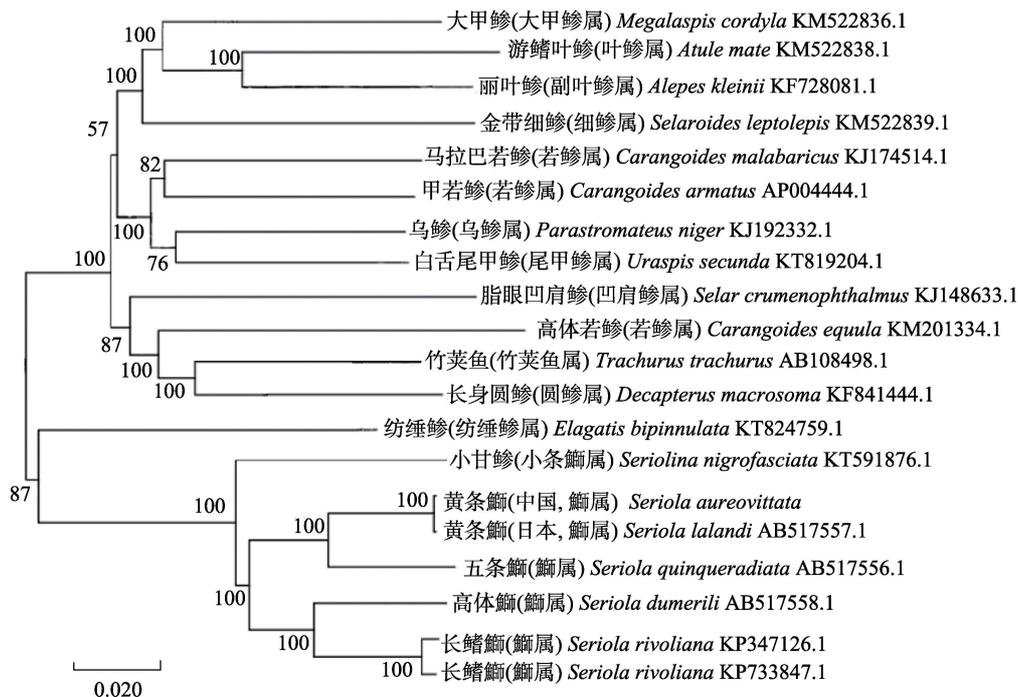


图 3 基于线粒体基因组全序列构建的系统进化树

Fig. 3 Molecular phylogenetic tree constructed based on mitochondrial genome by NJ method

小黄鱼(*Larimichthys polyactis*)(16%)<sup>[2]</sup>、圆斑星鲷(*Verasper variegatus*)(16.1%)<sup>[3]</sup>等鱼类线粒体基因中的 G 碱基含量接近; C 碱基含量为 30.12%, A 碱基含量为 26.68%, T 碱基含量为 25.36%, 这与典型脊椎动物碱基组成特点相符合<sup>[21]</sup>。钟东等<sup>[22]</sup>研究表明, 由于 DNA 双链之间存在自然突变及选择压力差异, 全基因组上碱基的分布常存在不均现象。在黄条鲷线粒体基因组中, H 链和 L 链上都有蛋白编码基因, 但是除 *NAD6* 外, 其余蛋白编码基因均在 H 链上。Brown 等<sup>[23]</sup>研究显示, 分布在 H 链上的线粒体基因容易形成没有保护的单链, 更易发生水解和氧化。黄条鲷线粒体基因组的蛋白编码基因和 RNA 主要分布在 H 链上, 说明黄条鲷线粒体基因组也容易发生水解和氧化, 而单独位于 L 链上的 *NAD6* 基因则相对稳定, 突显出 *NAD6* 基因的独特性与重要性。

黄条鲷线粒体基因组具有标准数量的 22 个 tRNA, 长度范围为 66~75 bp, 大多数符合三叶草型二级结构。茎环结构上的碱基突变频繁。黄条鲷线粒体基因组中 16S rRNA 和 12S rRNA 均为单拷贝且基因内部无间隔区, 符合后生动物的典型性特征<sup>[24]</sup>。线粒体基因组中的串联重复序列、大量插入序列与序列缺失等都是物种进化的特征<sup>[25]</sup>。因此, 此重复单元的特征及基因的序列多态性, 可作为种质鉴定和区分不同种群的重要分子标记。本研究中, 黄条鲷具有两个散在的重复序列, 这表明其属于进化的物种。而在日本海域黄条鲷中未报道发现具有异质性的重复序列。

Iguchi 等<sup>[11]</sup>报道的日本海域黄条鲷线粒体基因组特征, 与本研究中国黄海海域黄条鲷的线粒体基因组特征高度相似, 均含有 2 个 rRNA、22 个 tRNA 基因、13 个蛋白编码基因和 1 个控制区, 且各基因起止位点基本吻合。但是两者在基因组总体长度与蛋白编码基因上存在差异。日本海域黄条鲷线粒体基因组总长 16532 bp, 而中国黄海海域黄条鲷线粒体基因组总长 16609 bp, 出现这种差异可能是由于不同群体之间存在的非编码基因的长度多态性所致<sup>[26]</sup>。在蛋白编码基因方面, 中国黄海海域黄条鲷的 *COX1* 蛋白基因与 *NAD5* 蛋白基因皆略短于日本海域黄条鲷, 表现出位点

多态性。Martinez-Takeshita 等<sup>[12]</sup>根据黄条鲷的主要产区, 将其分为 3 个不同的地理种群, 即主要位于西北太平洋的种群“*Seriola aureovittata*”、主要位于东北太平洋的种群“*Seriola dorsalis*”和主要位于南太平洋的种群“*Seriola lalandi*”。而 Premachandra 等<sup>[27]</sup>则对此持怀疑态度, 并研究发现此 3 个群间尽管存在明显的遗传差异, 但仍未显著到足以支撑上述论点的程度; 同时该研究将中国、日本黄条鲷划归于位于西北太平洋的种群中, 而其对西北太平洋海域种群的研究结果显示样本的分组间存在差异。结合本研究, 我们推测此分组差异源于中国、日本黄条鲷群体间存在遗传差异。

随着基因领域的探索日益深入, 由物种单个基因构建的系统发生树已经不足以满足研究者们对更加详细、精确数据的需求。因此, 通过使用更长的线粒体 DNA 序列来构建更精确、可靠的系统发生树, 以获得能够更加全面反映生物分子进化水平的系统发育信息, 逐渐成为研究者的共识与主流<sup>[28-29]</sup>。本研究通过将中国海域黄条鲷线粒体基因组全序列与其他 18 种隶属于 13 属的鲷科鱼类线粒体基因组进行序列比对分析发现, 在鲷科鱼类类群进化体系中, 鲷属、小条鲷属及纺锤鲷属同属一个大支, 且分化较早, 与其他鲷科鱼类类群进化距离较远, 与鲷亚科类群互为姐妹群<sup>[30]</sup>。而鲷属(黄条鲷、五条鲷、高体鲷、长鳍鲷)与小条鲷属(小甘鲷)的亲缘关系较近。在鲷属类群中, 黄条鲷与五条鲷位于同一细枝, 亲缘关系最近, 二者可能具有相似的遗传背景。

#### 参考文献:

- [1] 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.
- [2] 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. [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.]
- [3] 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. [赫崇波, 高祥刚, 王效敏, 等. 圆斑星鲷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.]
- [4] Mao M G, Gu J, Liu R T, et al. Analysis of complete mitochondrial genome sequences of *Gadus macrocephalus*[J/OL]. Acta Hydrobiologica Sinica, 2018(06-18)/[2018-9-6]. <http://kns.cnki.net/kcms/detail/42.1230.Q.20180618.2206.002.html>. [毛明光, 顾杰, 刘瑞婷, 等. 太平洋鲳线粒体全基因组测序及结构特征分析[J/OL]. 水生生物学报, 2018(06-18)/[2018-9-6]. <http://kns.cnki.net/kcms/detail/42.1230.Q.20180618.2206.002.html>.]
- [5] Huang X L, Yang Y K, Lin H Z, et al. Phylogenetic information analysis of mitochondrial genome sequences in *Siganus* (Perciformes: Siganidae)[J]. Journal of Biology, 2018, 53(2): 207-219. [黄小林, 杨育凯, 林黑着, 等. 篮子鱼属线粒体基因组序列系统发育信息分析[J]. 生物学杂志, 2018, 53(2): 207-219.]
- [6] Jiang Z L, Zhang M, Lin Y, et al. Complete mitochondrial genome sequence and phylogenetic analysis of *Oreochromis aureus*[J]. Journal of Fishery Sciences of China, 2011, 18(4): 790-800. [蒋宗良, 张明, 林勇, 等. 奥利亚罗非鱼线粒体基因组全序列测定与系统进化分析[J]. 中国水产科学, 2011, 18(4): 790-800.]
- [7] Yang Y, Song X J, Tang W Q, et al. Complete mitochondrial genome of *Acrossocheilus kreyenbergii*, with phylogenetic analysis of genus *Acrossocheilus*[J]. Chinese Journal of Zoology, 2018, 53(2): 207-219. [杨杨, 宋小晶, 唐文乔, 等. 克氏光唇鱼线粒体基因组测定及光唇鱼属的系统发育分析[J]. 动物学杂志, 2018, 53(2): 207-219.]
- [8] Schuster S C. Next-generation sequencing transforms today's biology[J]. Nature Methods, 2008, 5: 16-18.
- [9] Li R, Xu Y J, Liu X Z, et al. Morphometric analysis and internal anatomy of yellowtail kingfish (*Seriola aureovittata*)[J]. Progress in Fishery Sciences, 2017, 38(1): 142-149. [李荣, 徐永江, 柳学周, 等. 黄条鲷(*Seriola aureovittata*)形态度量与内部结构特征[J]. 渔业科学进展, 2017, 38(1): 142-149.]
- [10] Quéro J C, Hureau J C, Post A, et al. Check-list of the fishes of the eastern tropical Atlantic[J]. Reviews in Fish Biology and Fisheries, 1992, 2(2): 182-184.
- [11] Iguchi J, Takashima Y, Namikoshi A, et al. Species identification method for marine products of *Seriola* and related species[J]. Fisheries Science, 2012, 78(1): 197-206.
- [12] Martinez-Takeshita N, Purcell C M, Chabot C L, et al. A tale of three tails: Cryptic speciation in a globally distributed marine fish of the genus, *Seriola*[J]. Copeia, 2015, 103(2): 357-368.
- [13] Shi B, Liu Y S, Liu X Z, et al. Study on the karyotype of yellowtail kingfish (*Seriola aureovittata*)[J]. Progress in Fishery Sciences, 2017, 38(1): 136-141. [史宝, 刘永山, 柳学周, 等. 黄条鲷(*Seriola aureovittata*)染色体核型分析[J]. 渔业科学进展, 2017, 38(1): 136-141.]
- [14] Meng Q W, Su J X, Miao X Z. Systematics of Fishes[M]. Beijing: China Agriculture Press, 1995: 672-673. [孟庆闻, 苏锦祥, 缪学祖. 鱼类分类学[M]. 北京: 中国农业出版社, 1995: 672-673.]
- [15] Balanov A A. On the species composition of fish of the genus *Seriola* (Carangidae) in the northwestern part of the Sea of Japan[J]. Journal of Ichthyology, 2008, 48(6): 415-421.
- [16] Beaumont A. Genetics and evolution of aquatic organisms[J]. Reviews in Fish Biology and Fisheries, 1995, 5(3): 385-386.
- [17] Burland T G. DNASTAR's Lasergene sequence analysis software[M]//Bioinformatics Methods and Protocols. Totowa: Humana Press, 2000, 132: 71-91.
- [18] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Research, 1997, 25(24): 4876-4882.
- [19] Kumar S, Stecher G, Peterson D, et al. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis[J]. Bioinformatics, 2012, 28(20): 2685-2686.
- [20] Xiao W H, Zhang Y P. Genetics and evolution of mitochondrial DNA in fish[J]. Acta Hydrobiologica Sinica, 2000, 24(4): 384-391. [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.]
- [21] Zhang F, Mi Z Y. Advance in molecular biology of animal mitochondrial DNA[J]. Progress in Biotechnology, 1998, 18(3): 25-31. [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.]
- [22] Zhong D, Zhao G J, Zhang Z S, et al. Advance in the entire balance and local unbalance of base distribution in genome[J]. Hereditas, 2002, 24(3): 351-355. [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, 24(3): 351-355.]
- [23] 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.
- [24] Wolstenholme D R, Jeon K W. Mitochondrial Genomes[M]. San Diego: Academic Press, 1992: 22.
- [25] Gao X G, Lu Z C, Tian J S, et al. Analysis and comparison of complete mitochondrial genome in spotted seal *Phoca largha*[J]. Fisheries Science, 2017, 36(4): 509-513. [高祥刚, 鹿志创, 田甲申, 等. 斑海豹线粒体基因组序列分析及比较研究[J]. 水产科学, 2017, 36(4): 509-513.]
- [26] 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. [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.]
- [27] Premachandra H K A, la Cruz F L, Takeuchi Y, et al. Genomic DNA variation confirmed *Seriola lalandi* comprises three different populations in the Pacific, but with recent divergence[J]. Scientific Reports, 2017, 7: Article No. 9386.
- [28] Zhang L L, Cheng Q Q. Mitochondrial genome characteri-

- stics and phylogenetic information of family Engraulidae (Clupeiformes: Clupeoidei) fishes[J]. *Marine Fisheries*, 2012, 34(1): 7-14. [张丽丽, 程起群. 鰺科鱼类线粒体全基因组序列结构特征及系统发育信息分析[J]. *海洋渔业*, 2012, 34(1): 7-14]
- [29] He H Y, Yu W D, Jiang W B. Research progress in mitochondrial genomics of butterfly[J]. *Chinese Bulletin of Life Sciences*, 2016, 28(9): 978-985. [何海燕, 俞伟东, 蒋韦斌. 蝶类线粒体基因组学研究进展[J]. *生命科学*, 2016, 28(9): 978-985.]
- [30] Zhu S H, Zheng W J, Zou J X, et al. Mitochondrial DNA control region structure and molecular phylogenetic relationship of Carangidae[J]. *Zoological Research*, 2007, 28(6): 606-614. [朱世华, 郑文娟, 邹记兴, 等. 鰺科鱼类线粒体DNA控制区结构及系统发育关系[J]. *动物学研究*, 2007, 28(6): 606-614.]

## Complete sequence and gene organization of the mitochondrial genome of *Seriola aureovittata*

SHI Bao<sup>1</sup>, LIU Xuezhou<sup>1,2</sup>, LIU Yongshan<sup>1,2</sup>, ZHANG Yanxiang<sup>3</sup>, GAO Quanyi<sup>3</sup>, XU Yongjiang<sup>1</sup>, WANG Bin<sup>1</sup>, JIANG Yan<sup>1</sup>, SONG Xuesong<sup>1,2</sup>

1. Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao); Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China;
2. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai, 201306, China;
3. Dalian Fugu Fishery Co., Ltd, Dalian 116400, China

**Abstract:** *Seriola aureovittata* is a large marine pelagic predatory fish species in the global warm temperate ocean, which is mainly distributed near the offshore rocks and reefs of the Bohai Sea and Yellow Sea of China. In recent years, most research has focused on its breeding and breeding technology at home and abroad, whereas few studies have been conducted on its genome and group composition. The complete mitochondrial DNA was extracted from the somatic side muscles of *S. aureovittata*. The complete mitochondrial genome of *S. aureovittata* was obtained using second generation sequencing and splicing with software of DNASTar. The sequence length was 16609 bp, and the base composition was A (26.68%), G (17.84%), C (30.12%), and T (25.36%). It contained 13 protein-coding genes, 22 tRNA genes and 2 rRNA genes. Except for *NAD6*, *trnQ*, *trnA*, *trnN*, *trnC*, *trnY*, *trnS*, *trnE*, and *trnP*, all the other genes were encoded on the heavy strand (H strand). The A+T content of the complete mitochondrial genome sequence and protein-coding genes were 52.05% and 51.085%, respectively. There was a strong bias toward A+T in the genome of *S. aureovittata*. There were two repeated sequences scattered in the genes. One was located in the middle part of the sense strand of the *NAD1* gene sequence, and the other was located in the upper part of the antisense strand of *COX2* gene sequence. All tRNA genes had typical cloverleaf structure except tRNA<sup>Gly</sup>. Its start-stop sites and codons of protein-coding genes, except *COX1* and *NAD5*, were the same as that of *Seriola lalandi*, which inhabits in the oceans of Japan. The *COX1* and *NAD5* protein genes of *S. aureovittata* were shorter than those of *S. lalandi*. It was preliminarily assumed that there were genetic differences between *S. aureovittata* and *S. lalandi*. The NJ-tree based on the complete mitochondrial DNA genomes showed the evolutionary position of *S. aureovittata* relative to that of 18 other Carangidae species. The results showed that the *Seriola* spp. belong to a secondary tree with *Seriolina nigrofasciata*. Furthermore, *S. aureovittata* and *Seriolina quinquerradiata* had a closer relationship than others in the tree, whereas *S. aureovittata* had the most distant relationship with *S. nigrofasciata*. This study provides basic data for solving different geographical population and kinship relations of *S. aureovittata*.

**Key words:** *Seriola aureovittata*; mitochondrial genome; sequence analysis; phylogenetic tree

**Corresponding author:** LIU Xuezhou. E-mail: liuxz@ysfri.ac.cn