

# 高体鲷(*Seriola dumerili*)线粒体全基因组测定 及结构特征分析\*

王开杰<sup>1,2,3</sup> 徐永江<sup>1,2</sup> 崔爱君<sup>1,2</sup> 柳学周<sup>1,2</sup> 姜燕<sup>1,2</sup> 王滨<sup>1,2</sup>

(1. 中国水产科学研究院黄海水产研究所 山东青岛 266071; 2. 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 山东青岛 266237; 3. 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)

**摘要** 为快速有效鉴别鲷属鱼类物种、加强鲷鱼遗传多样性管理与种质资源保护,通过 Illumina 测序技术,获得了东海海域养殖高体鲷(*Seriola dumerili*)线粒体基因组全序列(16 530 bp),碱基组成为 A (26.83%)、G (17.6%)、C (30.04%)和 T (25.53%), A+T 含量为 52.36%,且非编码控制区(D-loop) A+T 富含 61.64%,表现明显的 AT 偏好性。与其他硬骨鱼一样,高体鲷线粒体基因组包含 13 条蛋白编码基因,22 个 tRNA 基因,2 个 rRNA 基因,除 *ND6*、*tRNA<sup>Gln</sup>*、*tRNA<sup>Ala</sup>*、*tRNA<sup>Asn</sup>*、*tRNA<sup>Cys</sup>*、*tRNA<sup>Tyr</sup>*、*tRNA<sup>Ser</sup>*、*tRNA<sup>Glu</sup>*、*tRNA<sup>Pro</sup>* 基因外,其余均位于 H 链编码;蛋白编码基因中,除 *CO I*、*CO II* 和 *ND5* 的起始密码子分别为 ATC、ATA 和 ATA 外,其余 10 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 *ND4* 和 *Cytb* 中存在不完全密码子 T;除 *tRNA<sup>Ser</sup>-GCT* 外,其余 21 个 tRNA 均为典型三叶草二级结构。比较中国和日本海域高体鲷线粒体基因组发现,*CO I*、*CO II* 和 *ND5* 蛋白编码基因在起止位置、片段长度及起止密码子上存在显著差异。此外,与同属的黄条鲷(*Seriola aureovittata*)和五条鲷(*Seriola quinqueradiata*)的线粒体基因组 13 个蛋白编码基因进行两两对比分析,结果表明 3 种鲷属鱼类的蛋白编码基因的相似性在 85%~100%之间。基于线粒体基因组全序列构建的系统发育树,成功将高体鲷与其他鲷属鱼类有效区分,高体鲷与长鳍鲷同属一支,亲缘关系最近;黄条鲷和五条鲷聚为一支,亲缘关系最近。

**关键词** 高体鲷;线粒体基因组;序列比较分析;系统发育分析

中图分类号 S917 doi: 10.11693/hyhz20210700160

线粒体 DNA (Mitochondria DNA)具有结构简单、高拷贝数、分子量小、偏母性遗传且进化速度快等优点(Curole *et al.*, 1999)。鱼类线粒体 DNA 大小一般为 15~20 kb,多为双链闭合的环状结构,排列紧密,不同物种间线粒体基因组存在很大差异,序列中存在串联重复序列以及少数的散在重复序列(周传江等, 2019),与其他高等动物相似,鱼类线粒体基因组也是由 13 个蛋白质编码基因,2 个 rRNA,22 个 tRNA,以及非编码区重链复制起始相关的控制区(D-loop)和

轻链复制起始区(O<sub>L</sub>)组成(Satoh *et al.*, 2016)。随着 DNA 测序技术的逐渐成熟,鱼类线粒体基因组已被作为分子标记广泛应用于鱼类种质资源保护,群体多态性以及系统进化发育等领域(陈四海等, 2011)。黄小林等(2018)探讨了线粒体基因组在 6 种篮子鱼(*Siganus*)系统发育中的适用性;程佩琳等(2021)通过线粒体全基因组对 6 种鲟形目鱼类(Acipenseriformes)进行了种群划分;Bernal-Ramírez 等(2003)利用线粒体基因组控制区分析了新西兰金赤鲷(*Pagrus auratus*)

\* 国家重点研发计划项目,2019YFD0900901 号,2018YFD0901204 号;山东省支持青岛海洋科学与技术试点国家实验室重大科技专项,2018SDKJ0303-1 号;中国水产科学研究院基本科研业务费,TD47 号,2021GH05 号;国家海洋水产种质资源库项目;财政部和农业农村部:国家现代农业产业技术体系资助,CARS-47 号。王开杰,硕士研究生,E-mail: jasons@qq.com

通信作者:徐永江,研究员,E-mail: xuyj@ysfri.ac.cn

收稿日期:2021-07-09,收修改稿日期:2021-08-27

群体遗传结构的稳定性。

高体鲷(*Seriola dumerili*)属鲈形目(Perciformes)、鲈亚目(Percoidae)、鲷科(Carangidae)、鲷属(*Seriola*), 又名杜氏鲷、红甘鲷(陈大刚等, 2015)。高体鲷是暖水性中上层鱼类, 体侧从吻端到尾鳍有一金黄色纵带, 具有生长速度快、体型大、肉质鲜美等特点, 深受广大消费者的喜爱(廖志强, 2003)。高体鲷分布于全球的亚热带地区, 在中国主要分布于东海和南海。在形态上, 鲷属鱼类种间体色和条纹相似, 基于传统的外部形态特征来准确鉴定物种具有一定难度, 特别是受精卵和苗种在形态上更难区分, 给渔业资源调查物种及种群判别等带来了较大困难。Nugroho 等(2000)通过微卫星 DNA 和 D-Loop 分子标记发现日本高体鲷在遗传上至少分为两个不同的群体; Premachandra 等(2017)通过微卫星标记、基因组测序及线粒体 DNA 等方法认为太平洋海域黄条鲷由于地理隔离分离出的三个群体间不存在基因流动; Iguchi 等(2012)基于线粒体基因组结构特征, 探讨了日本海域的高体鲷、黄条鲷、五条鲷和长鳍鲷四种鲷属鱼类的种间系统进化关系。我国尚未有高体鲷遗传特性方面的研究报道。

本研究测定了东海海域养殖高体鲷的线粒体全基因组序列, 通过与日本海域高体鲷及其他两种鲷属鱼类(黄条鲷、五条鲷)线粒体基因组特征进行比较, 分析三者之间的遗传差异; 通过构建分子系统发育树, 探讨了高体鲷与其他鲷属鱼类的亲缘关系, 为鲷属鱼类的种质鉴定和系统进化提供依据。

## 1 材料与方法

### 1.1 实验材料及 DNA 提取

试验高体鲷样品采集于福建宁德(26°62'N, 119°75'E)海上养殖网箱, 经 MS-222 麻醉后, 对其进行体质量及形态学性状测量, 用无菌解剖剪剪取一部分胸鳍, 放入无水乙醇中保存。取胸鳍组织 100 mg, 按照 DNA 提取试剂盒(OMEGA, USA)使用说明书的操作步骤提取基因组 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

委托青岛欧易生物科技有限公司进行测序, 流程如下: 将提取的 DNA 样品参照 TruSeq™ Nano DNA Sample Prep Kit 试剂盒方法构建文库, 采用 Covaris M220 超声波破碎仪分割成长度为 300~500 bp 左右的片段并在两端加接头, 进行桥式 PCR 扩增, 采用 Illumina NovaSeq 6000 测序技术对样品 DNA 进行

paired-end 测序。

### 1.3 序列组装与分析

使用 SPAdes v3.14.1 (<http://bioinf.spbau.ru/spades>) 软件对 Clean Data 进行校对、排列、拼接后获得高体鲷的线粒体基因组序列。通过 Trimmomatic v0.39 (<http://www.usadellab.org/cms/index.php?page=trimmomatic>) 软件对其修剪后, 使用 MITOS (<http://mitos.bioinf.uni-leipzig.de/index.py>) 软件对线粒体基因组预测编码蛋白、tRNA 和 rRNA 基因, 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。利用 tRNA scan-SE 在线软件(<http://www.Genetics.Wustl.Edu/eddy/tRNAscan-SE>) 进行 tRNA 定位与 RNAstructure 预测, 利用软件 CGView ([http://stothard.afns.ualberta.ca/cgview\\_server/](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、Swiss-Prot、eggNOG、KEGG、GO 注释分析。

通过 MEGA 7.0 软件对高体鲷碱基组成、密码子偏好性及系统进化分析。以日本花鲈(*Lateolabrax japonicus*)为外类群, 从 GenBank 数据库中下载 18 种鲷科鱼类线粒体全基因组序列见表 1, 通过 Clustal W 进行序列多重比对, 自举检验 1 000 次, 构建 Neighbor-joining 进化树,

## 2 结果与分析

### 2.1 线粒体基因组结构特征

本研究获得的高体鲷线粒体全基因组序列经注释后提交 GenBank (Accession No.: MZ398238)。高体鲷线粒体全基因组序列为典型的环状结构, 全长 16 530 bp, 碱基组成为 A (26.83%)、G (17.6%)、C (30.04%)和 T (25.53%); A+T 含量为 52.36%, 显著大于 G+C 含量(47.64%), 表现明显的 AT 偏好性。与其他鱼类线粒体 DNA 结构一致, 共 37 个基因: 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 基因和一个控制区(D-loop 区)。

### 2.2 高体鲷线粒体基因组 rRNA、tRNA 及 D-loop 区结构分析

高体鲷线粒体基因组含有 *12S rRNA* 和 *16S rRNA* 两个 rRNA, *12S rRNA* 位于 *tRNA<sup>Phe</sup>* 和 *tRNA<sup>Val</sup>* 之间, 由 950 个碱基组成; *16S rRNA* 位于 *tRNA<sup>Val</sup>* 和 *tRNA<sup>Leu</sup>* 之间, 由 1 700 个碱基组成, 且 *12S rRNA* 比 *16S rRNA* 更加保守。



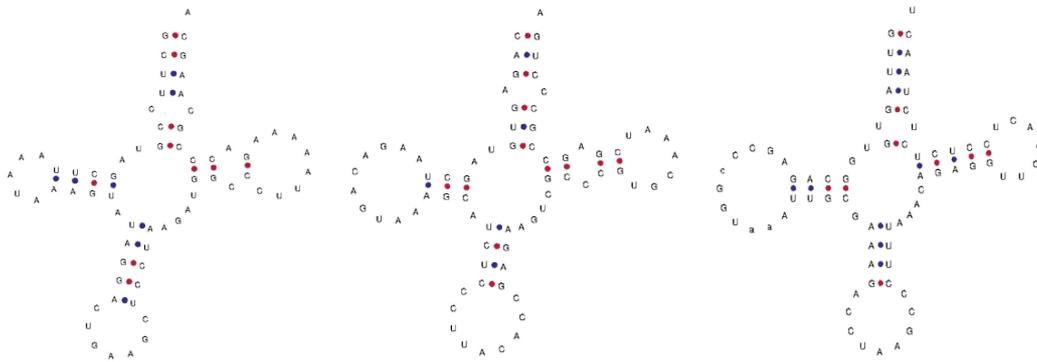
表 2 高体鲷线粒体基因组结构特征  
Tab.2 Characteristics of the mitogenomes of *S. dumerili*

基因名称	起止位置	长度/bp	编码链	间隔长度	起始密码子	终止密码子
<i>tRNA<sup>Phe</sup></i>	1~68	68	H	—	—	—
<i>12S rRNA</i>	69~1 018	950	H	0	—	—
<i>tRNA<sup>Val</sup></i>	1 019~1 090	72	H	0	—	—
<i>16S rRNA</i>	1 109~2 808	1 700	H	18	—	—
<i>tRNA<sup>Leu</sup></i>	2 809~2 883	75	H	0	—	—
<i>ND1</i>	2 884~3 858	975	H	0	ATG	TAA
<i>tRNA<sup>Ile</sup></i>	3 863~3 932	70	H	4	—	—
<i>tRNA<sup>Gln</sup></i>	3 932~4 002	71	L	-1	—	—
<i>tRNA<sup>Met</sup></i>	4 002~4 070	69	H	-1	—	—
<i>ND2</i>	4 071~5 117	1 047	H	0	ATG	TAA
<i>tRNA<sup>Tyr</sup></i>	5 117~5 187	71	H	-1	—	—
<i>tRNA<sup>Ala</sup></i>	5 189~5 257	69	L	1	—	—
<i>tRNA<sup>Asn</sup></i>	5 259~5 331	73	L	1	—	—
<i>tRNA<sup>Cys</sup></i>	5 366~5 431	66	L	34	—	—
<i>tRNA<sup>Tyr</sup></i>	5 432~5 501	70	L	0	—	—
<i>CO I</i>	5 509~7 053	1 545	H	7	ATC	TAA
<i>tRNA<sup>Ser</sup></i>	7 054~7 124	71	L	0	—	—
<i>tRNA<sup>Asp</sup></i>	7 129~7 199	71	H	4	—	—
<i>CO II</i>	7 205~7 912	708	H	5	ATA	TAA
<i>tRNA<sup>Lys</sup></i>	7 899~7 973	75	H	-14	—	—
<i>ATP8</i>	7 975~8 142	168	H	1	ATG	TAA
<i>ATP6</i>	8 133~8 816	684	H	-10	ATG	TAA
<i>CO III</i>	8 816~9 601	786	H	-1	ATG	TAA
<i>tRNA<sup>Gly</sup></i>	9 601~9 671	71	H	-1	—	—
<i>ND3</i>	9 672~10 022	351	H	0	ATG	TAG
<i>tRNA<sup>Arg</sup></i>	10 021~10 089	69	H	-2	—	—
<i>ND4L</i>	10 090~10 386	297	H	0	ATG	TAA
<i>ND4</i>	10 380~11 760	1 381	H	-7	ATG	T—
<i>tRNA<sup>His</sup></i>	11 761~11 829	69	H	0	—	—
<i>tRNA<sup>Ser</sup></i>	11 830~11 896	67	H	0	—	—
<i>tRNA<sup>Leu</sup></i>	11 901~11 973	73	H	4	—	—
<i>ND5</i>	11 992~13 812	1 821	H	18	ATA	TAA
<i>ND6</i>	13 809~14 330	522	L	-4	ATG	TAG
<i>tRNA<sup>Glu</sup></i>	14 332~14 400	69	L	1	—	—
<i>Cytb</i>	14 406~15 546	1 141	H	5	ATG	T--
<i>tRNA<sup>Thr</sup></i>	15 547~15 618	72	H	0	—	—
<i>tRNA<sup>Pro</sup></i>	15 618~15 688	71	L	-1	—	—
<i>D-loop</i>	15 689~16 530	842	H	0	—	—

22 个 tRNA 的序列总长度为 1 552 bp, 长度范围为 66~75 bp, 长度最短的为 *tRNA<sup>Cys</sup>*, 最长的为 *tRNA<sup>Leu</sup>* 和 *tRNA<sup>Lys</sup>*。其中 *tRNA<sup>Ser</sup>* 和 *tRNA<sup>Leu</sup>* 基因各有两个, 除 *tRNA<sup>Gln</sup>*、*tRNA<sup>Ala</sup>*、*tRNA<sup>Asn</sup>*、*tRNA<sup>Cys</sup>*、*tRNA<sup>Tyr</sup>*、*tRNA<sup>Ser</sup>*、*tRNA<sup>Glu</sup>*、*tRNA<sup>Pro</sup>* 位于 L 链外, 其余 14 个

tRNA 皆位于 H 链上。22 个 tRNA 结构均可通过 tRNAscan-SE 在线预测, 除 *tRNA<sup>Ser</sup>*-GCT 基因缺失二氢尿嘧啶臂(DHU 臂)外, 其余 21 个 tRNA 均含有典型的三叶草二级结构(图 2)。

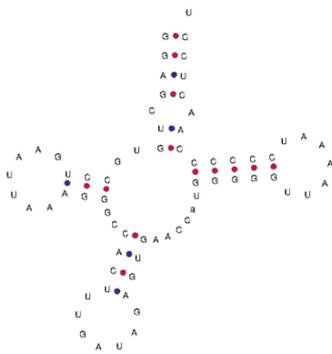
在氨基酸臂中由于 C-T 转换造成的 *tRNA<sup>Val</sup>*、



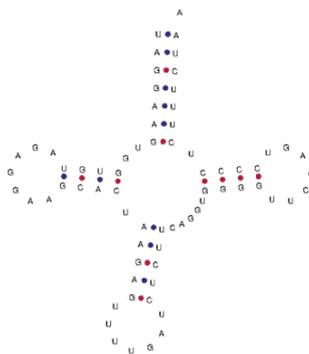
tRNA<sup>Phe</sup>-GAA

tRNA<sup>Val</sup>-TAC

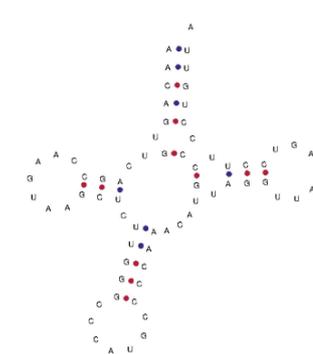
tRNA<sup>Leu</sup>-TAA



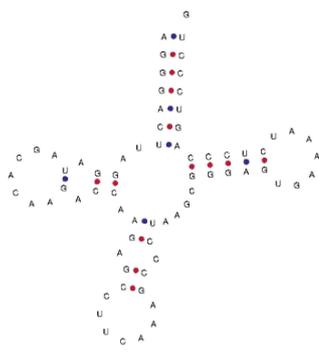
tRNA<sup>Ile</sup>-GAT



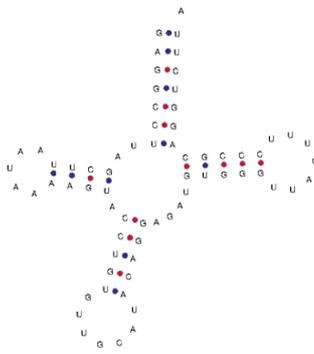
tRNA<sup>Gln</sup>-TTG



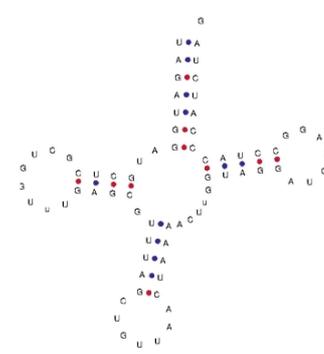
tRNA<sup>Met</sup>-CAT



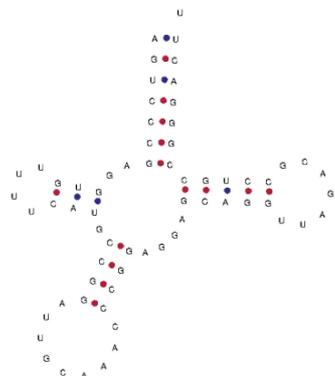
tRNA<sup>Trp</sup>-TCA



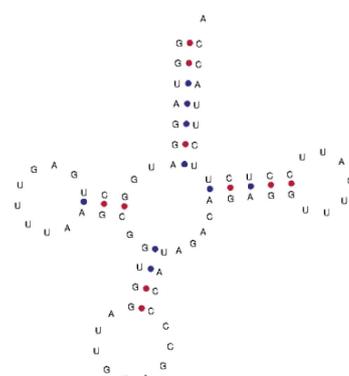
tRNA<sup>Ala</sup>-TGC



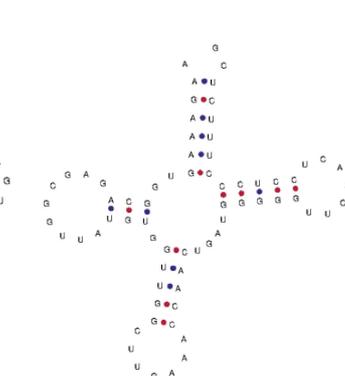
tRNA<sup>Asn</sup>-GTT



tRNA<sup>Cys</sup>-GCA



tRNA<sup>Tyr</sup>-GTA



tRNA<sup>Ser</sup>-TGA

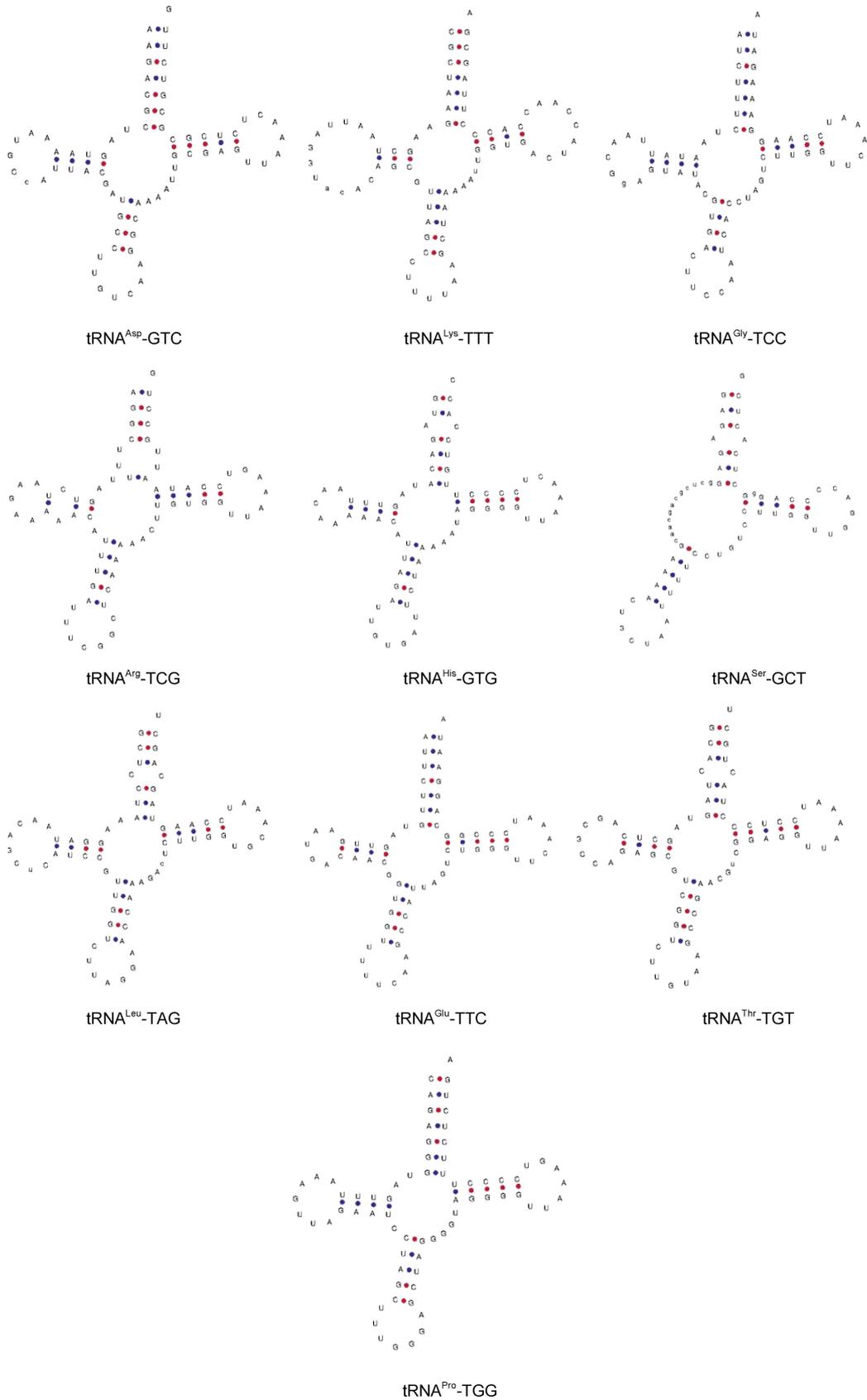


图 2 高体鲮线粒体 tRNA 二级结构  
Fig.2 Secondary structure of *S. dumerili* mitochondrial tRNA

*tRNA<sup>Ile</sup>* 和 *tRNA<sup>His</sup>* 中 A-C 不配对, 还导致 *tRNA<sup>Gln</sup>*、*tRNA<sup>Ala</sup>*、*tRNA<sup>Tyr</sup>* 和 *tRNA<sup>Pro</sup>* 基因中 G-T 发生错配。且在氨基酸臂中 *tRNA<sup>Ser</sup>*-GCT 存在 1 对 A-A 不配对, *tRNA<sup>Phe</sup>*、*tRNA<sup>Leu</sup>*、*tRNA<sup>Thr</sup>* 存在 1 对 C-C 不配对, *tRNA<sup>Arg</sup>* 存在 2 对 T-T 不配对。在反密码子茎中也存在由于 C-T 转换造成的 *tRNA<sup>Trp</sup>* 中 A-C 不配对。在 TΨC 茎上, 同样也存在 C-T 转换现象, 除此之外, *tRNA<sup>Met</sup>* 有 1 对 U-U 不配对。DHU 环也存在 G-T 错配现象, *tRNA<sup>Trp</sup>* 有 1 对 A-A 不配对, 且 *tRNA<sup>Ser</sup>*-GCT 基因该环缺失。

高体鲮控制区(D-loop)位于 *tRNA<sup>Pro</sup>* 和 *tRNA<sup>Phe</sup>* 基因之间, 全长为 842 bp。利用 Tandem repeat finder 软件搜索, 未发现串联重复序列。其碱基含量为(T

31.24%, C 21.97%, A 30.40%, G 16.39%), 该区域 A+T 富含 61.64%; 黄条鲮和五条鲮的 A+T 含量分别为 63.65%和 62.13%, 均表现明显的 A+T 偏好性和抗鸟嘌呤现象。

### 2.3 高体鲮蛋白编码基因密码子偏好性分析

本研究对高体鲮 13 个蛋白编码基因密码子的使用频率及相对同义密码子使用度(RSCU)进行了统计分析(表 3), 共编码 3 808 个密码子, 表中加粗字体为同种氨基酸密码子使用度最多的, 除 UGC、UGG 和 AGA 外, 其 RSCU 均大于 1, 均为偏好性密码子。其中, 以 NNC 类型的密码子占总密码子数的 38.08%, 除 UGC (0.99)外, 其余 NNC 类型的密码子 RSCU 均大于 1, 表明密码子第三位点为 C 的使用频率较高。

表 3 高体鲮 13 个蛋白编码基因密码子偏好性统计  
Tab.3 Codon preference statistics of 13 protein-coding genes of *S. dumerili*

密码子	出现次数	RSCU	密码子	出现次数	RSCU	密码子	出现次数	RSCU	密码子	出现次数	RSCU
UUU(F)	88	0.8	UCU(S)	32	0.7	UAU(Y)	29	0.55	UGU(C)	7	0.4
<b>UUC(F)</b>	<b>154</b>	<b>1.2</b>	<b>UCC(S)</b>	<b>73</b>	<b>2.21</b>	<b>UAC(Y)</b>	<b>81</b>	<b>1.3</b>	<b>UGC(C)</b>	<b>17</b>	<b>0.99</b>
UUA(L)	79	0.77	UCA(S)	50	1.11	UAA(*)	9	0.47	UGA(*)	100	2.44
UUG(L)	25	0.28	UCG(S)	25	0.62	UAG(*)	2	0.08	<b>UGG(W)</b>	<b>20</b>	<b>0.85</b>
CUU(L)	148	1.41	CCU(P)	59	1.3	CAU(H)	25	0.72	CGU(R)	8	0.55
<b>CUC(L)</b>	<b>174</b>	<b>1.51</b>	<b>CCC(P)</b>	<b>103</b>	<b>1.62</b>	<b>CAC(H)</b>	<b>82</b>	<b>1.28</b>	CGC(R)	13	1.13
CUA(L)	169	1.32	CCA(P)	43	0.78	<b>CAA(Q)</b>	<b>77</b>	<b>1.55</b>	<b>CGA(R)</b>	<b>46</b>	<b>3.01</b>
CUG(L)	65	0.72	CCG(P)	19	0.3	CAG(Q)	17	0.3	CGG(R)	12	0.79
<b>AUU(I)</b>	<b>127</b>	<b>1.11</b>	ACU(T)	53	0.91	AAU(N)	28	0.59	AGU(S)	9	0.22
AUC(I)	124	1.1	ACC(T)	105	1.27	<b>AAC(N)</b>	<b>90</b>	<b>1.41</b>	<b>AGC(S)</b>	<b>43</b>	<b>1.15</b>
AUA(I)	77	0.79	<b>ACA(T)</b>	<b>115</b>	<b>1.39</b>	<b>AAA(K)</b>	<b>61</b>	<b>1.28</b>	<b>AGA(R)</b>	<b>1</b>	<b>0.07</b>
<b>AUG(M)</b>	<b>75</b>	<b>1</b>	ACG(T)	23	0.43	AAG(K)	15	0.41	AGG(R)	0	0
GUU(V)	69	1.18	GCU(A)	62	0.71	GAU(D)	17	0.59	GGU(G)	34	0.5
<b>GUC(V)</b>	<b>83</b>	<b>1.45</b>	<b>GCC(A)</b>	<b>166</b>	<b>1.79</b>	<b>GAC(D)</b>	<b>62</b>	<b>1.26</b>	<b>GGC(G)</b>	<b>80</b>	<b>1.16</b>
GUA(V)	56	0.94	GCA(A)	115	1.3	<b>GAA(E)</b>	<b>71</b>	<b>1.46</b>	GGA(G)	73	1.09
GUG(V)	25	0.44	GCG(A)	17	0.21	GAG(E)	29	0.54	GGG(G)	52	1.05

注: \*表示终止密码子; 括号内的字母表示各氨基酸名称的缩写; 表中加粗字体为氨基酸偏好密码子

### 2.4 中国海域高体鲮与日本海域高体鲮蛋白编码基因差异

本研究发现东海海域与日本海域高体鲮在线粒体基因组序列中长度一致, 均为 16 530 bp, 比较分析了两者在蛋白编码基因上的差异(表 4)。除 *CO I*、*CO II* 和 *ND5* 蛋白基因在起止位置、片段长度以及起止密码子选择上存在差异外, 两者在其他 11 个蛋白编码基因中表现出很高的相似性。中国海域高体

鲮 *CO I* 蛋白基因长度比日本海域高体鲮 *CO I* 蛋白基因短 6 bp, 起始密码子分别为 ATG 和 GTG; *CO II* 蛋白编码基因长度比日本高体鲮长 17 bp, 起始密码子分别为 ATA 和 ATG; *ND5* 蛋白编码基因长度比日本海域高体鲮 *ND5* 蛋白基因短 18 bp, 起始密码子分别为 ATA 和 ATG。中国海域高体鲮 *CO II* 蛋白终止密码子为完全的“TAA”, 而日本海域高体鲮 *CO II* 蛋白的终止密码子为不完全的“T”。

表 4 中国海域高体鲷与日本海域高体鲷蛋白编码基因对比  
Tab.4 Protein coding gene comparisons between *S. dumerili* populations from China and Japan

基因	中国海域高体鲷				日本海域高体鲷			
	起止位置(bp)	长度(bp)	起始密码	终止密码	起止位置(bp)	长度(bp)	起始密码	终止密码
<i>ND1</i>	2 884~3 858	975	ATG	TAA	2 884~3 858	975	ATG	TAA
<i>ND2</i>	4 071~5 117	1 047	ATG	TAA	4 071~5 117	1 047	ATG	TAA
<i>CO I</i>	5 509~7 053	1 545	ATC	TAA	5 503~7 053	1 551	GTG	TAA
<i>CO II</i>	7 205~7 912	708	ATA	TAA	7 208~7 898	691	ATG	T--
<i>ATP8</i>	7 975~8 142	168	ATG	TAA	7 975~8 142	168	ATG	TAA
<i>ATP6</i>	8 133~8 816	684	ATG	TAA	8 133~8 816	684	ATG	TAA
<i>CO III</i>	8 816~9 601	786	ATG	TAA	8 816~9 601	786	ATG	TAA
<i>ND3</i>	9 672~10 022	351	ATG	TAG	9 672~10 022	351	ATG	TAG
<i>ND4L</i>	10 090~10 386	297	ATG	TAA	10 090~10 386	297	ATG	TAA
<i>ND4</i>	10 380~11 760	1 381	ATG	T--	10 380~11 760	1 381	ATG	T--
<i>ND5</i>	11 992~13 812	1 821	ATA	TAA	11 974~13 812	1 839	ATG	TAA
<i>ND6</i>	13 809~14 330	522	ATG	TAG	13 809~14 330	522	ATG	TAG
<i>Cytb</i>	14 406~15 546	1 141	ATG	T--	14 406~15 546	1 141	ATG	T--

## 2.5 三种鲷鱼线粒体基因组比较分析

比较分析了三种鲷属鱼类的线粒体基因组全序列(表 5), 长度分别为: 黄条鲷(16 609 bp)、高体鲷(16 530 bp)、五条鲷(16 537 bp)。其基因排列顺序与其他大多数鱼类线粒体基因组一致, 均包括 13 个蛋白质编码基因、22 个 tRNA 基因、2 个 rRNA 基因及非编码区的轻链复制起始区( $O_L$ )以及控制区(D-loop)。其中 *tRNA<sup>Gln</sup>*、*tRNA<sup>Pro</sup>*、*tRNA<sup>Ala</sup>*、*tRNA<sup>Glu</sup>*、*tRNA<sup>Cys</sup>*、*tRNA<sup>Asn</sup>*、*tRNA<sup>Tyr</sup>* 和 *tRNA<sup>Ser(UCN)</sup>* 八个 tRNA 基因与 *ND6* 共 9 个基因由 L 链编码, 其余 28 个基因则是由 H 链编码。黄条鲷、高体鲷和五条鲷线粒体基因组序列的 A+T 含量分

别为 52.05%、52.36%、51.8%, 平均 A+T 含量 52.07%, 明显高于 G+C 含量, 表现出明显的 A+T 偏好性。13 个蛋白编码基因中, 同一编码基因的起止位点上有差异, 基于 *CO I* 基因多以 GTG 为起始密码子, 而黄条鲷和高体鲷分别以 ATG 和 ATC 为起始密码子; 基于 *CO II* 基因, 高体鲷以 ATA 为起始密码子, 而黄条鲷和五条鲷以 ATG; 基于 *ND5* 基因, 黄条鲷和高体鲷以 ATA 为起始密码子, 而五条鲷以 ATG, 其余 10 个编码基因均以 ATG 为起始密码子。其中高体鲷 *CO II* 基因的终止密码子为完全的“TAA”, 黄条鲷和五条鲷 *CO II* 基因的终止密码子为不完全的“T”。

表 5 三种鲷鱼线粒体基因组结构特点  
Tab.5 Characteristics of the mitogenomes of three *Seriola* species

基因	黄条鲷			高体鲷			五条鲷			密码子	
	起始	终止	长度/bp	起始	终止	长度/bp	起始	终止	长度/bp	起始	终止
<i>tRNA<sup>Phe</sup></i>	1	68	68	1	68	68	1	68	68		
<i>12S rRNA</i>	69	1 019	951	69	1 018	950	69	1 023	955		
<i>tRNA<sup>Val</sup></i>	1 020	1 092	72	1 019	1 090	72	1 024	1 095	72		
<i>16S rRNA</i>	1 110	2 808	1 699	1 109	2 808	1 700	1 096	2 813	1 718		
<i>tRNA<sup>Leu(UUR)</sup></i>	2 809	2 883	75	2 809	2 883	75	2 814	2 888	75		
<i>ND1</i>	2 884	3 858	975	2 884	3 858	975	2 889	3 863	975	ATG	TAA
<i>tRNA<sup>Ile</sup></i>	3 863	3 932	70	3 863	3 932	70	3 868	3 937	70		
<i>tRNA<sup>Gln(L)</sup></i>	3 932	4 002	71	3 932	4 002	71	3 937	4 007	71		
<i>tRNA<sup>Met</sup></i>	4 002	4 070	69	4 002	4 070	69	4 007	4 075	69		
<i>ND2</i>	4 071	5 117	1 047	4 071	5 117	1 047	4 076	5 122	1 047	ATG	TAA
<i>tRNA<sup>Trp</sup></i>	5 117	5 187	71	5 117	5 187	71	5 122	5 192	71		
<i>tRNA<sup>Ala(L)</sup></i>	5 189	5 257	69	5 189	5 257	69	5 194	5 262	69		

续表

基因	黄条鲈			高体鲈			五条鲈			密码子	
	起始	终止	长度/bp	起始	终止	长度/bp	起始	终止	长度/bp	起始	终止
<i>tRNA<sup>Asn</sup>(L)</i>	5 259	5 331	73	5 259	5 331	73	5 264	5 336	73		
O <sub>L</sub>	5 332	5 365	33	5 332	5 365	33	5 337	5 370	34		
<i>tRNA<sup>Cys</sup>(L)</i>	5 366	5 431	66	5 366	5 431	66	5 371	5 436	66		
<i>tRNA<sup>Tyr</sup>(L)</i>	5 432	5 501	70	5 432	5 501	70	5 437	5 506	70		
<i>CO I</i>	5 584	7 053	1 470	5 509	7 053	1 545	5 508	7 058	1 551	ATG/ATC/GTG	TAA
<i>tRNA<sup>Ser(UCN)</sup>(L)</i>	7 054	7 124	71	7 054	7 124	71	7 059	7 129	71		
<i>tRNA<sup>Asp</sup></i>	7 129	7 199	71	7 129	7 199	71	7 134	7 204	71		
<i>CO II</i>	7 208	7 898	691	7 205	7 912	708	7 213	7 903	691	ATG/ATA/ATG	T/TAA/T
<i>tRNA<sup>Lys</sup></i>	7 899	7 973	75	7 899	7 973	75	7 904	7 978	75		
<i>ATP8</i>	7 975	8 142	168	7 975	8 142	168	7 980	8 147	168	ATG	TAA
<i>ATP6</i>	8 133	8 816	684	8 133	8 816	684	8 138	8 821	684	TAA	
<i>CO III</i>	8 816	9 601	786	8 816	9 601	786	8 821	9 606	786	ATG	TAA
<i>tRNA<sup>Gly</sup></i>	9 601	9 671	71	9 601	9 671	71	9 606	9 676	71		
<i>ND3</i>	9 672	10 022	351	9 672	10 022	351	9 677	10 027	351	ATG	TAG
<i>tRNA<sup>Arg</sup></i>	10 021	10 089	69	10 021	10 089	69	10 026	10 094	69		
<i>ND4L</i>	10 090	10 386	297	10 090	10 386	297	10 095	10 391	297	ATG	TAA
<i>ND4</i>	10 380	11 760	1 381	10 380	11 760	1 381	10 385	11 765	1 381	ATG	T
<i>tRNA<sup>His</sup></i>	11 761	11 829	69	11 761	11 829	69	11 766	11 834	69		
<i>tRNA<sup>Ser(AGY)</sup></i>	11 830	11 896	67	11 830	11 896	67	11 835	11 901	67		
<i>tRNA<sup>Leu(CUN)</sup></i>	11 901	11 973	73	11 901	11 973	73	11 906	11 978	73		
<i>ND5</i>	11 992	13 812	1 821	11 992	13 812	1 821	11 979	13 817	1 839	ATA/ATA/ATG	TAA
<i>ND6 (L)</i>	13 809	14 330	522	13 809	14 330	522	13 814	14 335	522	ATG	TAG
<i>tRNA<sup>Glu</sup>(L)</i>	14 332	14 400	69	14 332	14 400	69	14 337	14 405	69		
<i>Cytb</i>	14 405	15 545	1 141	14 406	15 546	1 141	14 410	15 550	1 141	ATG	T
<i>tRNA<sup>Thr</sup></i>	15 546	15 617	72	15 547	15 618	72	15 551	15 622	72		
<i>tRNA<sup>Pro</sup>(L)</i>	15 617	15 687	71	15 618	15 688	71	15 622	15 692	71		
<i>Control region</i>	15 689	16 609	921	15 689	16 530	842	15 693	16 537	845		

## 2.6 三种鲈属鱼类蛋白编码基因碱基组成及相似性分析

三种鲈属鱼类的蛋白编码基因的长度分别为：黄条鲈(11 838 bp)、高体鲈(11 426 bp)、五条鲈(11 433 bp)，分别占全序列总长度的 71.3%、69.12%、69.14%。且相邻蛋白编码基因中存在 1~10 bp 的碱基重叠区域。在 *CO I*、*CO II*、*CO III*、*ATP6*、*ATP8* 和 *ND5* 基因中，三种鲈属鱼类 A+T 含量均高于 50% (表 6)，*CO II* 基因中 A+T 平均含量最高(56.74%)。13 个蛋白编码基因的 A+T 总含量分别为黄条鲈(51.09%)、高体鲈(51.72%)、五条鲈(52.04)，可见其在蛋白编码基因组中也具有 AT 偏好性。除了 *ND6* 基因外，三种鲈属鱼类其他蛋白编码基因的 G 含量都在 20% 以下，表现明显的反 G 现象。

对三种鲈属鱼类的线粒体全基因组 13 个蛋白编码基因序列进行两两比对，结果表明三种鲈属鱼类的蛋白编码基因的相似性在 85%~100% 之间(图 3)。三者之间在 *ATP6* 基因中相似性最低，但也在 85% 以上。中国和日本高体鲈在 *ND2* 和 *ATP8*、*ATP6*、*ND3*、*ND6* 和 *Cytb* 六个基因上的相似性高达 100%；中国和日本黄条鲈在 *ND1*、*CO II*、*ATP8* 和 *ND4L* 基因上的相似性也是 100%。黄条鲈种群中除了日本黄条鲈的 *Cytb* 基因与五条鲈相似性低于 90%，其他所有蛋白编码基因相似性均在 90% 以上，说明黄条鲈和五条鲈的亲缘关系最近。由于差异性不显著，这些相似性 100% 的蛋白编码基因不适用于不同地理区域同种物种的鉴别，但所有蛋白编码基因对于三种鲈属鱼类都能有效区分。

表 6 三种鲷鱼蛋白编码基因碱基组成  
Tab.6 Nucleotide composition of protein-coding genes in three *Seriola* species

种类	碱基	ND1	ND2	CO I	CO II	ATP8	ATP6	CO III	ND3	ND4L	ND4	ND5	ND6	Cytb
黄条鲷 <i>S. aureovittata</i>	A+T	48.10	48.62	52.24	57.16	52.98	52.63	50.89	50.14	48.82	50.62	52.39	49.43	49.78
	T	26.36	24.83	28.30	28.08	23.81	27.63	26.97	29.63	30.64	24.55	26.03	36.02	26.73
	C	33.44	36.77	28.84	26.63	32.74	32.46	31.04	32.48	35.02	32.95	32.24	16.86	34.36
	A	21.74	23.78	23.95	29.09	29.17	25.00	23.92	20.51	18.18	26.07	26.36	13.41	23.05
	G	18.46	14.61	18.91	16.21	14.29	14.91	18.07	17.38	16.16	16.44	15.38	33.72	15.86
高体鲷 <i>S. dumerili</i>	A+T	49.95	50.15	52.16	56.77	51.78	51.61	51.40	49.29	49.50	51.48	52.83	51.72	51.00
	T	27.18	25.60	28.41	27.82	22.02	27.78	27.23	27.35	30.64	25.63	26.03	36.97	26.99
	C	32.92	35.05	29.19	26.84	34.52	32.75	30.92	34.47	35.35	32.37	31.96	15.33	33.39
	A	22.77	24.55	23.75	28.95	29.76	23.83	24.17	21.94	18.86	25.85	26.80	14.75	24.01
	G	17.13	14.80	18.64	16.38	13.69	15.64	17.68	16.24	15.15	16.15	15.21	32.95	15.60
五条鲷 <i>S. quinqueradiata</i>	A+T	48.51	47.27	52.22	56.29	51.19	50.81	51.53	50.99	48.82	49.46	53.12	50.19	49.96
	T	26.77	24.16	28.69	27.35	23.21	27.01	27.10	29.34	30.64	24.48	26.26	13.41	26.73
	C	32.72	37.15	28.24	27.35	33.33	33.43	30.92	32.76	34.68	33.02	31.65	33.91	33.83
	A	21.74	23.11	23.53	28.94	27.98	23.80	24.43	21.65	18.18	24.98	26.86	36.78	23.23
	G	18.77	15.57	19.54	16.35	15.48	15.77	17.56	16.24	16.50	17.52	15.23	15.90	16.21

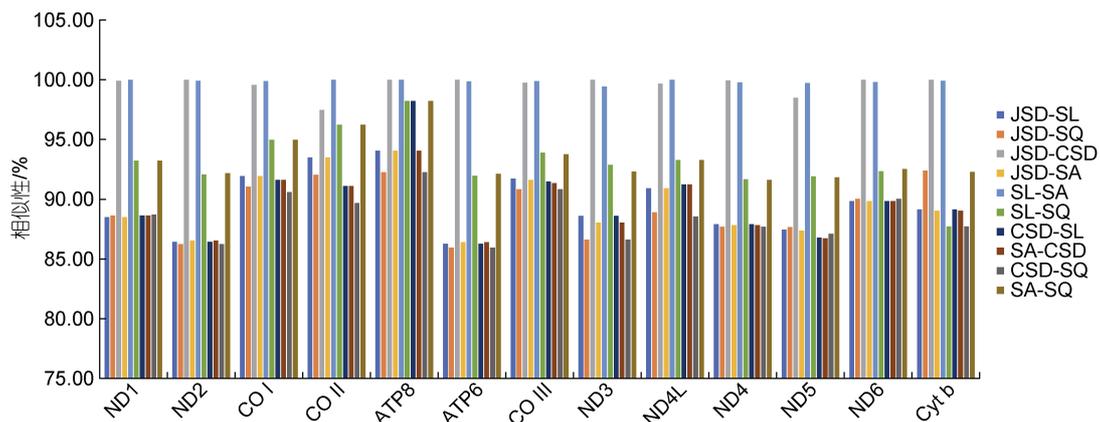


图 3 三种鲷属鱼类的 13 个蛋白编码基因之间的相似性比较

Fig.3 Similarity of 13 protein-coding genes of the three *Seriola* species

注: CSD: 中国高体鲷; JSD: 日本高体鲷; SA: 中国黄条鲷; SL: 日本黄条鲷; SQ: 五条鲷

## 2.7 系统进化分析

以日本花鲈为外类群, 构建了高体鲷与其他鲷科鱼类的系统发育树, 由图 4 可知, 每个鱼种都独立分支, 纺锤鲷属、小条鲷属和鲷属鱼类聚为单细分支, 且中国与日本高体鲷聚为一个独立小分支, 表明其具有最近的亲缘关系。在鲷属鱼类中, 分化时间上高体鲷略早于黄条鲷和五条鲷, 且高体鲷与长鳍鲷聚为一支, 亲缘关系最近; 黄条鲷和五条鲷同属一支。

## 3 讨论

本研究测定了高体鲷的线粒体基因组全序列,

全长为 16 530 bp, 其碱基组成为(T: 25.53%, C: 30.04%, A: 26.83%, G: 17.60%), GC 含量约 47.64%, 呈现出明显的 AT 偏好性, 这与脊椎动物碱基组成相似(张方等, 1998)。其中, 高体鲷G 碱基含量与其他硬骨鱼类如黄条鲷(17.84%) (史宝等, 2019)、斑鱧(*Sillago aeolus*) (18.75%) (肖家光, 2015)、西里伯斯青鳉(*Oryzias celebensis*) (17.60%) (马江茹, 2020)等含量相似, 表现出显著的抗鸟嘌呤现象。本研究基于高体鲷线粒体基因组的基础上, 对其基因组结构特征、碱基组成、密码子偏好性及蛋白编码基因等做了系统分析。结果表明, 在高体鲷线粒体基因组中, 13 个蛋白

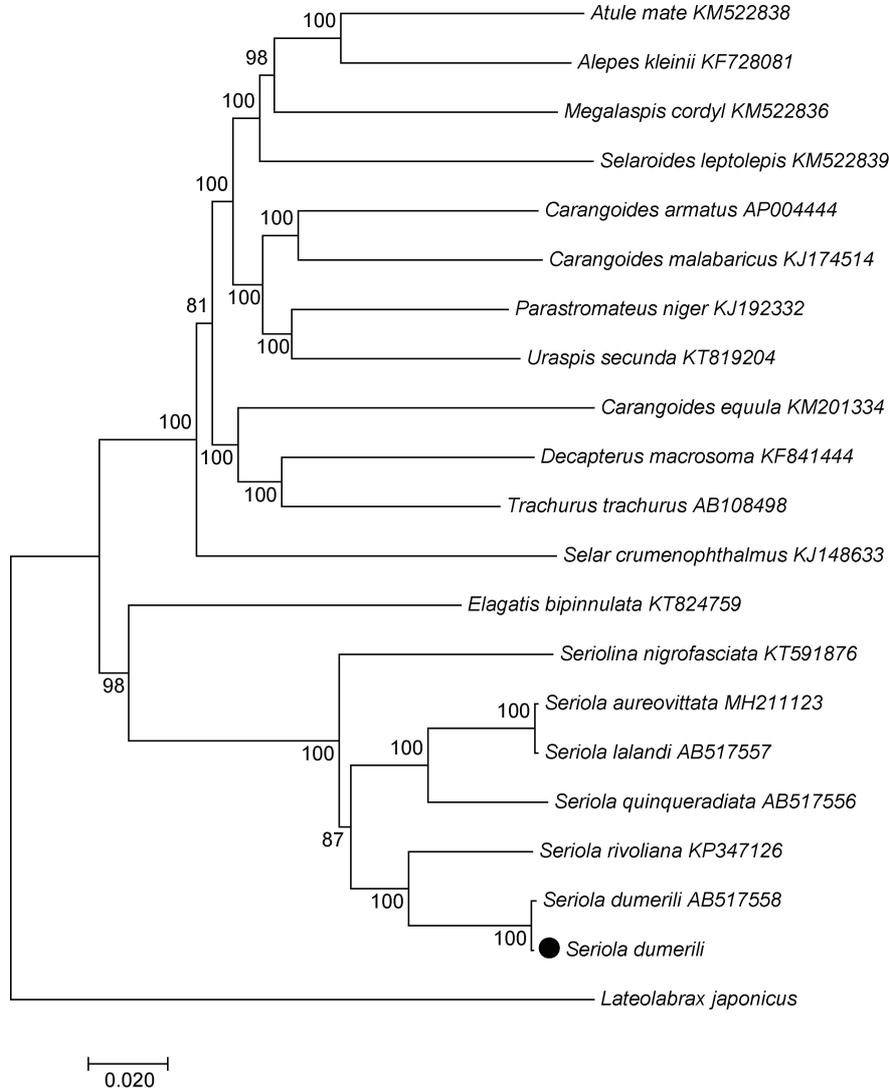


图 4 基于线粒体基因组全序列构建的 NJ 系统进化树

Fig.4 Molecular phylogenetic tree constructed by NJ method based on mitochondrial genome

编码基因除了 *ND6* 外, 其余均在 H 链上。基因组中重叠的片段在鱼类中一般只有 7~10 bp, 而在哺乳动物中一般可达 40~46 bp (Broughton *et al.*, 2001; Zhu *et al.*, 2013)。本研究发现高体鲷重叠片段长度为 1~14 bp, 其中, 蛋白编码基因 *CO II* 与 *tRNA<sup>Lys</sup>* 之间的重叠片段最大 (14 bp), 在 *tRNA<sup>Ile</sup>-tRNA<sup>Gln</sup>*、*tRNA<sup>Gln</sup>-tRNA<sup>Met</sup>*、*ND2-tRNA<sup>Trp</sup>*、*CO III-tRNA<sup>Gly</sup>* 和 *tRNA<sup>Thr</sup>-tRNA<sup>Pro</sup>* 之间重叠片段最小 (1 bp)。对于高体鲷密码子偏好性分析显示, 以 NNC 类型的密码子使用频率最高, 与其蛋白编码基因组成中碱基 C 偏好性一致, 与条石鲷 (*Oplegnathus fasciatus*) (孟乾等, 2020) 结果类似; 但在鲂属 (*Megalobrama*) (赖瑞芳等, 2014)、鲚亚科 (*Acheilognathinae*) (王尚洪, 2015) 等鱼类中以 NNA 类型密码子使用频率较高。可见, 不同鱼种在密码子偏

好性选择中存在差异。

本研究通过对中国海域高体鲷线粒体基因组的 37 个基因的起止位置、长度及蛋白编码基因等特征分析显示与 Iguchi 等 (2012) 报道的日本海域高体鲷线粒体基因组特征高度相似。但两者在 rRNA 与蛋白编码基因上存在差异。在 rRNA 中, 中国海域高体鲷 *16S rRNA* 较日本海域高体鲷 *16S rRNA* 基因长 18 bp; 蛋白编码基因中, 中国海域高体鲷的 *CO I* 蛋白基因与 *ND5* 蛋白基因皆略短于日本海域高体鲷, 而中国海域高体鲷的 *CO II* 蛋白基因略长于日本海域高体鲷, 可能由于中国和日本高体鲷存在一定的地理种群隔离而表现出的位点多态性。

此外, 本研究对三种鲷属鱼类线粒体基因组特征、蛋白编码基因及其相似性进行比较发现, 高体

鲷在 *CO II* 基因中以完全的“TAA”为终止密码子, 黄条鲷和五条鲷在 *CO II* 基因中的终止密码子为不完全的“T”, 而其余的蛋白质编码基因则是以常见的 TAA 和 TAG 作为终止密码子。*tRNA<sup>Asn</sup>* 和 *tRNA<sup>Cys</sup>* 之间的轻链复制起始区( $O_L$ )在 3 种鲷属鱼类中是一个 34 bp 的 DNA 片段, 控制区是一个变异累积多, 进化速度最快的基因之一, 本研究中的三种鲷属鱼类, 其 D-loop 区 A+T 含量都相对比较高, 该区域调控线粒体的复制和转录(郭新红等, 2004)。在 tRNA 结构中, *tRNA<sup>Ser</sup>*-GCT 基因缺失二氢尿嘧啶臂(DHU 臂)不能形成三叶草结构, 这在斑石鲷(*Oplegnathus punctatus*) (孟乾等, 2020)等鱼类中也有类似报道。每个蛋白编码基因在进化速率上是不同的, Zardoya 等(1996)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*、*Cytb* 和 *ND5* 这 5 个基因作为好, *CO II*、*CO III*、*ND1* 和 *ND6* 作为中等, 而 *ATP6*、*ATP8*、*ND3* 和 *ND4L* 则作为差的一组。本研究通过对三种鲷属鱼类蛋白编码基因进行比较分析, *Cytb* 和 *ND* 基因均可用于鲷属鱼类分子标记, 还可筛选合适蛋白编码基因用作不同地理种群划分, 为鲷属鱼类的物种鉴定和系统进化分析提供依据。

本研究基于线粒体基因组全序列通过邻接法构建了部分鲷科鱼类系统发育树, 鲷属鱼类与小条鲷属先聚为一类, 再与其他鲷科鱼类聚为一大支, 这与郑文娟等(2008)采用单个基因 *16S rRNA* 分类结果一致, 在分化时间上高体鲷要晚于黄条鲷, 且黄条鲷与五条鲷亲缘关系最近, 同属于一个细支。高体鲷和长鳍鲷同属一个分支, 亲缘关系最近。

#### 4 结论

本研究基于线粒体基因组结构特征比较, 推测中国海域高体鲷与日本海域高体鲷可能属于同一种群, 但可能存在一定的地理隔离。结合形态学与生物信息学成功将高体鲷与其他鲷属鱼类属准确区分, 为鲷属鱼类物种鉴别、种群划分和种质资源保护及可持续利用提供技术支撑, 助力我国鲷鱼养殖产业的持续健康发展。

#### 参 考 文 献

马江茹, 2020. 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: 29-30.  
王尚洪, 2015. 鲷亚科三种鱼类线粒体全基因组测定及其比较基因组学分析[D]. 南昌: 南昌大学: 54-76.  
史宝, 柳学周, 刘永山, 等, 2019. 黄条鲷线粒体全基因组测

序及结构特征分析[J]. 中国水产科学, 26(3): 405-415.  
肖家光, 2015. 基于线粒体基因组全序列的鳊属鱼类系统发育研究[D]. 青岛: 中国海洋大学: 24-25.  
张方, 米志勇, 1998. 动物线粒体 DNA 的分子生物学研究进展[J]. 生物工程进展, 18(3): 25-31.  
陈大刚, 张美昭, 2015. 中国海洋鱼类[M]. 青岛: 中国海洋大学出版社: 1096-1097.  
陈四海, 区又君, 李加儿, 2011. 鱼类线粒体 DNA 及其研究进展[J]. 生物技术通报(3): 13-20.  
周传江, 马爱磊, 汪曦, 等, 2019. 鱼类线粒体基因组研究进展[J]. 河南师范大学学报(自然科学版), 47(2): 74-82.  
郑文娟, 朱世华, 邹记兴, 等, 2008. 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, 32(6): 847-854.  
孟乾, 张志勇, 张志伟, 等, 2020. 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, 39(5): 702-709.  
郭新红, 刘少军, 刘巧, 等, 2004. 鱼类线粒体 DNA 研究新进展[J]. 遗传学报, 31(9): 983-1000.  
黄小林, 杨育凯, 林黑着, 等, 2018. 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, 35(5): 33-36.  
程佩琳, 俞丹, 刘焕章, 等, 2021. 基于线粒体基因组全序列的鲟形目鱼类(Pisces: Acipenseriformes)的分子系统发育重建[J]. 水生生物学报, 45(3): 487-494.  
赖瑞芳, 张秀杰, 李艳和, 等, 2014. 鲂属鱼类线粒体基因组的比较及其系统发育分析[J]. 水产学报, 38(1): 1-14.  
廖志强, 2003. 高体鲷网箱养殖技术[J]. 中国水产(12): 60-61.  
BERNAL-RAMÍREZ J H, ADCOCK G J, HAUSER L, et al, 2003. Temporal stability of genetic population structure in the New Zealand snapper, *Pagrus auratus*, and relationship to coastal currents [J]. Marine Biology, 142(3): 567-574.  
BROUGHTON R E, MILAM J E, ROE B A, 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA [J]. Genome Research, 11(11): 1958-1967.  
CUROLE J P, KOCHER T D, 1999. Mitogenomics: digging deeper with complete mitochondrial genomes [J]. Trends in Ecology & Evolution, 14(10): 394-398.  
IGUCHI J, TAKASHIMA Y, NAMIKOSHI A, et al, 2012. Species identification method for marine products of *Seriola* and related species [J]. Fisheries Science, 78(1): 197-206.  
NUGROHO E, TANIGUCHI N, KATO K, et al, 2000. Genetic difference among seed populations of greater amberjack used in aquaculture farm of Japan [J]. Aquaculture Science, 48(4): 665-674.  
PREMACHANDRA H K A, LAFARGA-DE LA CRUZ F, TAKEUCHI Y, et al, 2017. Genomic DNA variation confirmed *Seriola lalandi* comprises three different populations in the Pacific, but with recent divergence [J]. Scientific Reports, 7(1): 9386.  
SATO T P, MIYA M, MABUCHI K, et al, 2016. Structure and variation of the mitochondrial genome of fishes [J]. BMC Genomics, 17(1): 719.  
ZARDOYA R, MEYER A, 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving

relationships among vertebrates [J]. *Molecular Biology and Evolution*, 13(7): 933-942.  
ZHU Y X, CHEN Y, CHENG Q Q, *et al*, 2013. The complete

mitochondrial genome sequence of *Schizothorax macropogon* (Cypriniformes: Cyprinidae) [J]. *Mitochondrial DNA*, 24(3): 237-239.

## COMPLETE SEQUENCE AND GENE ORGANIZATION OF THE MITOCHONDRIAL GENOME OF *SERIOLA DUMERILI*

WANG Kai-Jie<sup>1,2,3</sup>, XU Yong-Jiang<sup>1,2</sup>, CUI Ai-Jun<sup>1,2</sup>, LIU Xue-Zhou<sup>1,2</sup>, JIANG Yan<sup>1,2</sup>, WANG Bin<sup>1,2</sup>  
(1. *Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China*; 2. *Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China*; 3. *National Engineering Research Center For Marine Aquaculture, Zhejiang Ocean University, Zhoushan 316022, China*)

**Abstract** To identify the species of *Seriola* and strengthen the management of genetic diversity and protection of germplasm resources of the *Seriola*, the complete mitochondrial genome sequence of *Seriola dumerili* cultured in the East China Sea was successfully obtained by second-generation sequencing and bioinformatic analysis. The total length of the mitochondrial genome sequence was 16 530 bp with typical mtDNA components including 13 proteins and 22 tRNAs, 2 rRNAs and a D-Loop region, in which the A content was 26.83%, G was 17.6%, C was 30.04% and T was 25.53%. Except for *ND6*, *tRNA<sup>Gln</sup>*, *tRNA<sup>Ala</sup>*, *tRNA<sup>Asn</sup>*, *tRNA<sup>Cys</sup>*, *tRNA<sup>Tyr</sup>*, *tRNA<sup>Ser</sup>*, *tRNA<sup>Glu</sup>*, and *tRNA<sup>Pro</sup>*, all the other genes were encoded on the H-strand. The A+T content of the complete mitochondrial genome sequence and the protein coding gene were 52.36% and 51.72%, respectively. Among the 22 tRNA genes, except for *tRNA<sup>Ser</sup>*-GCT, the other 21 tRNA genes contained typical secondary structure of clover. The results show that the similarity of 13 protein coding genes in the whole mitochondrial genome of the three species was over 85% to 100%. The genes encoding *CO I*, *CO II*, and *ND5* showed difference at the start point, length, and starting codon between Chinese and Japanese *S. dumerili*. Furthermore, the phylogenetic tree was constructed with the complete mitochondrial genome sequences of four species from *Seriola* genus. The phylogenetic tree showed that *S. aureovittata* and *S. quinqueradiata* were clustered in one branch, and the phylogenetic tree of *S. dumerili* and *S. rivoliana* belonged to the same proximal branch. The phylogenetic relationship between *S. aureovittata* and *S. quinqueradiata*, and between *S. dumerili* and *S. rivoliana* are the closest.

**Key words** *Seriola dumerili*; mitochondrial genome; sequence analysis; phylogenetic tree