

银色裂腹鱼线粒体基因组序列特征与系统进化分析

刘玉萍¹, 胡建勇², 宁子君¹, 肖佩怡¹, 杨天燕^{1*}

1 浙江海洋大学水产学院, 浙江 舟山 316022

2 新疆维吾尔自治区水产科学研究所, 新疆 乌鲁木齐 830000

刘玉萍, 胡建勇, 宁子君, 肖佩怡, 杨天燕. 银色裂腹鱼线粒体基因组序列特征与系统进化分析[J]. 生物工程学报, 2023, 39(7): 2965-2985.

LIU Yuping, HU Jianyong, NING Zijun, XIAO Peiyi, YANG Tianyan. Mitochondrial genome sequence characteristics and phylogenetic analysis of *Schizothorax argentatus*[J]. Chinese Journal of Biotechnology, 2023, 39(7): 2965-2985.

摘要: 银色裂腹鱼(*Schizothorax argentatus*)在我国仅分布于新疆地区的伊犁河流域, 是我国裂腹鱼类中珍稀濒危品种之一, 具有较高的科研和经济价值。本研究采用高通量测序技术获得了银色裂腹鱼长度为 16 580 bp 的线粒体基因组全序列, 其基因组成和排列顺序均与典型的脊椎动物相似, 共有 13 个蛋白质编码基因、22 个 tRNA 基因、2 个 rRNA 基因和 1 个非编码区(D-loop)。碱基组成分别为 A (30.25%)、G (17.28%)、C (27.20%)和 T (25.27%), 呈现明显的 AT 偏好性和反 G 偏倚。tRNA 基因中仅 tRNA-Ser^(GCU)因缺少二氢尿嘧啶茎而无法形成典型的三叶草结构。ND6 基因的 AT-skew 和 GC-skew 值波动最大, 揭示该基因经历的选择和突变压力可能与其他基因不同。银色裂腹鱼线粒体控制区包含了 3 个不同的结构域: 终止序列区(ETAS)、中央保守区(CSB-F、CSB-E、CSB-D 和 CSB-B)和保守序列区(CSB1、CSB2 和 CSB3), 且在 CSB3 下游约 50 bp 处识别到鲤形目(Cypriniformes)鱼类中普遍存在的保守序列片段 TT (AT) nGTG。基于 28 种裂腹鱼属鱼类线粒体基因组全序列构建的系统发育关系表明银色裂腹鱼分化时间较早, 与其他类群亲缘关系较远, 这可能与其所生活的水域地理位置和水文环境有密切关系。

关键词: 银色裂腹鱼; 线粒体基因组; 序列分析; 系统进化

资助项目: 新疆土著鱼类种质工程技术研究中心补助资金(202210229); 浙江省大学生科技创新活动计划暨新苗人才计划(2021R411008)

This work was supported by the Subsidy Fund for Xinjiang Indigenous Fish Germplasm Engineering Technology Research Center (202210229) and by the Science and Technology Innovation Project of College Students in Zhejiang Province (2021R411008).

*Corresponding author. E-mail: hellojelly1130@163.com

Received: 2022-09-28; Accepted: 2023-02-08

Mitochondrial genome sequence characteristics and phylogenetic analysis of *Schizothorax argenteus*

LIU Yuping¹, HU Jianyong², NING Zijun¹, XIAO Peiyi¹, YANG Tianyan^{1*}

¹ Fisheries College, Zhejiang Ocean University, Zhoushan 316022, Zhejiang, China

² Xinjiang Fishery Research Institute, Urumqi 830000, Xinjiang, China

Abstract: *Schizothorax argenteus* that only distributes in the Ili River basin in Xinjiang is one of the rare and endangered species of schizothorax in China, thus has high scientific and economic values. In this study, the complete mitochondrial genome sequence of *S. argenteus* with a length of 16 580 bp was obtained by high-throughput sequencing. The gene compositions and arrangement were similar to those of typical vertebrates. It contained 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes, and a non-coding region (D-loop). The nucleotide compositions were A (30.25%), G (17.28%), C (27.20%), and T (25.27%), respectively, showing obvious AT bias and anti-G bias. Among the tRNA genes, only tRNA-Ser^(GCU) could not form a typical cloverleaf structure due to the lack of dihydrouracil arm. The AT-skew and GC-skew values of the *ND6* gene were fluctuating the most, suggesting that the gene may experience different selection and mutation pressures from other genes. The mitochondrial control region of *S. argenteus* contained three different domains, i.e., termination sequence region (ETAS), central conserved region (CSB-F, CSB-E, CSB-D, and CSB-B), and conserved sequence region (CSB1, CSB2, and CSB3). The conserved sequence fragment TT (AT) nGTG, which was ubiquitous in Cypriniformes, was identified at about 50 bp downstream CSB3. Phylogenetic relationships based on the complete mitochondrial genome sequence of 28 *Schizothorax* species showed that *S. argenteus* had differentiated earlier and had a distant relationship with other species, which may be closely related to the geographical location and the hydrological environment where it lives.

Keywords: *Schizothorax argenteus*; mitochondrial genome; sequence analysis; phylogenetics

裂腹鱼类是在第三纪晚期青藏高原阶段性隆起抬升和气候变化过程中,逐渐适应高寒冷水环境的鲤科鱼类中为数不多的类群,这些鱼类呈现出以青藏高原为中心,由高原边缘向高原腹地辐射分布的格局,因臀鳍至肛门两侧各有一列特化的臀鳞,形似一条裂缝而得名^[1-2]。统计资料表明,我国新疆现有裂腹鱼亚科(*Schizothoracinae*)鱼类5属11种,占我国裂腹鱼总数的15%,占新疆土著鱼类的24%^[2-3]。银色裂腹鱼(*Schizothorax argenteus* Kessler, 1874),是仅分布于中亚地区

的巴尔喀什湖-阿拉湖水系及我国新疆伊犁河流域的一种珍稀濒危鱼类,也是新疆特有的7种裂腹鱼类之一^[3-4]。根据李思忠^[5]对我国淡水鱼类区系组成的划分标准,银色裂腹鱼属于中亚高山复合体类型,常栖息于砾石底质的山涧冷水溪流中,能够适应高海拔、严寒和水流急湍的恶劣环境。因繁殖期间该鱼类从哈萨克斯坦的巴尔喀什湖洄游至我国伊犁河流域产卵,20世纪60年代前曾是伊犁河主要捕捞对象,可占渔获物60%以上^[6]。20世纪70年代起,由于哈萨克斯坦境

内卡甫恰盖水库的修建以及伊犁河外来鱼类的人工移植,银色裂腹鱼野生资源锐减至濒临灭绝的境地^[6-7],亟待抢救性开展其种质资源的保护工作。目前,围绕伊犁河流域的其他几种珍稀裂腹鱼类,如伊犁裂腹鱼(*Schizothorax pseudoaksaiensis*)、斑重唇鱼(*Diptychus maculatus*)和新疆裸重唇鱼(*Gymnodiptychus dybowskii*)已开展大量保护生物学和分子生物学研究工作,而由于银色裂腹鱼资源濒临绝境,其研究仅集中于物种分布和形态描述领域^[4,6-7],遗传学背景资料十分匮乏。

线粒体 DNA (mitochondrial DNA, mtDNA) 是长度为 15–20 kb 的成闭环状的独立于核基因组的“小基因组”,由于其呈母系遗传、结构简单、能独立复制、具有较高的突变率且变异发生的几率相对稳定等特点被广泛应用于物种演化、亲缘关系和群体遗传学的研究,已成为解决近缘种间及种内群体间遗传分化关系的有力工具^[8-10]。近年来, DNA 测序技术的进步为快速准确地获取鱼类线粒体基因组数据提供了极大的便捷,越来越多的裂腹鱼类线粒体 DNA 全序列被相继报道,也为进一步开展该类群的分类和系统进化研究提供了丰富的资料。本研究采用高通量测序技术对采自我国新疆伊犁河水系的银色裂腹鱼线粒体基因组全序列进行了测定和结构特征分析,通过与 28 种近缘种鱼类线粒体基因组进行对比,进一步探讨了该类群鱼类的分子系统发生关系,研究结果以期对银色裂腹鱼的种质鉴定和遗传多样性保护提供基础信息。

1 材料与方法

1.1 实验材料和 DNA 提取

本研究银色裂腹鱼样品于 2017 年 9 月采自我国新疆维吾尔自治区伊犁河支流喀什河(43°72'N, 81°85'E),剪取部分背部新鲜肌肉组织于 1.5 mL EP 离心管内,浸泡于 95%酒精固定

48 h 后,置于 -20 °C 冰箱冷冻保存备用。采用传统的氯仿-Tris 饱和酚方法提取基因组 DNA^[11],使用 NanoDrop ND-1 000 超微量分光光度计测定基因组 DNA 浓度,并使用 1%琼脂糖凝胶电泳检测 DNA 完整性。

1.2 高通量测序和基因注释

质检合格的 DNA 样品采用 Covaris 超声波破碎仪随机打断成长度为 300 bp 左右的片段,构建银色裂腹鱼小片段基因组 DNA 文库,委托深圳市恒创基因科技有限公司基于 HiSeq 2500 高通量测序平台进行测序。利用 SOAPdenovo 2.04 软件(<http://soap.genomics.org.cn/soapdenovo.html>)^[12]对测得的 clean reads 进行组装,并基于 reads 的双末端(paired-end)和重叠(overlap)关系实现局部组装结果的优化;使用 GapCloser 1.12 软件(<http://soap.genomics.org.cn/soapdenovo.html>)^[13]将拼接 scaffold 过程中引入的 gap 进行弥补和修复,最后去除冗余的段序列得到最后的组装结果。使用 NOVOPlasty 软件^[14]对线粒体基因组进行提取和组装,将拼接得到的完整线粒体基因组序列上传至 MITOS 网页服务器(<http://mitofish.aori.u-tokyo.ac.jp/>)进行蛋白质编码基因、RNA 和非编码区的注释,并辅以人工比对校正最终确定每个基因的位置和长度信息。

1.3 高通量测序结果的验证

为了进一步验证高通量测序数据的准确性,以组装获得的银色裂腹鱼和近缘物种线粒体全序列为模板,使用 Primer Premier 6.0 软件设计覆盖银色裂腹鱼线粒体全基因组的 11 对 PCR 扩增引物如表 1 所示。所有引物均由生工生物工程(上海)股份有限公司合成。

25 μL 的 PCR 反应体系为: TaKaRa LA Taq (5 U/μL) 0.25 μL, 10×LA PCR Buffer II (Mg²⁺ Plus) 2.5 μL, dNTPs (2.5 mmol/L) 4 μL, 10 μmol/L 的引物各 1 μL, 50 ng/μL 的 DNA 模板 1 μL,

表 1 银色裂腹鱼线粒体基因组的 PCR 引物信息

Table 1 Primers designed according to the mitochondrial genome sequence of *Schizothorax argentatus*

Primer name	Primer sequence	Primer location (bp)	Length (bp)	Annealing temperature (°C)
1F	CAAGGGAATTCAGCAGTG	227-244	2 115	52
1R	TAGGAAGTTTTCCAGTC	2 360-2 377		
2F	GCAGAAGCGGGTATAATA	2 200-2 217	2 447	50
2R	GCGGAGGTCATAATAATG	4 665-4 682		
3F	AACTTGCCCCATTCGCAC	4 445-4 462	2 181	51
3R	TCAGTGCACAAAGGCTGC	6 644-6 661		
4F	TCCTATTACAGTGGGAG	6 510-6 527	545	50
4R	AGTGACAGAGTGGTTATG	7 073-7 090		
5F	AATCAAATTAACGAGAAA	7 014-7 031	2 476	57
5R	CGAAGTGTCAGTATCAGG	9 508-9 525		
6F	ACATCTGAACACCATTTT	9 474-9 491	2 225	55
6R	GCACCACCTCATATGAG	11 717-11 734		
7F	GACTTCAACCATTCCACA	11 634-11 651	1 351	59
7R	TTGCTCGTCGTTTAGGCT	13 003-13 020		
8F	CCTTAACCCAGAATGATA	12 833-12 850	814	55
8R	TTTGCCATGGTCATTTGT	13 665-13 682		
9F	ACTCAAACCTCACCTAGG	13 578-13 595	1 789	53
9R	ATGTTCTACTGGCATTCC	15 385-15 404		
10F	ATTCTCTCCAACCTGCT	15 099-15 116	1 066	56
10R	AGTGTTTCGACCTTGAATA	16 183-16 200		
11F	TGTCTCATTACCCACAT	16 085-16 102	1 041	50
11R	TAAGCTGATGCTCGTAGT	564-581		

ddH₂O 15.25 μL。PCR 反应程序为：94 °C 预变性 5 min；94 °C 变性 45 s，50 °C-59 °C 退火 45 s-1 min (不同引物选用的退火温度不同)，72 °C 延伸 1 min，共 35 个循环；72 °C 延伸 10 min。最后保存温度设置为 4 °C。获得的 PCR 目的片段经浓度为 1.5% 的琼脂糖凝胶电泳检测后，使用 Tanon 2500 系列全自动凝胶图像分析系统拍照，合格的样品送至生工生物工程(上海)股份有限公司进行双向测序。

1.4 数据分析

采用在线软件 tRNAscan-SE (<http://trna.ucsc.edu/tRNAscan-SE/>)^[15] 预测 tRNA 基因的二级结构；利用 RNACentral 网站(<https://rnacentral.org/>)^[16] 的 Visualise RNA 2D Structure (R2DT) 工具绘制 rRNA 二级结构；用 RNAfold 软件(<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) 计算 rRNA 基因

的自由能值；采用 MEGA-X 软件^[17] 分析碱基组成和密码子使用情况，计算 AT-skew 和 GC-skew 值并运用 Microsoft Excel 软件进行构图分析。从 GenBank (<https://www.ncbi.nlm.nih.gov/>) 数据库检索并下载 28 种裂腹鱼属鱼类线粒体基因组全序列(表 2)，以鲤鱼(*Cyprinus carpio*: KF856965) 和中华倒刺鲃(*Spinibarbus sinensis*: KF214722) 作为外群，基于去除终止密码子后的 12 个蛋白编码基因(L 链编码的 ND6 基因除外)串联序列，分别使用最大似然法(maximum likelihood method, ML)^[18] 和贝叶斯法(Bayesian method, BI)^[19] 构建裂腹鱼属鱼类的系统进化树。使用 MEGA-X^[17] 进行多重比对并构建 ML 树。构建 BI 树的步骤如下，将比对后的 FASTA 文件转为 NEXUS 文件；将其导入 CIPRES 软件(<https://www.phylo.org/>)^[20] 后设置参数并进行构树；最后采用

FigTree v1.4.4 软件对系统发育树进行编辑。构建 ML 系统树之前,先使用 jModelTest 2.1.7 软件^[21]基于赤池信息标准(Akaike information criteria, AIC)筛选最适核苷酸替代模型 HKY+I+G。

2 结果与分析

2.1 高通量测序及线粒体组装效果

高通量测序结果显示银色裂腹鱼的原始测序数据量为 5 065 Mb, 过滤处理后的有效数据量为 4 518 Mb, 插入片段平均长度为 458 bp。

有效数据量中平均 GC 含量为 38.17%, Q20 值为 96.11, Q30 值为 91.10。银色裂腹鱼基因组 A、T 碱基含量相近, 高于 C、G 碱基, 不存在混合碱基 N (图 1A)。而在起始、终止与中间位置处的碱基错误百分率明显高于其他 read 位置(图 1B)。以上结果均表明高通量测序数据质量较好, 符合后续线粒体基因组挖掘和分析。一般而言 GC-depth 在理想条件下都呈泊松分布。通过比较 reads 和组装的线粒体基因组序列, 并将组装序列的 GC 含量 reads 覆盖深度进行统计。从图 2

表 2 系统发生分析中使用的物种分类列表

Table 2 Species taxonomy used in phylogenetic analysis

Species	A+T (%)	Length (bp)	GenBank accession No.
<i>Schizothorax esocinus</i>	55.25	16 591	KT210882
<i>Schizothorax wangchiachii</i>	55.40	16 593	KC292197
<i>Schizothorax lissolabiatus</i>	54.95	16 583	KP796150
<i>Schizothorax plagiostomus</i>	55.23	16 563	KT184924
<i>Schizothorax griseus</i>	55.08	16 586	MF688995
<i>Schizothorax macropogon</i>	55.32	16 588	KC020113
<i>Schizothorax grahami</i>	54.93	16 584	KU234535
<i>Schizothorax waltoni</i>	55.30	16 589	JX202592
<i>Schizothorax lantsangensis</i>	54.92	16 580	KP143725
<i>Schizothorax richardsonii</i>	55.36	16 592	KC790369
<i>Schizothorax nepalensis</i>	55.42	16 589	AP011207
<i>Schizothorax nukiangensis</i>	55.16	16 585	KT223584
<i>Schizothorax prenanti</i>	55.03	16 587	KJ126773
<i>Schizothorax progastus</i>	55.22	16 575	KF739399
<i>Schizothorax labiatus</i>	55.24	16 590	KT944287
<i>Schizothorax kozlovi</i>	54.98	16 585	KJ755668
<i>Schizothorax chongi</i>	55.01	16 584	KJ718889
<i>Schizothorax pseudoaksaiensis</i>	55.55	16 582	KM079630
<i>Schizothorax oconnori</i>	55.45	16 590	KC513575
<i>Schizothorax argentatus</i>	55.52	16 580	Study
<i>Schizothorax yunnanensis</i>	55.10	16 585	KP892531
<i>Schizothorax dolichonema</i>	54.91	16 585	KJ577589
<i>Schizothorax sinensis</i>	54.99	16 571	MW191514
<i>Schizothorax davidi</i>	54.89	16 576	KM879227
<i>Schizothorax taliensis</i>	54.93	16 578	MH094667
<i>Schizothorax curvilabiatus</i>	55.49	16 578	MF804977
<i>Schizothorax gongshanensis</i>	55.12	16 591	KT946652
<i>Schizothorax eurystomus</i>	55.09	16 588	KY436758
<i>Cyprinus carpio</i> (outgroup)	56.72	16 581	KF856965
<i>Spinibarbus sinensis</i> (outgroup)	57.39	16 591	KF214722

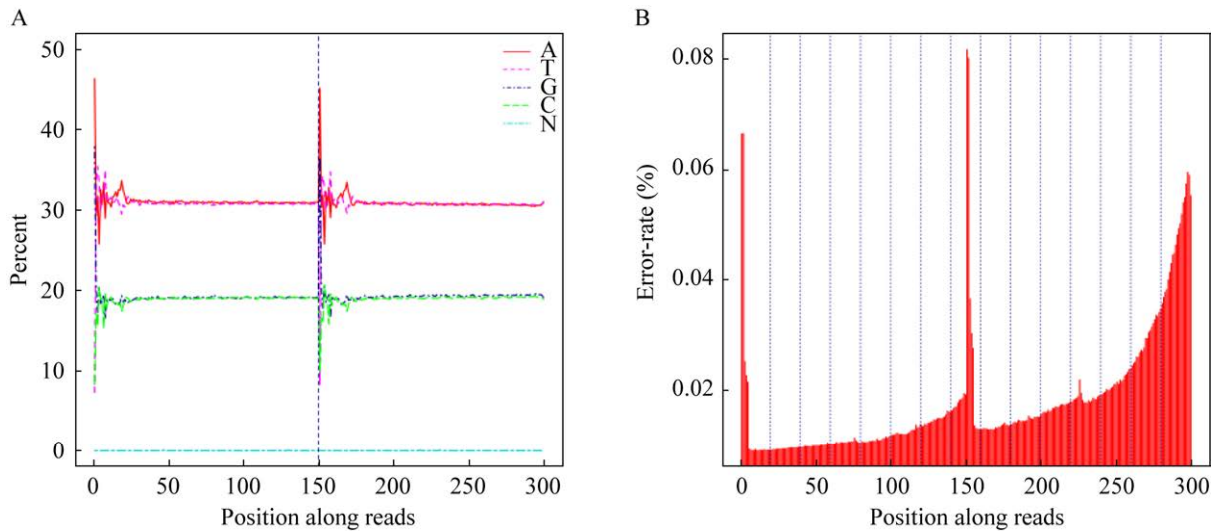


图 1 测序数据的碱基含量(A)和质量分布(B)

Figure 1 Base content (A) and quality distribution (B) of the sequencing data.

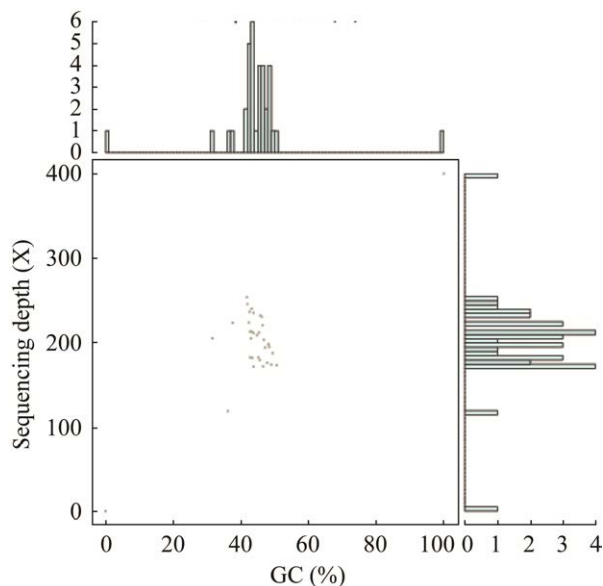


图 2 GC 含量与测序深度关联分析统计图 横坐标表示 GC 含量, 纵坐标表示测序深度, 右方是测序深度分布, 上方是 GC 含量分布

Figure 2 Correlation analysis between GC content and sequencing depth. The abscissa indicates GC content, the ordinate indicates sequencing depth, the right is sequencing depth distribution, and the top is GC content distribution.

可以看出测序长度在 150–300 bp 范围内, reads 覆盖度相对集中且 GC 无明显偏向性, 序列组装结果正常且不存在外源性污染。

为了进一步保障用于后续分析数据的完整度和准确性, 将“引物步移法”扩增得到的 DNA 片段截去两端重叠区域后分别得到长度为 1 985、2 231、1 963、437、2 360、2 018、1 255、762、1 664、959、946 bp 的 11 条序列片段, 将其进行串联拼接后获得总长度为 16 580 bp 的线粒体全序列, 与高通量测序组装获得的序列数据进行 BLAST 比对结果完全一致。

2.2 线粒体基因组结构与组成

本研究获得银色裂腹鱼线粒体基因组序列总长度为 16 580 bp, 位于 28 种已报道的裂腹鱼类线粒体全基因组序列长度范围之间(16 563–16 593 bp), 其结构高度保守, 符合典型的脊椎动物线粒体基因组结构^[22], 即包含 22 个 tRNA 基因、13 个蛋白编码基因(protein-coding genes, PCGs): 7 个脱氢酶亚单位(*ND1*、*ND2*、*ND3*、*ND4*、*ND4L*、*ND5*、*ND6*); 3 个细胞色素 C 氧

化酶(*COI*、*COII*、*COIII*); 2 个 *ATP* 酶亚基: *ATP6*、*ATP8*; 细胞色素 b 基因(*Cyt b*)、2 个 rRNA 基因 (12S rRNA、16S rRNA) 以及 1 个控制区(D-loop) (图 3)。其中, *ND6* 和 8 个 tRNA (tRNA-Gln、tRNA-Ala、tRNA-Asn、tRNA-Cys、tRNA-Tyr、tRNA-Ser^(UGA)、tRNA-Glu、tRNA-Pro) 基因位于轻链(L-链)上, 其余基因由重链(H-链)编码, 而控制区是线粒体全序列中唯一的大片段非编码区。

银色裂腹鱼线粒体基因组共存在 13 个间隔区域和 6 个重叠区域(表 3)。13 个间隔区域总长度为 69 bp, 占线粒体全长的 0.416%。其中 tRNA-Asn 和 tRNA-Cys 之间的间隔最大(33 bp), 其次为 tRNA-Asp 和 *COII* 之间的间隔(13 bp); 最小的间隔仅为 1 bp, 共有 5 个, 分别位于 tRNA-Leu^(UAA) 和 *ND1*、tRNA-Ala 和 tRNA-Asn、tRNA-Tyr 和 *COI*、tRNA-Lys 和 *ATP8*、tRNA-Ser^(GCU)

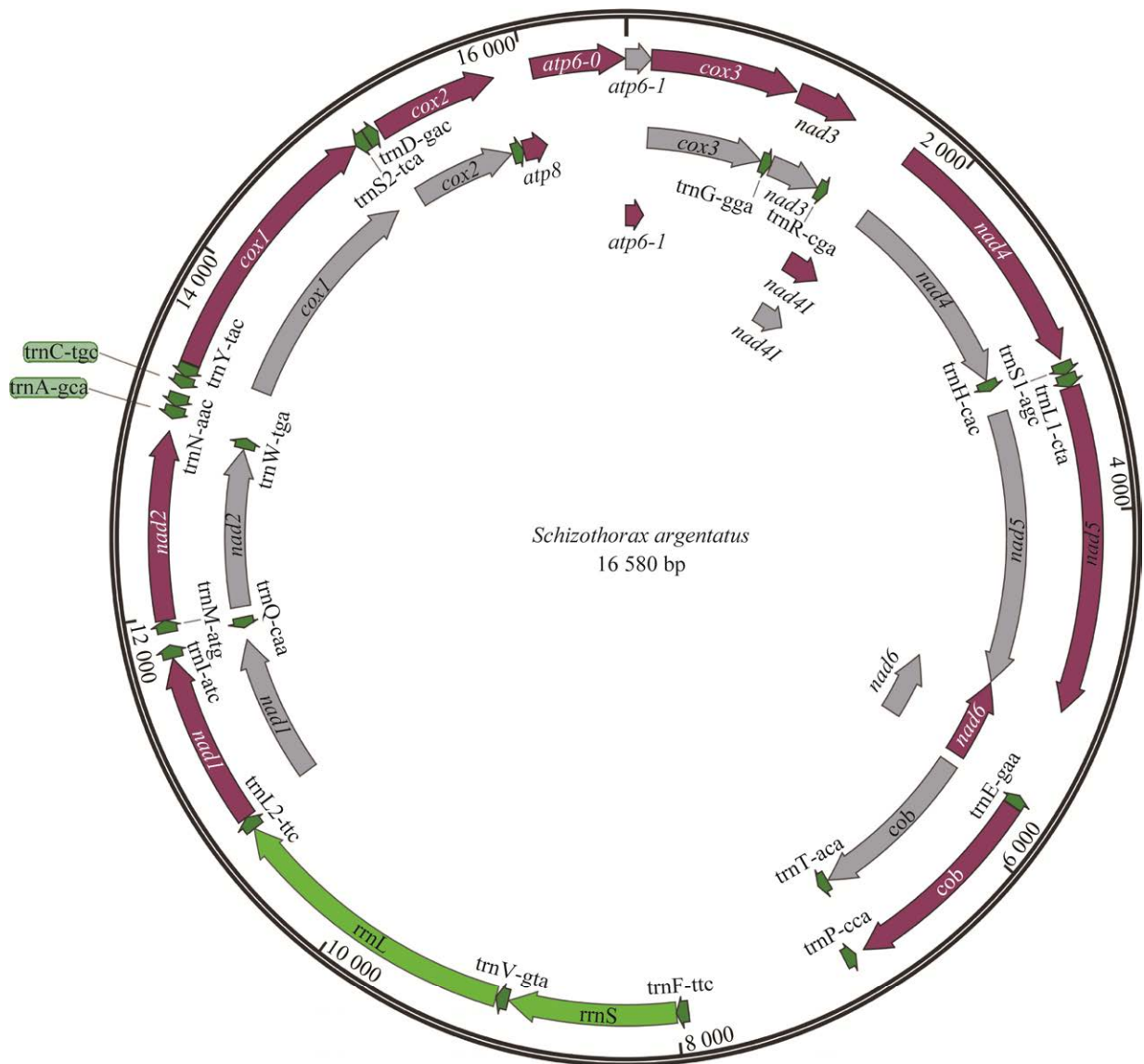


图 3 银色裂腹鱼线粒体基因组结构图

Figure 3 Mitochondrial genome structure of *Schizothorax argentatus*.

表3 银色裂腹鱼线粒体基因组结构特征

Table 3 Structure and characteristics of mitochondrial genome of *Schizothorax argentatus*

Gene	Coding strand	Start	End	Length (bp)	Space or overlap (bp)	Amino acid	Start codon	End codon	Anticodon	Anticodon site
tRNA-Phe	H	1	69	69	0				GAA	32-34
12S rRNA	H	70	1 024	955	0					
tRNA-Val	H	1 025	1 096	72	0				TAC	34-36
16S rRNA	H	1 097	2 773	1 677	0					
tRNA-Leu ^(UAA)	H	2 774	2 849	76	1				TAA	37-39
<i>ND1</i>	H	2 851	3 825	975	4	324	ATG	TAA		
tRNA-Ile	H	3 830	3 901	72	-2				GAT	33-35
tRNA-Gln	L	3 900	3 970	71	2				TTG	33-35
tRNA-Met	H	3 973	4 041	69	0				CAT	31-33
<i>ND2</i>	H	4 042	5 086	1 045	0	348	ATG	T--		
tRNA-Trp	H	5 087	5 157	71	2				TCA	33-35
tRNA-Ala	L	5 160	5 228	69	1				TGC	31-33
tRNA-Asn	L	5 230	5 302	73	33				GTT	34-36
tRNA-Cys	L	5 336	5 402	67	-1				GCA	29-31
tRNA-Tyr	L	5 402	5 472	71	1				GTA	33-35
<i>COI</i>	H	5 474	7 024	1 551	0	516	GTG	TAA		
tRNA-Ser ^(UGA)	L	7 025	7 095	71	3				TGA	33-35
tRNA-Asp	H	7 099	7 170	72	13				GTC	34-36
<i>COII</i>	H	7 184	7 874	691	0	230	ATG	T--		
tRNA-Lys	H	7 875	7 950	76	1				TTT	37-39
<i>ATP8</i>	H	7 952	8 116	165	-7	54	ATG	TAG		
<i>ATP6</i>	H	8 110	8 792	683	0	227	ATG	TA-		
<i>COIII</i>	H	8 793	9 576	784	0	261	ATG	T--		
tRNA-Gly	H	9 577	9 648	72	0				TCC	34-36
<i>ND3</i>	H	9 649	9 997	349	0	116	ATG	T--		
tRNA-Arg	H	9 998	10 067	70	0				TCG	32-34
<i>ND4L</i>	H	10 068	10 364	297	-7	98	ATG	TAA		
<i>ND4</i>	H	10 358	11 738	1 381	0	460	ATG	T--		
tRNA-His	H	11 739	11 807	69	0				GTG	31-33
tRNA-Ser ^(GCU)	H	11 808	11 875	68	1				GCT	27-29
tRNA-Leu ^(UAG)	H	11 877	11 949	73	3				TAG	34-36
<i>ND5</i>	H	11 953	13 776	1 824	-4	607	ATG	TAA		
<i>ND6</i>	L	13 773	14 294	522	0	173	ATG	TAA		
tRNA-Glu	L	14 295	14 363	69	4				TTC	31-33
<i>Cyt b</i>	H	14 368	15 508	1 141	0	380	ATG	T--		
tRNA-Thr	H	15 509	15 580	72	-1				TGT	33-35
tRNA-Pro	L	15 580	15 649	70	0				TGG	32-34
D-loop	H	15 650	16 580	931	0					

和 tRNA-Leu^(UAG)之间; 剩余间隔大小分别是 4 bp (*ND1* 和 tRNA-Ile、tRNA-Glu 和 *Cyt b*)、3 bp [tRNA-Ser^(UGA)和 tRNA-Asp、tRNA-Leu^(UAG)和 *ND5*]、2 bp (tRNA-Gln 和 tRNA-Met、tRNA-Trp 和 tRNA-Ala)。6 个重叠区域总长度为 22 bp, 其中 *ATP8* 和 *ATP6* 以及 *ND4L* 和 *ND4* 之间的重叠区域最大(7 bp); tRNA-Cys 和 tRNA-Tyr 以及 tRNA-Thr 和 tRNA-Pro 之间的重叠区域最小(1 bp)。

银色裂腹鱼线粒体全序列中各碱基含量如表 4 所示, A 碱基含量最高, 为 30.25%; 其次为 C 碱基, 占比 27.20%; 然后是 T 碱基(25.27%) 和 G 碱基(17.28%)。银色裂腹鱼的线粒体基因组 AT 碱基占比较高(55.52%), 在裂腹鱼属线粒体基因组中仅次于伊犁裂腹鱼(55.55%), 表现出明显的 AT 偏好性。

2.3 蛋白编码基因

银色裂腹鱼线粒体蛋白质编码基因总长为 11 408 bp, 占线粒体全基因组的 68.81%, 有明显的 AT 偏好性(A+T 占 55.09%), 编码 3 794 个氨基酸残基。除 *ND6* 基因由 L 链编码外, 其余蛋白编码基因均位于 H 链上。其中 *ND5* 的碱基序列最长(1 824 bp), 编码 607 个氨基酸; *ATP8* 最短(165 bp), 仅编码 54 个氨基酸。所有的蛋白编码基因中, 除 *COI*以 GTG 为起始密码子外, 其余基因均以 ATG 为起始密码子。*ND1*、*COI*、*ND4L*、*ND5*、*ND6* 以 TAA 为终止密码子, *ATP8* 以 TAG 为终止密码子, 剩余基因均为不完全密码子(*ATP6* 为 TA-; *ND2*、*COII*、*COIII*、*ND3*、*ND4* 和 *Cyt b* 为 T-), 这种不完全密码子现象常见于后生动物线粒体基因组中。

由表 4 可以看出, 银色裂腹鱼线粒体基因组蛋白编码基因中碱基分布不均, 全部蛋白编码基因序列的总 AT 和 GC 占比分别为 55.09%和

44.91%。其中 T 碱基含量最高的是位于 L 链上的 *ND6* (38.89%), 最低的是 *ND2* (22.30%); C 碱基含量最高的是 *ND2* (31.96%), 最低的是 *ND6* (15.90%); A 碱基含量最高的是 *ATP8* (34.55%), 最低的是 *ND6* (13.79%); G 碱基含量最高的是 *ND6* (31.42%), 最低的是 *ATP8* (11.52%)。AT 碱基含量最高的为 *ATP8* (63.03%), 最低的为 *ND1* (51.28%); CG 碱基含量最高的是 *ND1* (48.72%), 最低的是 *ATP8* (36.97%)。密码子碱基组成上来看, 密码子第 1 位点上 4 种碱基含量较为接近, 碱基使用偏好不明显; 密码子第 2 位点上 T 碱基含量最高(40.67%) 而 G 碱基含量最低(13.60%), 表现出明显的 T 偏倚和反 G 偏倚; 而密码子第 3 位点上 A 碱基含量最高(38.71%), G 碱基含量最低(11.13%)。密码子第 1 位点的 AT 含量为 47.54%明显低于第 2 位和第 3 位 AT 含量(59.14%和 58.61%)。根据 MEGA 软件分析所得, 银色裂腹鱼线粒体蛋白编码基因中 *ND6* 表现为较强的 GC 偏好性, 其余基因表现为 AT 偏好性(图 4)。

初步分析了银色裂腹鱼线粒体基因组 13 个蛋白编码基因的相对同义密码子使用度(relative synonymous codon usage, RSCU)和氨基酸种类与数量。结果表明在编码的 3 794 个氨基酸中, 亮氨酸(Leu)是使用频率最高的氨基酸(16.50%), 丙氨酸(Ala)次之(8.83%), 半胱氨酸(Cys)数量最少, 占比仅为 0.66%(图 5)。13 个蛋白质编码基因中存在 29 个偏好密码子(RSCU \geq 1)^[23]。密码子 AGA、AGG 没有被使用, 因此除密码子 AGA 外, 第 3 位点是 A 碱基的密码子均为偏好密码子。编码亮氨酸(Leu)的密码子 CUA 最为常见(RSCU=2.34), 其次为编码异亮氨酸(Ile)的密码子 AUU (RSCU=1.12)。密码子 UAA 和 UAG 为终止密码子, 不编码氨基酸(表 5)。

表 4 银色裂腹鱼线粒体蛋白编码基因、转运 RNA、核糖体 RNA 和控制区碱基组成

Table 4 Base composition of mitochondrial protein-coding genes (PCGs), tRNAs, rRNAs, and D-loop of *Schizothorax argenteus*

Gene name/Codon site	Length (bp)	Base composition						AT-skew	GC-skew
		T (%)	C (%)	A (%)	G (%)	(A+T) (%)	(C+G) (%)		
<i>ND1</i>	975	25.85	29.85	25.44	18.87	51.28	48.72	-0.01	-0.23
<i>ND2</i>	1 045	22.30	31.96	30.72	15.02	53.01	46.99	0.18	-0.36
<i>COI</i>	1 551	28.95	26.11	26.69	18.25	55.64	44.36	-0.05	-0.18
<i>COII</i>	691	27.79	25.33	30.54	16.35	58.32	41.68	0.07	-0.22
<i>ATP8</i>	165	28.48	25.45	34.55	11.52	63.03	36.97	0.16	-0.38
<i>ATP6</i>	683	26.94	28.70	29.87	14.49	56.81	43.19	0.07	-0.33
<i>COIII</i>	784	26.40	28.70	27.04	17.86	53.44	46.56	0.01	-0.23
<i>ND3</i>	349	28.94	28.08	28.65	14.33	57.59	42.41	-0.01	-0.32
<i>ND4L</i>	297	28.28	30.30	25.59	15.82	53.87	46.13	-0.06	-0.31
<i>ND4</i>	1 381	26.21	28.75	28.46	16.58	54.67	45.33	0.05	-0.27
<i>ND5</i>	1 824	25.60	29.11	30.59	14.69	56.20	43.80	0.11	-0.33
<i>ND6</i>	522	38.89	15.90	13.79	31.42	52.68	47.32	-0.53	0.33
<i>Cyt b</i>	1 141	28.57	27.87	27.34	16.21	55.92	44.08	-0.03	-0.26
1 st codon	-	21.15	26.24	26.40	26.21	47.54	52.46	0.10	0.00
2 nd codon	-	40.67	27.26	18.47	13.60	59.14	40.86	-0.54	-0.33
3 rd codon	-	19.89	30.26	38.71	11.13	58.61	41.39	0.45	-0.46
PCGs	11 408	27.24	27.92	27.86	16.99	55.09	44.91	0.01	-0.24
tRNAs	1 562	26.95	21.13	28.49	23.43	55.44	44.56	0.03	0.05
rRNAs	2 632	19.53	25.00	34.42	21.05	53.95	46.05	0.32	-0.09
D-loop	931	32.65	19.87	34.05	13.43	66.70	33.30	0.04	-0.19
mtDNA	16 580	25.27	27.20	30.25	17.28	55.52	44.48	0.11	-0.22

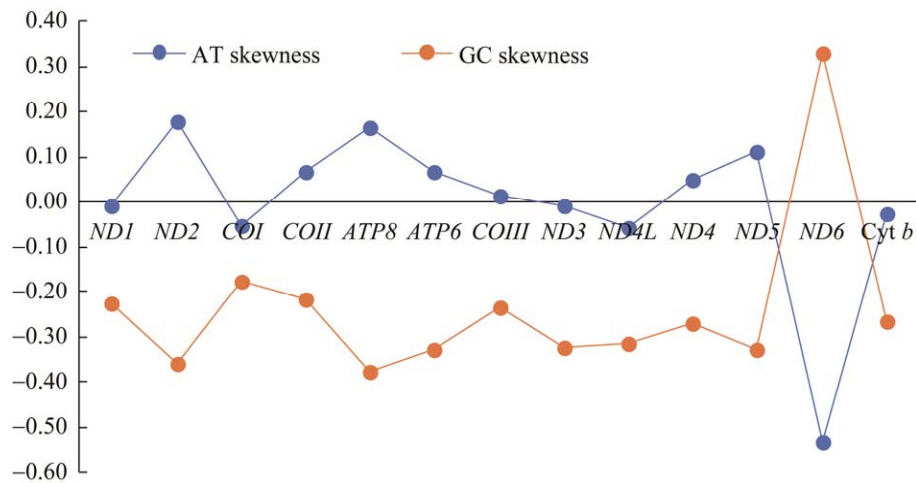


图 4 银色裂腹鱼线粒体基因组蛋白编码基因偏倚分布

Figure 4 Bias distribution of protein coding genes in mitochondrial genome of *Schizothorax argenteus*.

表 5 银色裂腹鱼线粒体编码蛋白密码子使用频率

Table 5 Codon usage frequency of mitochondrial encoded protein in *Schizothorax argentatus*

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU (F)	95	0.83	UCU (S)	39	1.00	UAU (Y)	49	0.87	UGU (C)	6	0.48
UUC (F)	133	1.17	UCC (S)	53	1.36	UAC (Y)	64	1.13	UGC (C)	19	1.52
UUA (L)	109	1.04	UCA (S)	82	2.11	UAA (*)	5	3.33	UGA (W)	96	1.60
UUG (L)	14	0.13	UCG (S)	9	0.23	UAG (*)	1	0.67	UGG (W)	24	0.40
CUU (L)	94	0.90	CCU (P)	17	0.32	CAU (H)	31	0.59	CGU (R)	9	0.47
CUC (L)	96	0.92	CCC (P)	75	1.40	CAC (H)	74	1.41	CGC (R)	13	0.68
CUA (L)	244	2.34	CCA (P)	106	1.98	CAA (Q)	91	1.80	CGA (R)	39	2.05
CUG (L)	69	0.66	CCG (P)	16	0.30	CAG (Q)	10	0.20	CGG (R)	15	0.79
AUU (I)	159	1.12	ACU (T)	43	0.57	AAU (N)	49	0.82	AGU (S)	8	0.21
AUC (I)	124	0.88	ACC (T)	107	1.41	AAC (N)	71	1.18	AGC (S)	42	1.08
AUA (M)	109	1.27	ACA (T)	128	1.68	AAA (K)	60	1.58	AGA (*)	0	0.00
AUG (M)	63	0.73	ACG (T)	26	0.34	AAG (K)	16	0.42	AGG (*)	0	0.00
GUU (V)	54	0.91	GCU (A)	52	0.62	GAU (D)	23	0.61	GGU (G)	28	0.46
GUC (V)	43	0.73	GCC (A)	140	1.67	GAC (D)	52	1.39	GGC (G)	44	0.72
GUA (V)	97	1.64	GCA (A)	118	1.41	GAA (E)	83	1.58	GGA (G)	104	1.69
GUG (V)	43	0.73	GCG (A)	25	0.30	GAG (E)	22	0.42	GGG (G)	70	1.14

Preference codons are indicated in bold. Asterisks represent termination codons and capital letters in brackets mean the abbreviations of amino acids.

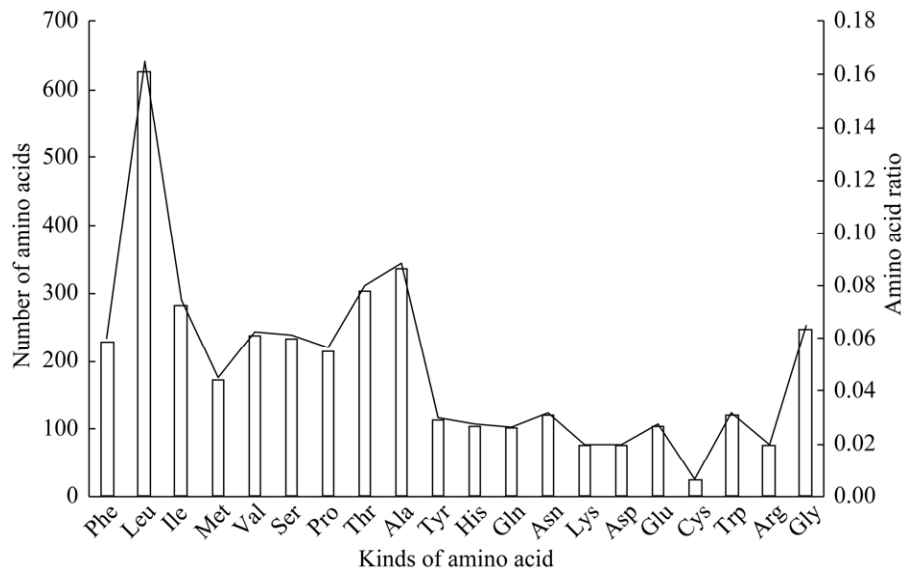


图 5 蛋白编码基因编码的氨基酸种类与数量

Figure 5 Types and quantities of amino acids encoded by mitochondrial protein-coding genes.

2.4 rRNA 和 tRNA 结构特征

银色裂腹鱼线粒体基因组 12S rRNA 位于 tRNA-Phe 和 tRNA-Val 间, 而 16S rRNA 位于 tRNA-Val 和 tRNA-Leu 间, 长度分别是 955 bp 和 1 677 bp。二者均由 H 链编码, 且前后既无间

隔区域也无重叠区域。rRNAs 的 AT 含量为 53.95%、GC 含量为 46.05%, 其中 A 碱基的占比最大(34.42%)。银色裂腹鱼 rRNA 基因的二级结构比较保守, 形成了多个大小不一的茎环结构, 预测得到的 12S rRNA 基因的自由能值为

-214.70 kcal/mol, 16S rRNA 基因的自由能值为 -409.40 kcal/mol。由于该基因进化速度在线粒体各基因组中也相对较慢,故适用于科及以上分类阶元系统进化关系的研究^[24]。

以 CRW 网站^[25]提供的鲤鱼(*Cyprinus carpio*)线粒体 12S rRNA (*b.16.m.C.carpio*)文件和 RiboVision 网站提供的人(*Homo sapiens*)线粒体 16S rRNA (*mHS_LSU_3D*)文件为模板对银色裂腹鱼的 rRNA 二级结构进行预测。银色裂腹鱼线粒体的 rRNA 结构整体上与两者的结构类似,只在部分位点存在差异。12S rRNA 二级结构中包

含 4 个结构域(图 6A), 其中第 I、第 II 结构域称为可变区, 第 III、第 IV 结构域则为保守区。16S rRNA 二级结构包含 6 个结构域(图 6B), 其中第 I-第 III 和第 VI 结构域为可变区, 第 IV 与第 V 结构域为保守区。与鲤鱼相比, 银色裂腹鱼线粒体 12S rRNA 在第 I、第 II 结构域多处出现碱基的替换、插入与重排现象, 而第 III、第 IV 结构域则相对较少。其中重排和插入现象均只发生在第 I、II、IV 结构域。第 III 结构域仅出现替换现象。重排区域有 7 处, 其中 A 碱基占比较大; 插入的碱基仅 6 个, 且均为 G、C 碱基。重排和插入现象均发

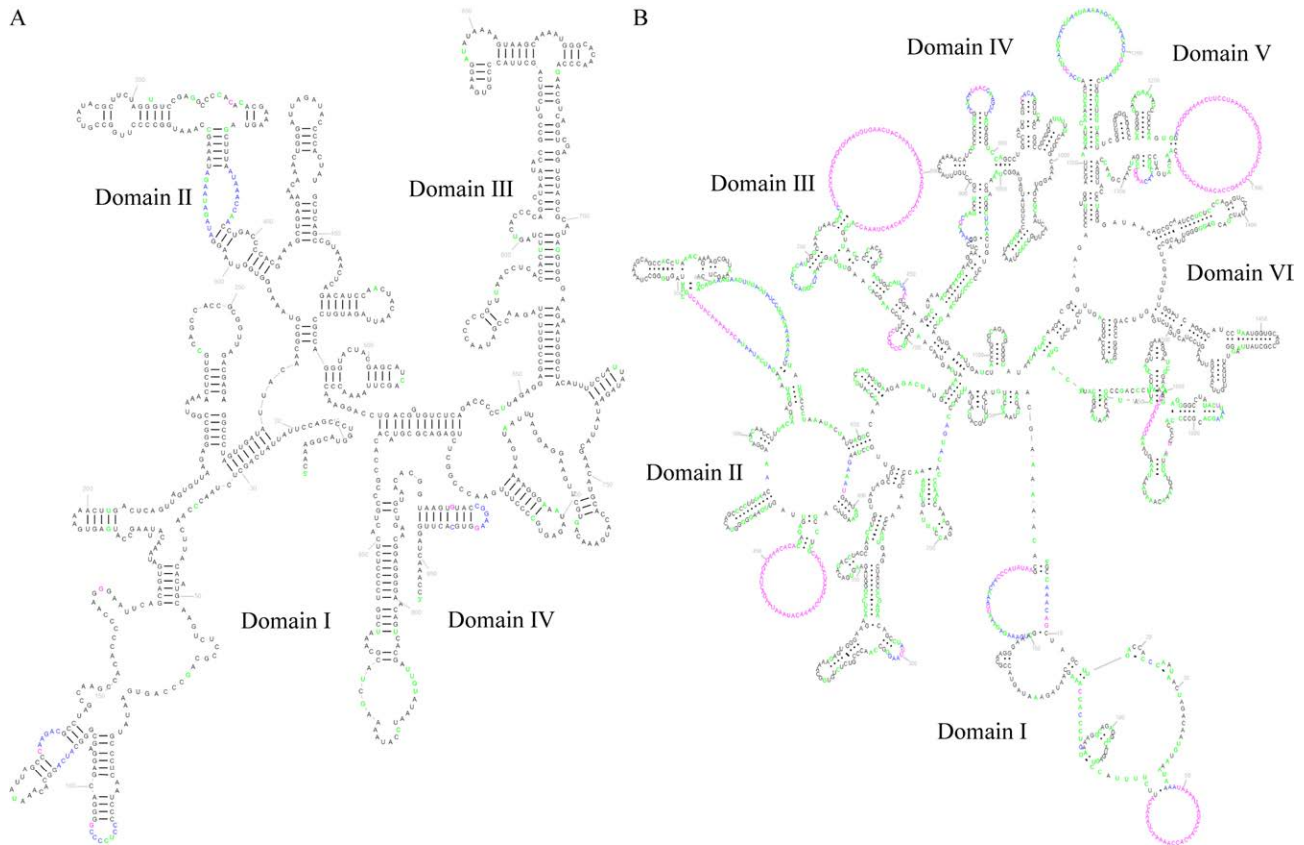


图 6 银色裂腹鱼线粒体 12S rRNA (A)和 16S rRNA (B)二级结构预测图 黑色表示碱基与模板相同; 绿色表示碱基与模板不同; 粉色表示插入的碱基; 蓝色表示碱基的重排

Figure 6 Predicted secondary structure of mitochondrial 12S rRNA (A) and 16S rRNA (B) of *Schizothorax argentatus*. Black indicates that those nucleotides were the same as the template; Green indicates those nucleotides were modified compared to the template; Pink indicates that those nucleotides were inserted nucleotides; Blue indicates that those nucleotides were repositioned compared to the template.

生在环区上, 而替换现象茎环结构中均有发生。相较于人的 16S rRNA, 由于物种差异较大, 碱基的替换、插入与重排现象较为频繁, 且均只发生在环区。因此 rRNA 茎区的碱基序列比环区的碱基序列要相对保守, 较少出现碱基的替换、插入与重排现象。

银色裂腹鱼 mtDNA 包含长度在 67–76 bp 之间的 22 个 tRNA 基因, 可运载 20 种氨基酸。这些 tRNA 在 H 链和 L 链上均有编码, 其中

tRNA-Gln、tRNA-Ala、tRNA-Asn、tRNA-Cys、tRNA-Tyr、tRNA-Ser^(UGA)、tRNA-Glu、tRNA-Pro 位于 L 链, 其余 14 个 tRNA 在 H 链上编码(表 3)。全部 tRNA 碱基含量分别为: T=26.95%、C=21.13%、A=28.49%、G=23.43%, A+T 含量高达 55.44%, 而 G+C 含量为 44.56% (表 4)。除 tRNA-Ser^(GCU) 缺少二氢尿嘧啶(DHU)茎外, 其余 21 个 tRNA 都能够折叠成典型的三叶草二级结构(图 7)。22 个 tRNA 中存在 4 种碱基错配, 分

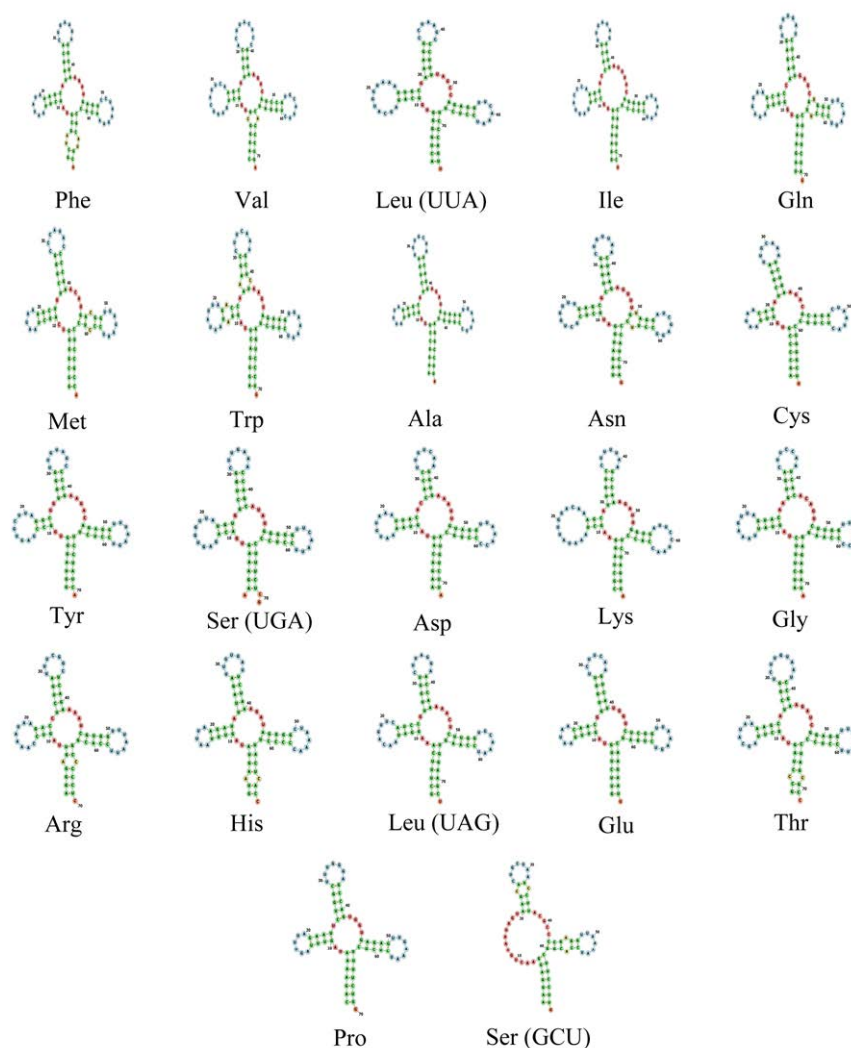


图 7 银色裂腹鱼线粒体 22 个 tRNA 二级结构预测图

Figure 7 Secondary structure prediction of 22 tRNAs in mitochondrion of *Schizothorax argentatus*.

别为 A-C、U-U、A-A 和 C-C。除 tRNA-Trp 和 tRNA-Ser^(GCU)在 DHU 茎和反密码子臂上存在错配外,其余错配主要发生在氨基酸接受臂和 T Ψ C 臂上。这些 tRNA 基因的部分错配可通过后期的 RNA 编辑进行校正,不会引起氨基酸转运障碍^[26]。tRNAs 的二级结构中,tRNA-Val、tRNA-Ile、tRNA-Cys、tRNA-Tyr、tRNA-Ser^(UGA)、tRNA-Lys 和 tRNA-Arg 的 DHU 茎为 3 对碱基(其余为 4 bp);tRNA-Phe 和 tRNA-Lys 的 T Ψ C 茎为 4 对碱基,tRNA-Ser^(GCU)的 T Ψ C 茎为 6 对碱基(其余为 5 bp);tRNA-Val、tRNA-Ile、tRNA-Tyr、tRNA-Asp、tRNA-Gly 和 tRNA-Thr 的反密码子臂为 4 对碱基(其余为 5 bp)。T Ψ C 环碱基数除 tRNA-Phe 为 8 个碱基以及 tRNA-Lys 为 9 个碱基外,其余均为 7 个碱基;反密码环除 tRNA-Val 和 tRNA-Thr 为 9 个碱基外,其余也均为 7 个碱基,DHU 环长度变化较大,在 5-13 个碱基之间。

2.5 线粒体基因组控制区结构特征

线粒体控制区(D-loop)属于线粒体 DNA 中的非编码区,是线粒体 DNA 中进化速率最快的序列,被广泛应用于鱼类群体遗传学和分子系统学的研究^[27-28]。银色裂腹鱼线粒体控制区位于 tRNA-Pro 和 tRNA-Phe 之间,长度为 931 bp。D-loop 区的 A (34.05%)、T (32.65%)碱基含量较高,而 C (19.87%)、G (13.43%)碱基含量相对偏低。线粒体控制区序列可分为终止序列区(terminal associated sequence, TAS)、中央保守区(central conserved domain, CD)、保守序列区(conserved sequence block, CSB)。其中,中央保守区又可依次分为 CSB-F、CSB-E、CSB-D、CSB-C、CSB-B 区,保守序列区也可分为 CSB1、CSB2、CSB3 区^[28-30]。

本研究参考了几种裂腹鱼属和鲤科鱼类控制区序列,识别到银色裂腹鱼的终止序列区长度为 242 bp,包含与 DNA 复制终止相关的 1 个扩展终止相关序列区(extended terminal associated sequence, ETAS)和 3 个重复的核心 TACAT 序列及 2 个反向互补序列 ATGTA 序列,而 TACAT 与其互补序列 ATGTA 可形成发夹结构。与终止序列区和保守序列区相比,中央保守区相对更加保守。银色裂腹鱼控制区的中央保守区长度为 332 bp,其内可识别相似的保守序列,分别对应 CSB-F、CSB-E、CSB-D 和 CSB-B,其中 CSB-F 为划分终止序列区和中央保守区的标志。CSB-F 的关键序列为 5'-ATATTAATGTAGTAA GAAACCACCAA-3',紧随其后的 CSB-E 可通过 GTGGG-box 序列进行识别。识别到的 CSB-E 的关键序列为 5'-AGGGACAATCAGTGTGGGGGT-3',其上游为 CSB-D,关键序列为 5'-TATTACTGGCA TCTGGTTCCTATTTTCAGG-3',该序列在多种鲤科鱼类间均无差异,具有高度保守性。CSB-D 序列之后还识别了与哺乳类 CSB-B 相似的序列,该序列在裂腹鱼属间差异不大,其关键序列为 5'-TTATATGCATAACGTATTTTTT-3'。最后是保守序列区,其序列长度为 357 bp,可识别到 CSB1 (5'-TCAATTATCGTATGACATA-3')、CSB2 (5'-CAAACCCCCCTACCCCC-3')和 CSB3 (5'-TGTCAAACCCCTAAACCAA-3')序列。其中,CSB1 与线粒体 DNA 的复制起始相关,因而保守序列区被认为是整个控制区最关键的部分^[31]。在银色裂腹鱼的 CSB3 下游约 50 bp 处还发现了一段保守的 TT (AT) nGTG 序列片段,该片段在鲤形目其他鱼类中也普遍存在,AT 重复片段数目的不同,可导致不同鱼类保守序列区长度的变异。银色裂腹鱼线粒体控制区结构详情见图 8。

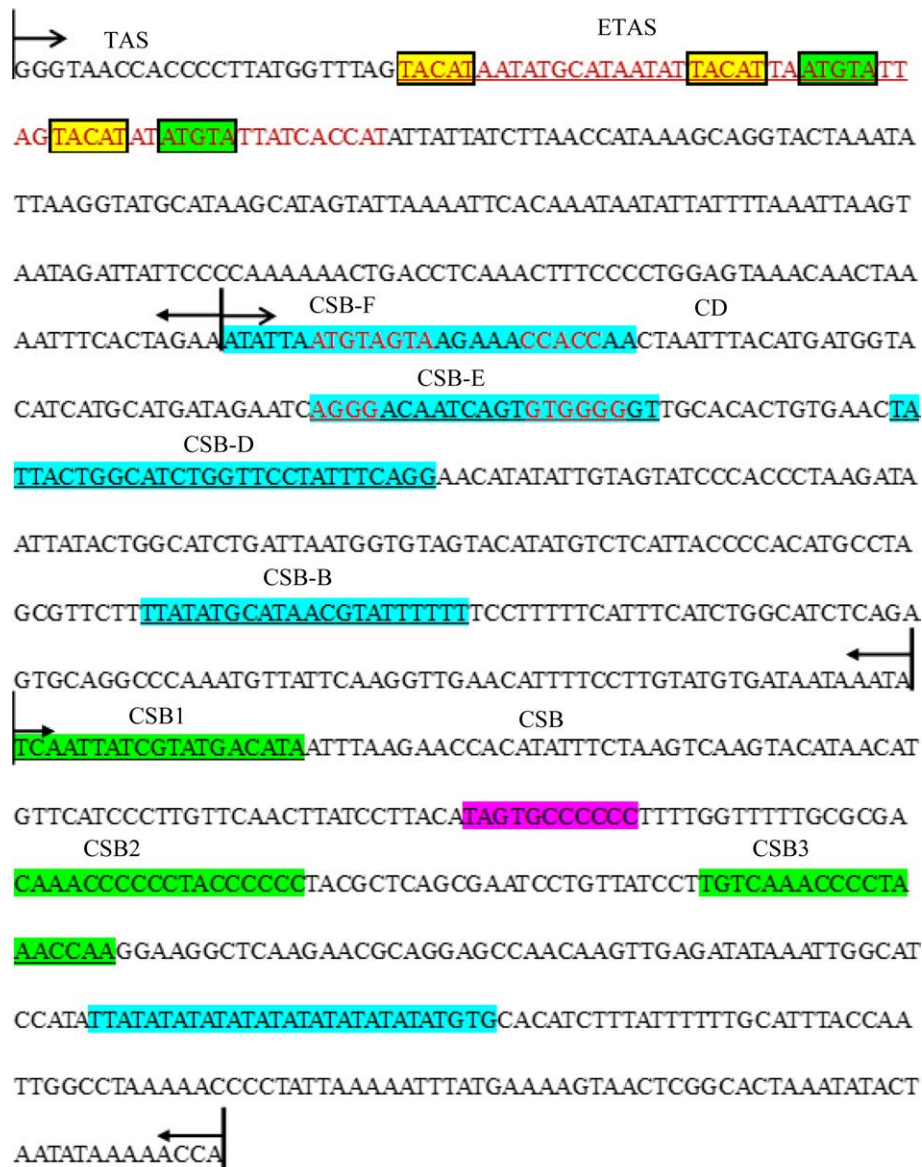


图 8 银色裂腹鱼控制区结构

Figure 8 Mitochondrial control region structure of *Schizothorax argentatus*.

2.6 系统发育分析

使用 28 种裂腹鱼属鱼类线粒体 H 链编码的 12 个 PCGs 串联序列, 基于 1 000 次自检重复分别构建 ML 和 BI 系统发育树(图 9)。结果显示两种类型的进化树有着相似的拓扑结构, 分布区相同的鱼类聚到同一个进化支中。族群 I 包含的裂腹鱼类主要分布于云南、贵州、四川三省区域, 由于所处的地理环境差异较大, 故其种类繁多且分

支复杂; 族群 II 均分布于喜马拉雅山脉水系; 族群 III 主要分布于雅鲁藏布江及其支流; 族群 IV 中的伊犁裂腹鱼、银色裂腹鱼和宽口裂腹鱼分布在我国新疆伊犁河和塔里木河, 与其他裂腹鱼类存在地理隔离。在 ML 树中, 银色裂腹鱼所在的族群 IV 是这些裂腹鱼中最早分化的一支, 其次为族群 III、族群 II, 最后分化的是族群 I。BI 树拓扑结构稍有不同, 最先分化的是族群 II, 其次为族群 IV。

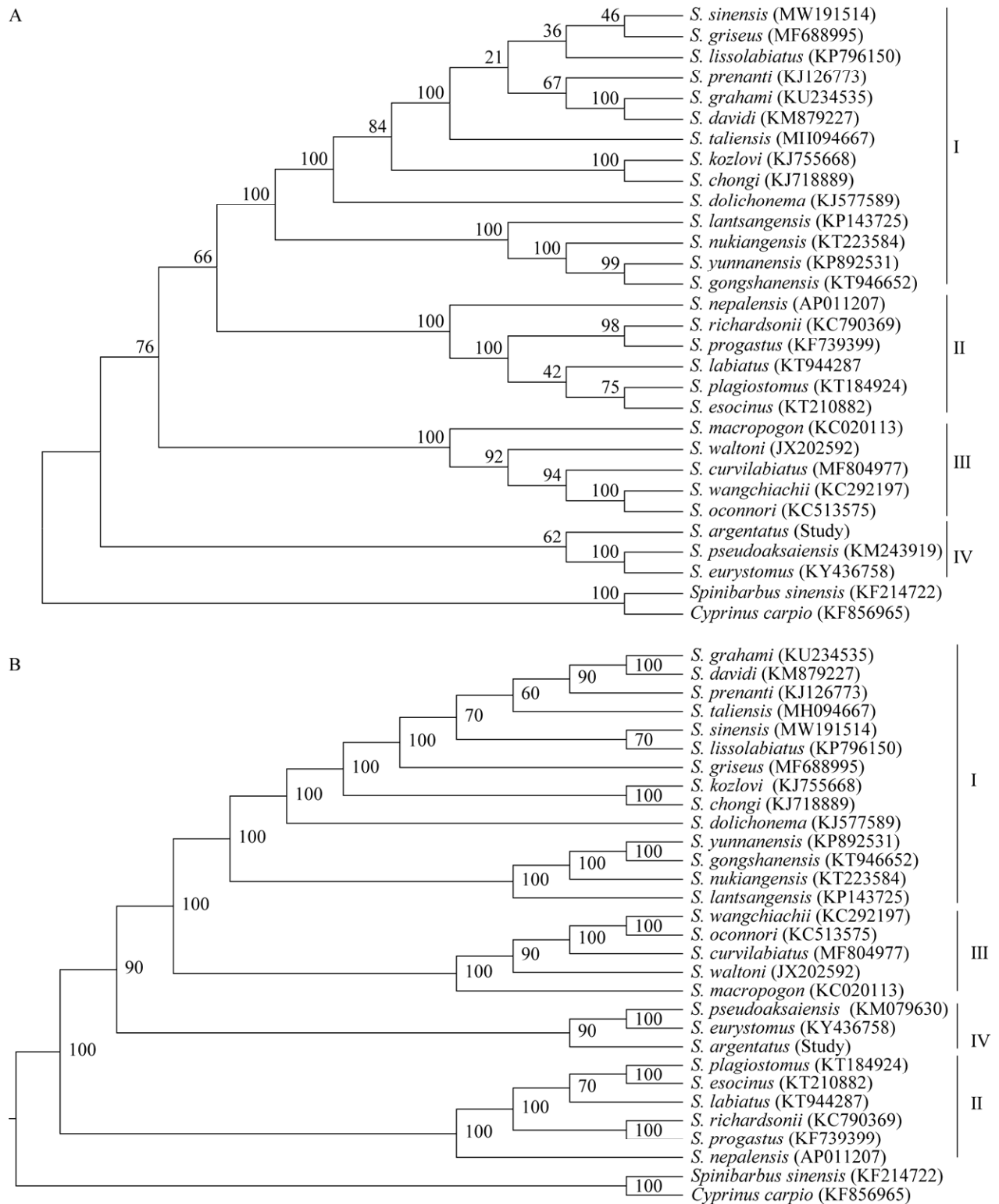


图9 基于12个PCGs串联序列构建的ML (A)和BI (B)分子系统树

Figure 9 Molecular phylogenetic trees constructed by ML (A) and BI (B) methods based on 12 PCGs concatenated sequences.

3 讨论与结论

本研究采用高通量测序技术获得了长度为 16 580 bp 的银色裂腹鱼线粒体基因组全序列, 其结构特征与已发布裂腹鱼属鱼类相似, 各基因长度也基本一致, 表明该类群鱼类线粒体基因组在进化上具有高度保守性^[2]。AT-skew 和 GC-skew 值可用来衡量重链与轻链之间碱基含的差异程度, 绝对值越大表明两者之间碱基组成差异越明显^[32]。银色裂腹鱼线粒体 DNA 序列具有较高的 A+T 含量和较低的 G+C 含量, 表现出明显的 AT 偏好性, 这一现象在其他裂腹鱼类中同样存在, 只是含量大小因物种不同而略有差异, 造成这种碱基偏好性的原因可能与复制和转录过程中的自然突变和选择压力等因素有关^[33]。除非编码区和 tRNA 碱基数量变化较大外, 各基因与裂腹鱼亚科鱼类的其他物种相似, 并且具有很高的同源性, 在其非编码区也存在 AT 重复的典型特征^[34]。蛋白质编码基因除密码子第 1 位点各碱基分布较为均匀, 第 2 和第 3 位点均呈现出明显的反 G 偏倚, 氨基酸对碱基分布的限制性及相应密码子使用频率差异是造成蛋白编码基因碱基分布不均衡的一个重要原因^[35]。对银色裂腹鱼密码子偏好性分析显示, NNA 类型的密码子使用频率最高, 与其蛋白编码基因组成中第 3 位点碱基 A 偏好性一致, 该结果与青藏高原 3 种裂腹鱼类线粒体全序列中密码子偏好性特征一致^[36]。在编码蛋白质的 20 种氨基酸中, 疏水型氨基酸 Leu 的使用频率最高, 这一现象在裂腹鱼属其他鱼类中也较为常见^[36-37], 可能与线粒体基因编码的跨膜蛋白其氨基酸大部分由疏水性氨基酸组成有关。值得一提的是, ND6 基因的碱基组成中 AT-skew 的绝对值最高且 GC-skew 值是唯一的正值, AT/GC 偏倚值波动较大。质子转运 NADH-醌氧化还原酶(proton-translocating

NADH-quinone oxidoreductase)又简称复合体 I (complex I), 是一种多亚基的膜蛋白复合体, 能够催化 NADH 呼吸链的电子转移, 并提供脊椎动物合成 ATP 所需的约 40%质子动力^[38-39]。线粒体 DNA 编码的 ND6 亚基作为复合体 I 不可或缺的组成部分, 推测该基因经历的与呼吸代谢相关的选择和突变压力可能与其他基因显著不同。

银色裂腹鱼线粒体基因组两种 rRNA 总长为 2 632 bp, 且 12S rRNA 和 16S rRNA 基因前后无间隔区, 符合后生动物的典型性特征^[24]。本研究得出 12S rRNA 基因的自由能值绝对值比 16S rRNA 基因低, 一般而言, 自由能值越低则其分子结构越稳定, 说明前者较后者更加保守。此外, 16S rRNA 的二级结构预测图也显示较为频繁的碱基替换、插入与重排现象, 其变异程度较 12S rRNA 更大。因此, 12S rRNA 基因也常被用作 DNA 宏条形码(DNA meta barcoding)来进行鱼类鉴定和系统发育研究^[40]。22 个 tRNA 中仅 tRNA-Ser^(GCU)无法形成三叶草结构, 与大多数脊椎动物 tRNA 结构预测图相似, 亦是鱼类线粒体基因组的共同特质, 有研究表明这种缺失 DHU 臂的 tRNA 可通过后期调整其结构形态和功能融入核糖体中, 发挥其携带并转运氨基酸的功能^[41]。此外, 碱基错配现象也普遍存在于 tRNA 的二级结构中, 由于线粒体基因组不受重组过程的影响, 因此允许存在碱基错配, 并且这些错配有助于消除有害突变^[42]。

线粒体控制区主要包含了控制线粒体基因组复制和表达的调控元件, 对其结构功能的研究将有助于了解 DNA 复制、转录机制和进化规律^[27-28]。通过对比分析, 我们在银色裂腹鱼控制区中识别出终止序列区、中央保守区和保守序列区, 与谢佳燕^[43]对齐口裂腹鱼控制区序列研究的结果一致。在终止序列区中识别了可形成发夹结构的特征序列 ETAS 和 3 个重复的核心

TACAT 及 2 个反向互补序列 ATGTA。银色裂腹鱼线粒体控制区的中央保守区 CSB-F、CSB-E、CSB-D、CSB-B 特征序列均相对比较保守,与已报道的近缘种序列结构相似^[43-44],其中 CSB-E 关键序列中含有 GTGGG-box^[29]。

目前,关于裂腹鱼类的系统发育研究多基于线粒体 Cyt *b* 或 *COI* 等单个基因^[45-47],考虑到只有 *ND6* 基因位于 L 链上,且其碱基组成的异质性容易导致系统发育分析效果不佳^[22],因此本研究使用 H 链编码的 12 个 PCGs 基因串连序列,基于 ML 和 BI 方法构建了部分裂腹鱼类系统发育树。从系统发育关系来看,银色裂腹鱼所处的族群 IV 的分化时间较早,与其他裂腹鱼类的亲缘关系较远,暗示银色裂腹鱼或为同属鱼类中较为原始的类群,与海萨·艾也力汗等^[48]基于 *COI*、*ND4* 和 16S rRNA 基因联合数据得到的结果一致,这可能与天山山脉的形成和阻隔效应造成的栖息地生态环境差异有关。族群 I 多分布于金沙江、怒江、澜沧江水域,这 3 条水系均发源于青藏高原唐古拉山脉,且在云南省内并行奔流,推测相似的水文和环境因素使得这里分布的裂腹鱼类在进化方向也存在相似之处,故其亲缘关系也较为密切。但随着流域范围的延伸,所处的地理环境不断改变,因而生活在这 3 条水系的裂腹鱼类在不断适应进化中演化出了种类繁多的族群 I。族群 II 栖息的水系发源于青藏高原的印度河流域,其中尼泊尔裂腹鱼主要分布于尼泊尔最大的淡水湖——拉拉湖(Rara Lake),因此与该族群其他裂腹鱼类亲缘关系较远,位于这一支系的基部,与迪丽娜·茹斯坦木等^[49]的研究结论一致。族群 III 分布在雅鲁藏布江及其支流,尽管在地域上与族群 II 相近,但由于所生活的水域环境不同,河流彼此间不存在交汇,因而分化为了不同支系。

综上所述,本研究首次开展了银色裂腹鱼这

一珍稀濒危裂腹鱼类线粒体 DNA 基因组结构特征分析,并基于多基因串联序列探讨了其系统发育关系,结合地质构造变化和环境因子的差异分析了银色裂腹鱼与其他水系裂腹鱼类遗传分化的原因。研究结果对于今后进一步开展该类群鱼类的物种进化、分类鉴定和多样性研究提供了新的线索和参考依据。

REFERENCES

- [1] 曹文宣, 陈宜瑜, 武云飞, 朱松泉. 裂腹鱼类的起源和演化及其与青藏高原隆起的关系[C]//中国科学院青藏高原综合科学考察队. 青藏高原隆起的时代、幅度和形式问题. 北京: 科学出版社, 1981, 118-130.
CAO WX, CHEN YY, WU YF, ZHU SQ. Origin and Evolution of Schizothoracinae fishes and the Relationship with the Uplift of the Qinghai-Xizang Plateau[C]//Tibetan Expedition Team of the Chinese Academy of Science Studies. On the period, amplitude and type of the Qinghai-Xizang Plateau. Beijing: Science Press, 1981, 118-130 (in Chinese).
- [2] 代应贵, 肖海. 裂腹鱼类种质多样性研究综述[J]. 中国农学通报, 2011, 27(32): 38-46.
DAI YG, XIAO H. Review of studies on the germplasm resources of the schizothoracinae fishes[J]. Chinese Agricultural Science Bulletin, 2011, 27(32): 38-46 (in Chinese).
- [3] 王德忠. 新疆的裂腹鱼亚科鱼类研究[J]. 干旱区研究, 1998, 15(4): 26-32.
WANG DZ. The schiz. othoracinae fishes in Xinjiang[J]. Arid Zone Research, 1998, 15(4): 26-32 (in Chinese).
- [4] 任慕莲. 伊犁河鱼类[J]. 水产学杂志, 1998, 11(1): 7-17.
REN ML. Fishes of River Yili[J]. Chinese Journal of Fisheries, 1998, 11(1): 7-17 (in Chinese).
- [5] 李思忠. 中国淡水鱼类的分布区划[M]. 北京: 科学出版社, 1981.
LI SZ. Studies on Zoogeographical Divisions for Fresh Water Fishes of China[M]. Beijing: Science Press, 1981 (in Chinese).
- [6] 郭焱, 张人铭, 蔡林钢. 新疆鱼类志[M]. 乌鲁木齐: 新疆科学技术出版社, 2012.
GUO Y, ZHANG RM, CAI LG. Xinjiang Ichthyology[M]. Urumqi: Xinjiang Science and Technology Press, 2012 (in Chinese).

- [7] 郭焱, 蔡林刚, 张人铭, 阿达克. 伊犁河(中国境内)土著鱼类的分布与演化[J]. 干旱区研究, 1999, 16(2): 31-35.
GUO Y, CAI LG, ZHANG RM, ADAK K. The native fisher distributed and made of evolution in the Yili River (China's section)[J]. Arid Zone Research, 1999, 16(2): 31-35 (in Chinese)
- [8] INOUE JG, MIYA M, TSUKAMOTO K, NISHIDA M. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences[J]. Molecular Phylogenetics and Evolution, 2001, 20(2): 275-285.
- [9] HARRISON RG. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology[J]. Trends in Ecology & Evolution, 1989, 4(1): 6-11.
- [10] BOORE JL. Animal mitochondrial genomes[J]. Nucleic Acids Research, 1999, 27(8): 1767-1780.
- [11] MAMIATIS T, FRITSCH EF, SAMBROOK J, ENGEL J. Molecular cloning-a laboratory manual. New York: Cold Spring Harbor Laboratory. 1982, 545 S., 42 \$[J]. Acta Biotechnologica, 1985, 5(1): 104.
- [12] LI RQ, ZHU HM, RUAN J, QIAN WB, FANG XD, SHI ZB, LI YR, LI ST, SHAN G, KRISTIANSEN K, LI SG, YANG HM, WANG J, WANG J. *De novo* assembly of human genomes with massively parallel short read sequencing[J]. Genome Research, 2010, 20(2): 265-272.
- [13] ZHAO QY, WANG Y, KONG YM, LUO D, LI X, HAO P. Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study[J]. BMC Bioinformatics, 2011, 12(14): 1-12.
- [14] DIERCKXSENS N, MARDULYN P, SMITS G. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data[J]. Nucleic Acids Research, 2017, 45(4): e18.
- [15] LOWE TM, CHAN PP. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes[J]. Nucleic Acids Research, 2016, 44(W1): W54-W57.
- [16] CONSORTIUM TR, SWEENEY BA, PETROV AI, BURKOV B, FINN RD, BATEMAN A, SZYMANSKI M, KARLOWSKI WM, GORODKIN J, SEEMANN SE, CANNONE JJ, GUTELL RR, FEY P, BASU S, KAY S, COCHRANE G, BILLIS K, EMMERT D, MARYGOLD SJ, HUNTLEY RP, et al. RNACentral: a hub of information for non-coding RNA sequences[J]. Nucleic Acids Research, 2019, 47(D1): D221-D229.
- [17] KUMAR S, NEI M, DUDLEY J, TAMURA K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences[J]. Briefings in Bioinformatics, 2008, 9(4): 299-306.
- [18] GUINDON S, GASCUEL O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood[J]. Systematic Biology, 2003, 52(5): 696-704
- [19] HUELSENBECK JP, RONQUIST F. MRBAYES: Bayesian inference of phylogenetic trees[J]. Bioinformatics, 2001, 17(8): 754-755.
- [20] MARK AM, TERRI S, WAYNE P. Embedding CIPRES science gateway capabilities in phylogenetics software environments[P]. Extreme Science and Engineering Discovery Environment, 2013.
- [21] POSADA D. jModelTest: phylogenetic model averaging[J]. Molecular Biology and Evolution, 2008, 25(7): 1253-1256.
- [22] MIYA M, TAKESHIMA H, ENDO H, ISHIGURO NB, INOUE JG, MUKAI T, SATOH TP, YAMAGUCHI M, KAWAGUCHI A, MABUCHI K, SHIRAI SM, NISHIDA M. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences[J]. Molecular Phylogenetics and Evolution, 2003, 26(1): 121-138.
- [23] BEHURA SK, SEVERSON DW. Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes[J]. Biological Reviews, 2013, 88(1): 49-61.
- [24] 宋洁, 张小芳, 刁雪涛, 金梅. 线粒体 DNA 12S rRNA, 16S rRNA 研究进展 [J]. 安徽农学通报, 2008, 14(9): 42-43, 37.
SONG J, ZHANG XF, DIAO XT, JIN M. Research progress of mitochondrial DNA, 12S rRNA and 16S rRNA[J]. Anhui Agricultural Science Bulletin, 2008, 14(9): 42-43, 37 (in Chinese).
- [25] JAMIE C, SANKAR S, MURRAY S, JAMES C, LISA D, DU YS, BRIAN F, LIN N, LAKSHMI M, KIRSTEN M, NUPUR P, SHANG ZD, YU N, ROBIN G. The comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs: correction[J]. BMC Bioinformatics, 2002, 3(1): 15.
- [26] TOMITA K, UEDA T, WATANABE K. RNA editing in the acceptor stem of squid mitochondrial tRNATyr[J]. Nucleic Acids Research, 1996, 24(24): 4987-4991.
- [27] 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 WH, ZHANG YP. Genetics and evolution of mitochondrial DNA in fish[J]. Acta Hydrobiologica Sinica, 2000, 24(4): 384-391 (in Chinese).
- [28] 刘红艳, 余来宁, 张繁荣. 鱼类线粒体 DNA 控制区的分子结构及应用进展[J]. 水利渔业, 2008, 28(2): 4-8.
- LIU HY, YU LN, ZHANG FR. Molecular structure and application progress of fish mitochondrial DNA control region[J]. Reservoir Fisheries, 2008, 28(2): 4-8 (in Chinese).
- [29] LEE WJ, CONROY J, HOWELL WH, KOCHER TD. Structure and evolution of teleost mitochondrial control regions[J]. Journal of Molecular Evolution, 1995, 41(1): 54-66.
- [30] 李易. 鱼类 mtDNA 控制区的结构研究[J]. 曲靖师专学报, 1996, 15(6): 36-39.
- LI Y. Study on the structure of mtDNA control region in fish[J]. Journal of Qujing Normal College, 1996, 15(6): 36-39 (in Chinese).
- [31] 苏天凤, 江世贵, 马之明, 李夏. 二长棘犁齿鲷线粒体 DNA 控制区结构和进化[J]. 南方水产科学, 2012, 8(6): 1-8.
- SU TF, JIANG SG, MA ZM, LI X. Structure and evolution of mitochondrial control regions in *Evyynniscardinalis*[J]. South China Fisheries Science, 2012, 8(6): 1-8 (in Chinese).
- [32] PERNA NT, KOCHER TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes[J]. Journal of Molecular Evolution, 1995, 41(3): 353-358.
- [33] 钟东, 赵贵军, 张振书, 徐安龙. 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.
- ZHONG D, ZHAO GJ, ZHANG ZS, XUN AL. Advance in the entire balance and local unbalance of base distribution in genome[J]. Hereditas (Beijing), 2002, 24(3): 351-355 (in Chinese).
- [34] 陈娟, 郭新异, 谢玲, 张涓泽, 庞礴, 郭松长, 祁得林. 裂腹鱼亚科不同等级代表物种线粒体基因组全序列测序与分析[C]. 第三届中国西部动物学学术研讨会, 2014: 80-81.
- CHEN J, GUO XY, XIE L, ZHANG XZ, PANG B, GUO SC, QI DL. Sequencing and analysis of complete mitochondrial genome sequence of representative species in different grades of Schizothoracinae[C]. The Third Symposium on Zoology in Western China, 2014: 80-81.
- [35] FRANK AC, LOBRY JR. Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms[J]. Gene, 1999, 238(1): 65-77.
- [36] 李亚莉. 青藏高原三种裂腹鱼线粒体全基因组的测定及分子进化分析[D]. 上海: 复旦大学博士学位论文, 2012.
- LI YL. Determination and molecular evolution analysis of mitochondrial genome of three species of *Schizothorax* in Qinghai-Tibet Plateau[D]. Shanghai: Doctoral Dissertation of Fudan University, 2012 (in Chinese).
- [37] 鲍明明. 伊犁裂腹鱼线粒体全基因组结构分析及微卫星标记开发[D]. 武汉: 华中农业大学硕士学位论文, 2015.
- BAO MM. The structure analysis of complete mitochondrial genome and development of microsatellite markers for *Schizothorax pseudaksaiensis*[D]. Wuhan: Master's Thesis of Huazhong Agricultural University, 2015 (in Chinese).
- [38] LAZAROU M, THORBURN DR, RYAN MT, MCKENZIE M. Assembly of mitochondrial complex I and defects in disease[J]. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2009, 1793(1): 78-88.
- [39] EFREMOV RG, BARADARAN R, SAZANOV LA. The architecture of respiratory complex I[J]. Nature, 2010, 465(7297): 441-445.
- [40] MIYA M, SATO Y, FUKUNAGA T, SADO T, POULSEN JY, SATO K, MINAMOTO T, YAMAMOTO S, YAMANAKA H, ARAKI H, KONDOH M, IWASAKI W. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species[J]. Royal Society Open Science, 2015, 2(7): 150088.
- [41] 连总强, 滚双宝, 李力, 张锋, 肖伟, 吴旭东. 基于第二代测序技术兰州鲇线粒体基因组全序列测定与分析[J]. 水生生物学报, 2017, 41(2): 334-345.
- LIAN ZQ, GUN SB, LI L, ZHANG F, XIAO W, WU XD. Sequencing and analysis of the complete mitochondrial genome of *Silurus lanzhouensis* based on next generation sequencing technologies[J]. Acta Hydrobiologica Sinica, 2017, 41(2): 334-345 (in Chinese).
- [42] LYNCH M. Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes[J]. Molecular Biology and Evolution, 1997, 14(9): 914-925.

- [43] 谢佳燕. 齐口裂腹鱼线粒体 DNA 控制区结构分析[J]. 动物学杂志, 2011, 46(2): 97-101.
XIE JY. Structure analysis of mtDNA control region in *Schizothorax prenanti*[J]. Chinese Journal of Zoology, 2011, 46(2): 97-101 (in Chinese).
- [44] SONG ZB, SONG J, YUE BS. Population genetic diversity of Prenant's schizothoracin, *Schizothorax prenanti*, inferred from the mitochondrial DNA control region[J]. Environmental Biology of Fishes, 2008, 81(3): 247-252.
- [45] QI D, GUO S, TANG J, ZHAO X, LIU J. Mitochondrial DNA phylogeny of two morphologically enigmatic fishes in the subfamily Schizothoracinae (Teleostei: Cyprinidae) in the Qinghai-Tibetan Plateau[J]. Journal of Fish Biology, 2007, 70(sa): 60-74.
- [46] 杨天燕, 孟玮, 张人铭, 郭焱, 海萨, 谢春刚. 2 种珍稀裂腹鱼类线粒体 DNA 部分序列片段的比较分析[J]. 动物学杂志, 2011, 46(3): 47-54.
YANG TY, MENG W, ZHANG RM, GUO Y, HAI S, XIE CG. Comparative analysis of mtDNA gene between two rare schizothoracine fishes[J]. Chinese Journal of Zoology, 2011, 46(3): 47-54 (in Chinese).
- [47] 海萨·艾也力汗, 张人铭, 牛建功, 胡江伟, 刘鸿, 张钰, 吐尔逊·提立瓦尔地, 封永辉. 基于 COI 基因探讨塔里木河流域裂腹鱼属鱼类的亲缘关系[J]. 中国水产科学, 2021, 28(11): 1417-1425.
HAYSA A, ZHANG RM, NIU JG, HU JW, LIU H, ZHANG Y, TURSUN T, FENG YH. Relationship of *Schizothorax* Heckel in Tarim River Basin based on *COI* gene[J]. Journal of Fishery Sciences of China, 2021, 28(11): 1417-1425 (in Chinese).
- [48] 海萨·艾也力汗, 郭焱, 孟玮, 杨天燕, 马燕武. 新疆裂腹鱼类的系统发生关系及物种分化时间[J]. 遗传, 2014, 36(10): 1013-1020.
HAYSA A, GUO Y, MENG W, YANG TY, MA YW. Phylogeny and divergence time estimation of Schizothoracinae fishes in Xinjiang[J]. Hereditas, 2014, 36(10): 1013-1020 (in Chinese).
- [49] 迪丽娜·茹斯坦木, 袁晓倩, 张琪, 韩洁. 基于线粒体基因组数据的裂腹鱼类系统发育研究[J]. 中国水产科学, 2022, 29(6): 781-791.
RUSTAM D, YUAN XQ, ZHANG Q, HAN J. Study on the phylogeny of Schizothoracids based on complete mitochondrial genome[J]. Journal of Fishery Sciences of China, 2022, 29(6): 781-791 (in Chinese).

(本文责编 郝丽芳)