



实验红鲫线粒体 DNA 条形码特征分析及应用

王余德¹, 李登科², 陈科洁², 吴端生^{1,2*}, 刘少军¹

(1. 湖南师范大学生命科学学院, 省部共建淡水鱼类发育生物学国家重点实验室, 湖南长沙 410081;

2. 南华大学实验动物学部, 湖南衡阳 421001)

摘要: 为探索建立具有实验红鲫品系特异性的线粒体 DNA 条形码, 对鲫属品种进行鉴定。本研究以实验红鲫和鲫属其他鱼类(洞庭湖鲫、日本白鲫、大眼鲫、兰氏鲫、金鱼、湘军金鱼和湘军花鲫)为测试群体, 利用 Sanger 测序对所有个体的线粒体 *CO I*、*CO III*、*Cytb* 这 3 个基因和 D-loop 进行测序, 同时构建 3 个基因的拼接序列, 并分析它们的序列特征和 10 种鲫属鱼类分子系统进化的聚类特征。结果显示, *CO I*、*CO III*、*Cytb* 这 3 个基因的拼接序列和 *CO I* 基因的进化树聚类能明显地发现 10 种不同的鲫属鱼类聚成不同的小支, 而 *CO III* 基因和 D-loop 序列聚类不能同时区分 10 种鲫属鱼类。研究表明, 通过 *CO I* 基因以及 3 个基因的拼接序列均能有效地鉴别实验红鲫与鲫属其他鱼类, 故该类序列能作为实验红鲫的有效分子标记, 同时 *CO I*、*CO III*、*Cytb* 和 D-loop 序列构建的进化树均可以反映 10 种不同鲫属鱼类之间的进化关系。

关键词: 实验红鲫; 线粒体 DNA; DNA 条形码; 系统进化树

中图分类号: Q 785; S 917.4

文献标志码: A

DNA 条形码是指生物体内能够代表该物种的、标准的、相对较短的 DNA 片段作为物种标签, 已在生物医药、食品科学、农业科学、生态学等许多研究领域中的应用, 尤其是应用于物种快速准确鉴定^[1-3]。目前, 已经报道了条形码在一些生物分类鉴定中的应用^[2]。Lopes 等^[4]应用 DNA 代谢条形码分析 2 种巴西栉鼠 (*Ctenomys brasiliensis*) 亚种的食物种类差别。Liu 等^[5]建立了一种通用的 DNA 条形码 ITS2 序列, 以识别中国药典中的有毒药用植物及其有毒相关物种或掺假物。DNA 条形码技术研究中至关重要的一步是选择合适的基因片段, 该片段需具备: ①包含能够用于种间区分的遗传变异; ②具有高度保守的区域, 便于开发可用范围更为广泛的通

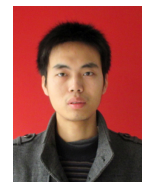
用引物; ③序列长度适中, 以利于 DNA 的扩增及测序; ④含有一定的物种进化信息来确定物种的分类地位^[6]。与核 DNA 相比, 线粒体基因占有更多优势。线粒体基因 *CO I*、*CO III*、*Cytb* 和 D-loop 序列被认为是动物 DNA 条形码首选的靶标基因。其一, 这些基因编码的氨基酸序列的演化速率比较慢, 可以区分出分化时间较久的物种。其次, 在 DNA 序列上又有足够的变异能够区分亲缘关系较近的姊妹种, 所以被广泛应用于动物中不同分类阶元水平的分子系统学研究中。Fernandes 等^[7]用 *CO I* 和 *Cytb* 基因来鉴定 4 种鳕的遗传变异, 得出 *Cytb* 序列能够以 99.3% 的高置信度对所研究鱼类进行有效识别。Huemer 等^[8]采集约 690 种奥地利夜蛾(鳞翅目,

收稿日期: 2019-11-04 修回日期: 2020-06-13

资助项目: “十二五”国家科技支撑计划(2015BAI09B05); 湖南省科学技术厅科技计划重点项目(2014TT1002); 中国博士后基金(2019M662788)

第一作者: 王余德(照片), 从事鱼类遗传育种研究, E-mail: 461669793@qq.com;
李登科, 从事实验动物标准化研究, E-mail: 2608704215@qq.com

通信作者: 吴端生, E-mail: wuduansheng@aliyun.com



Lepidoptera) 的 671 个物种 CO I 基因, 构建了多样性丰富的奥地利夜蛾的 DNA 条形码文库, 依此分类的结果与 Linnean 分类法几乎完美匹配。

理论上, 进行物种鉴定的前提是确定遗传距离阈值, 并将待测物种与已知物种群体进行匹配。Hebert 等^[9]提出, 同一物种内的遗传距离应小于 2%, 大多数物种的种内遗传距离普遍小于 1%, 并推荐将 2% 和 10 倍的种内遗传距离作为划分物种的标准阈值。线粒体 DNA 分子结构简单、几乎不发生重组、呈现母系遗传、进化速率快, 目前在系统进化、分类学、群体遗传以及动物保护生物学等研究方面得到普遍应用。近年来常用的分子标记有 *Cytb*、*ND*、*CO I*、*CO II* 和 D-loop 等^[10]。D-loop 区域在线粒体基因组中的碱基突变率最高, 长度变异最大, 其碱基替换率比 mtDNA 基因组的其它区域高 5~10 倍, D-loop 序列比较适合于近缘种以及种内不同群体之间的遗传差异分析。以线粒体条形码鉴别物种, 通常是利用单个基因来鉴别, 但常常碰到对亲缘关系特别近的物种难以区分的问题^[11]。

此外, 有的学者曾采用多种条形码标记联合使用, 如 2009 年国际生物条形码协会 (CBOL) 植物工作组对来自 550 个物种 907 个样品的 7 个序列 (*rbcL*、*matk*、*rpoC1*、*ropB*、*psbA-trnH*、*psbK-psbI* 和 *atpF-atpH*) 进行分析比较, 提出将 *matk+rbcL* 片段组合作为植物 DNA 条形码的核心序列^[12]。对有效识别亲缘关系很近的动物品种或品系的 DNA 条形码起到了重要作用。随着斑马鱼 (*Danio rerio*) 等鱼类实验动物资源的开发与利用, 实验鱼类越来越多地被用于发育生物学研究、水生态毒理学研究、水环境污染监测、药物与化学品环境安全性评价等领域。红鲫是一种鲫属鱼类, 在实验动物研究方面具有重要意义。实验红鲫 (*Carassius auratus*, red variety) 就是由红鲫驯化而来的^[13]。目前, 已经对实验红鲫的生物学特性进行了研究^[14], 建立了一些遗传标记, 并颁布了地方标准^[15]。为进一步完善实验红鲫的分子遗传标记, 又开展了本研究。

本研究以线粒体基因 (*CO I*、*CO III*、*Cytb*) 和 D-loop 序列为对象, 通过比较鉴别实验红鲫及同属的 9 种近缘鱼类, 筛选出具有实验红鲫品系特异性的线粒体 DNA 条形码标记, 并以此研究实验红鲫与其同属近缘鱼类之间的系统进化

关系, 从而进一步完善实验红鲫的遗传背景资料, 为实验红鲫种质遗传质量控制标准的建立提供数据支撑, 对我国实验鱼类资源库的建设具有重要的学术意义。

1 材料与方法

1.1 实验对象

实验红鲫 C1HD 系, 2 龄, 20 尾, 雌雄各半, 由南华大学实验动物学部培育并供应; 洞庭湖鲫, 2 龄鱼, 20 尾, 雌雄各半, 购买于长沙某市场; 金鱼、湘军金鱼及湘军花鲫为 2 龄鱼, 各 6 尾, 雌雄各半; 雌性日本白鲫 (*C. cuvieri*) 1 尾, 均由湖南师范大学省部共建淡水鱼类发育生物学国家重点实验室提供。

在 NCBI 数据库中下载其他鱼类的 DNA 序列, 用于 DNA 条形码比较分析。相关物种的 GenBank 序列号: 实验红鲫, MN502672; 日本白鲫, AB045144; 银鲫 (*C. gibelio*), GU138989; 大眼鲫 (*C. auratus grandoculis*), KJ874430; 兰氏鲫 (*C. auratus langsdorfii*), LC019789.1。

1.2 仪器与试剂

PCR 仪 (Eppendorf 公司, 22331 Hamburg, 德国); 台式高速冷冻离心机 (Eppendorf); -80 °C 冰箱 [SANYO 公司, MDF-382E(N), 日本]; 恒温培养箱 (北京中兴伟业仪器有限公司, DH-500A6, 中国); 台式恒温振荡器 (CRYSTAL 公司, IS-RSD3, 美国); 琼脂糖凝胶电泳 (北京百晶生物技术有限公司, BG-Power 600, 中国); 凝胶成像仪 (SYNGENE 公司, DR4V2/3255, 美国)。

pMD18-T vector [宝日医生物技术 (北京) 有限公司, SD4038]; 2×Taq PCR Mastermix [天根生化科技 (北京) 有限公司, KT201-02]; SanPrep 柱式 DNA 胶回收试剂盒 [生工生物工程 (上海) 股份有限公司, B518131-0100]; 全基因组 DNA 提取试剂盒 [宝日医生物技术 (北京) 有限公司, 9765]。引物由生工生物工程 (上海) 股份有限公司合成。

1.3 基因组 DNA 提取

实验样本按照 TaKaRa 全基因组 DNA 提取试剂盒步骤抽提 DNA。通过琼脂糖凝胶电泳和核酸分析仪, 检测所提取 DNA 的质量及浓度。

1.4 PCR 扩增

PCR 总扩增体系为 20 μL , 包括 2 \times Taq PCR Mastermix 10 μL , 模板 DNA 2 μL , 正反向引物各 1 μL (表 1), ddH₂O 6 μL 。PCR 扩增条件: 94 $^{\circ}\text{C}$

预变性 5 min; 94 $^{\circ}\text{C}$ 变性 30 s、退火 30 s、72 $^{\circ}\text{C}$ 延伸 (COI、COIII 和 D-loop 均延伸 60 s, Cytb 延伸 80 s), 循环 35 次; 最后 72 $^{\circ}\text{C}$ 延伸 10 min, 扩增产物于 16 $^{\circ}\text{C}$ 保存。

表 1 实验红鲫线粒体 DNA 条形码 PCR 引物

Tab. 1 Mitochondrial DNA barcode PCR primer of laboratory *C. carassius*, red variety

基因 gene	引物序列(5'→3') primer sequence (5'→3')	退火温度/ $^{\circ}\text{C}$ T_m	PCR产物/bp PCR product
CO I	CACAAAGACATTGGCACCCCT AGTCAGCTGAACACTTTTAC	51	940
COIII	ATGGCCACCAAGCACATGC TATGAGCCTCATCAATAGAT	51	788
Cytb	ATGGCAAGCCTACGAAAAAC AAGCTCATTTCAAGTGCTTA	51	1 143
D-loop	GGATAACCATCCCTGTATGG TAAATAATGTTAGAAAATT	46	923

1.5 目的基因产物检测、连接、转化及测序

PCR 扩增产物利用 1.5% 琼脂糖凝胶检测并回收, 连接到 pMD18-T 载体, 再转化到大肠杆菌 DH5 α 感受态细胞中之后涂布于培养皿中, 于培养箱中倒置培养 12~16 h, 再挑选单克隆菌落进行阳性检测, 并送生工生物工程(上海)股份有限公司进行测序。

1.6 数据分析

对各物种条形码测序结果首先依据 abi 峰图检查测序质量, 利用 Bioedit 软件 (version 7.0) 查找并去克隆接头, 拼接双向序列, 进行 BLAST 比对。利用 Mitofish (<http://mitofish.aori.u-tokyo.ac.jp/annotation/input.html>) 在线软件获得实验红鲫线粒体的 DNA 注释信息。

运用 MEGA 6.0 软件的 Kimura 双参数模型法 (Kimura 2-Parameter, K₂P) 计算遗传距离, 并依据遗传距离分别构建 NJ 树。自举检验 (bootstrap test) 1 000 次重复估计各分支的置信值。

2 结果

2.1 实验红鲫 COI 基因及拼接基因序列特征

实验红鲫 CO I 基因序列全长为 1 551 bp (GenBank no: MT559523) (图 1-a), 其中碱基 A 占

27.53%, 碱基 C 占 25.73%, 碱基 G 占 17.41%, 碱基 T 占 29.34%, A+T 含量为 56.87%, G+C 含量为 43.13%。编码 516 个氨基酸 (图 1-a)。由 CO I、COIII 及 Cytb 基因构成的拼接序列全长为 3 477 bp, 命名为 PCCC。其碱基 A 占 28.16%, 碱基 C 占 27.01%, 碱基 G 占 16.05%, 碱基 T 占 28.79%, A+T 含量为 56.95%, G+C 含量为 43.06%。碱基 A+T 含量高于 G+C 含量之和。这与脊椎动物线粒体 DNA 的碱基组成特征一致^[16](图 1-b)。

2.2 基于 CO I、COIII、Cytb 和 D-loop 序列的遗传距离分析及系统发育树构建

基于 CO I、COIII、Cytb 及 D-loop 序列的 10 种鱼类的 K₂P 遗传距离及进化树分析表明, CO I 基因序列能够对实验红鲫与其他 9 种鲫属鱼类进行区分 (表 2, 图 2-a)。10 种鲫属鱼类的种内遗传距离为 0.0%~0.7%, 均值为 0.18%, 其中实验红鲫和洞庭湖鲫的种内遗传距离为 0.0%, 湘军花鲫的种内遗传距离最大 (0.7%), 均值为 0.18%。10 种鲫属鱼类的种间遗传距离为 0.2%~13.8%, 均值为 5.21%, 其中金鱼与实验红鲫之间的遗传关系最近, 遗传距离为 0.2%。锦鲤 (*Cyprinus carpio* var. *koi*) 与大眼鲫、兰氏鲫之间的关系最远, 遗传距离为 13.8%。以锦鲤为外类群, 采用 NJ 法构建进化树显示, 外类群锦鲤独立成支,


```

COI
1 11 21 31 41 51 61 71
1 1 6909CAATCCGCGGATCTTCTTACCAACCAAGAACATGGTACCCCTTATCTAGTATG90CTG6A0CC9C
1 (+1) V A I T R Y F F S T H R K E D I G L I L Y L V F G A Y A G
1 AATAGTGGACCGCTTAAAGCTCTCCTCCTCAGCTGACATTAAGCAACCCGGGATCACTTCTAGGTGGGCAAAATT
28 (+1) M Y G T A L I R X A E L S L A C C P G S L L G E D Q M Z I
81 ACAATGTAATGTATGCCCCAGCCCTTGTAAATAATTTTCTTAAGTAAATGCTTCTTCAATGGAGGATGGAAC
54 (+1) Y R V I T Y T A H A F V M I T F F M V H P I L I G G F G W
161 TGACTGTACCCCTGATTAATGCGAGCTCAGACATG6EATFDDCAGAAATAAACAATTAAGCTTCTGACTTCTTCCDC
81 (+1) V L V P L M I G A P D N A P P R N N N N S F N L L P F
321 ATCATCTCTGTACTACTACTCTCTCCTCGGTGTGAAAGCCGAGACTGGCACDGGATGGACATATACCCCTCTGAC
108 (+1) S F L L L L A S S G V E A G A G T G R T V Y F P L A
401 GAAGCTGGCCACGAGGAGCAGTAGACATTAACAATTTTCCTACATCAITTAAGAGGTGTTTCATCAGCCGCG
134 (+1) G E L A R A G A S V E L T T F S L H L A G V S S I L G
481 GCAATCACTCTATTACTACCACTAACTAAACACTCCAGCATTTCCCACTCCCAACCCCTATTGTGTGAT
161 (+1) A T N F I T T T N H E P P A S I Q Y T F L F V Y V S
561 CGACTGTGACCGGCTGCTTCTCTCTATCACTACTGTTCTAGCTGCCGTATTACAATGTTTAAACAGATGAA
188 (+1) V L V T A V L L L L S L P V L A A G I T M L L T D R
641 ATCTAACACCACTCTTGTATCCCGCAAGCGGAGACCAATTCCTTAAACCACTACTTGTATCTTGTGTGAC
214 (+1) M L N T Y T F D P A G G G B P I L I Q R L F V Y F F G H
721 CUNAGATGTCATTTTATCTTCGAGGTTTGGAAATATTTCACGTGTGAGGATATATTCAGTAAAGAAAGAC
241 (+1) F F V Y T T T G P F G T T C H V A Y T R Y Y Y Y F F
801 ATTTGGTATATAGGAATGTATGACCCATAAGGCCAATG90CTTCAAG9TTCATGTATGAGCCCAACATATGTTA
268 (+1) F G Y N G W V Y A M T A H G L L C F I V W A H H H F
881 CTGTGGATGGADSTAGACACDCCGTCATATTTTACATCCGCAACATAATTAAGCAATTCUACCGG9TFAAAA9TA
294 (+1) T V G H B V H I R A I F S A T F I V G Y K V
961 TTAGCTGACTGGCTACACTTCAAGAGGATCAATTAAGAGAAACCAUACTATAGAGCTAGGATCAITTTTCT
321 (+1) F S W L A T L H G G S I E X W E T P N L H W A L G F I F L
1 041 GTTACAGTGGAGGACTTACAGGAATGTCTTCTTAACTCACTGTATATGTCTTCCAGCCACTATATATGAG
348 (+1) F T V G G L T G T V L S W S S L D I V L H D T Y T V
1 121 TAGCACATTTCTACTATGTACTATCAATGGTGGCTATGTGCAATATAGAGACTTTGTACTGATGCTTCCCTCA
374 (+1) V A H F H Y V L S H G A V P A I H A A P V H H P F L L
1 201 AGGGGTACACTGTCATAGGCTTGAACAAAATTCCTTGTGGGTTATATTTATCGAGATTAACCTCACTTCTTCC
401 (+1) T G Y T L H S A W T K I H F G V H F I G V H L T F P P
1 281 ACAACACTTCTG90CTAGGAGGATACACAGC9TATCTGATATCCAGGACTTATGCCCTAGAAATCAGTAT
428 (+1) Q H F L G L A G H F R R I T S B Y P A T A L W A L W T V
1 361 CACTAATGCTTCTTAACTCTTCTGAGGCTGATTAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
454 (+1) S S I G S L I S L Y A V I W F L F I L W X A F A A K R
1 441 GAGTGTATCTGAGACATCAACAATAGTGGAACTCCATGGTCCCTCTCTTCAACCAATGAGGAA
481 (+1) H V L S V E L T M T V H V E W L H G C P F P Y K T Y E E
1 521 ACAAGCATTTGTCAGATCAATCAAAATTA
508 (+1) P A F V Q T Q S H *

```

(a)

```

PCCC
1 1 1056CAATCCGCGGATCTTCTTACCAACCAAGAACATGGTACCCCTTATCTAGTATG90CTG6A0CC9C
81 AATAGTGGACCGCTTAAAGCTCTCCTCCTCAGCTGACATTAAGCAACCCGGGATCACTTCTAGGTGGGCAAAATT
161 ACATGTAATGTTTACGGCCACCGCTTCTGTAATTTTCTTAAAGTAAAGCTTACTTGGAGGATGGAGAAC
241 TGACTGTACCCCTGATTAATGCGAGCTCAGACATG6EATFDDCAGAAATAAACAATTAAGCTTCTGACTTCTTCCDC
321 ATCATCTCTGTACTACTACTCTCTCCTCGGTGTGAAAGCCGAGACTGGCACDGGATGGACATATACCCCTCTGAC
401 GAAGCTGGCCACGAGGAGCAGTAGACATTAACAATTTTCCTACATCAITTAAGAGGTGTTTCATCAGCCGCG
481 GCAATCACTCTATTACTACCACTAACTAAACACTCCAGCATTTCCCACTCCCAACCCCTATTGTGTGAT
561 CGACTGTGACCGGCTGCTTCTCTCTATCACTACTGTTCTAGCTGCCGTATTACAATGTTTAAACAGATGAA
641 ATCTAACACCACTCTTGTATCCCGCAAGCGGAGACCAATTCCTTAAACCACTACTTGTATCTTGTGTGAC
721 CUNAGATGTCATTTTATCTTCGAGGTTTGGAAATATTTCACGTGTGAGGATATATTCAGTAAAGAAAGAC
801 ATTTGGTATATAGGAATGTATGACCCATAAGGCCAATG90CTTCAAG9TTCATGTATGAGCCCAACATATGTTA
881 CTGTGGATGGADSTAGACACDCCGTCATATTTTACATCCGCAACATAATTAAGCAATTCUACCGG9TFAAAA9TA
961 TTAGCTGACTGGCTACACTTCAAGAGGATCAATTAAGAGAAACCAUACTATAGAGCTAGGATCAITTTTCT
1 041 GTTACAGTGGAGGACTTACAGGAATGTCTTCTTAACTCACTGTATATGTCTTCCAGCCACTATATATGAG
1 121 TAGCAATTTCTACTATGTACTATCAATGGTGGCTATGTCAGGATTAAGGAGGCTTGTACTTCTTCTTCTTCT
1 201 AGGGGTACACTGTCATAGGCTTGAACAAAATTCCTTGTGGGTTATATTTAICGGAGTTAACCTCAGATCTTCC
1 281 ACAACACTTCTG90CTAGGAGGATACACAGC9TATCTGATTAAGGAGGCTTACGCCCTTAAAGATACA9TAT
1 361 CACTAATGCTTCTTAACTCTTCTGAGGCTTATGTCATTTTCTTAACTCACTGATTAAGGAGGCTTGGCCCT
1 441 GAGTGTATCTGAGACATCAACAATAGTGGAACTCTATGGCGCCDCCCTCTTACCACACATACAGGAA
1 521 ACAAGCATTTGTCAGATCAATCAAAATTAAGTGGCCACCAAGCAGCATG6EATFDDCAGAAATAAACAAT
1 601 CACTAAGCGAGCTATGTCGCTTCACTAACTCAAGCAGCTTGTACTTCACTTCCACTCAACCAATTAATA
1 681 ACITTAGGACTAATTTCTCTACTTCTTACTATGTATCAATGATGAGTGCATCAITTTGGAGAGGAACTTCCAGGCTT
1 761 CCACACACCCCGAGTCAAAAAAGCACTACGATADGGAATAATCTTATTATCTCCCGGAAGTATCTTCTCTAGGAT
1 841 TCTCTGAGCATCTACCCTCGAGTGTAGCACTACCCCAAGACTAGGAGGATGCTGACCTCCACAGGAATCAACCC
1 921 CTGACCCCAITTTGAAAGTCCCTCTTCAACACAGCTGTACTATTAGCACTCCGGGTTACAGTTACATGAGCC
2 001 CATTAAGAGGGGAGCAAAAACAAGCTTCAATCTTACGATTAACCACTTACTTGGACTTCTTACTGCTTCC
2 081 AAGCTAAGGAATATTATGAGGACDCTTCAACAATTCAGAGAGGATATAGCTCAACATCTTTGTAGTACAGGAT
2 161 CAI9GACTCATGTTATTATTGGATCAACCTTCTAGCAGTATGCTTCAAGCCAAAATCAACACTTACTTACAT
2 241 ACACATTTGGCTTGAAGCCCTGCTGATCTGACTTGTGAGGATGATAGCAATTTCTTACTGCTTACTTCT
2 321 ATTTAGAGGCTCATAATGGCAGCTTACGAACAAACACCCCTCATTAAGATGCTATGAGCAGTATGAGCTAC
2 401 CCACACATCCACATTTGAGCATGATGAACCTTGGCTCTTACTAGGATATGCTTAACTCAATTTCAADCCG
2 481 CTATTCTAGCTATACATTAACCTCAGACATTTCAACCGCATCTTCACTGCTTCAACCACTGCGGAGATGAA
2 561 C9GCTGACATTTGTAATATTCAGCCAAATGGAGCTCATCTCTTCTCTGTTTACTACATATGCGGAGGCG
2 641 TATATATGAGTCACTTAAACAAGAAACCTGAACATTTGGATGTTCTTCTACTCTGATTTAATGAGCAGCCTT
2 721 GTGGTATGTTCTTCCATGAGGCAATGCTTCTTGGGGCTACAGTAATCAACAACCTTATCCCGCGTGCATA
2 801 TATAGGAGATATATGTTCAATGAATTTGAGGAGCTTCTCCGTAGACAAATGCAACATAACACGATCTTTG
2 881 ACTTCTTCAACATCAITATGCGCCGCTTCTGCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
2 961 ATCTGACTGACTACAGCCAGCAAAATTTCTTCCACCAACTTTTATCAAAAGGACTCTTGGGTTGTTGTAAT
3 041 ACTACTGACCTTCACTCTTGGCATAATCTCCCAAACTTTTGGAGAGCCAGAAACCTTCACTCCAGCCATCC
3 121 TGGTATCCCTCCATATAATTAACAGAGTATATCTTGTGTTGCTACGGCACTTCCGATCAATTTCCCAAACT
3 201 GGAGGAGTCTTGCACACTATCTTCCATCTGACTAATAGTGGTACCCCTTACTACATACTTAAACAAGAGGACT
3 281 AAGTCTCCCCCACTACCCAACTCTTATTGCACTTATGTCAGCAGCATGATTTCTGACATGAATGGAGGAA
3 361 CAGTAAACATCTCATATCATGAGCAAAATGCTATGCTATATTTGGACTGTTCTTCTTCTTCTTCTTCTTCT
3 441 GCGAGATGATTAAGAAATAAAGCACTGAAATGAGCTT

```

(b)

图 1 实验红鲫 COI 基因及拼接基因序列

(a) 实验红鲫 COI 基因碱基序列和氨基酸序列, (b) 实验红鲫拼接 DNA 序列全长

Fig. 1 COI and splicing gene sequence of laboratory *C. carassius*, red variety

(a) COI gene sequence of CIHD and amino acid sequence of CIHD, (b) full length of the CIHD concatenating sequence

其他鲫属鱼类聚为一大支。其中实验红鲫与金鱼聚为一支, 在小支上存在一定差异, 大眼鲫与兰氏鲫聚为一支, 湘军花鲫和湘军金鱼聚为一支(图 2-b)。

基于 COIII 基因序列的遗传距离分析及进化树构建结果显示, 实验红鲫可与湘军金鱼、湘军花鲫、洞庭湖鲫、银鲫等区别, 但不能与金鱼区分(表 2, 图 2-b)。所研究鱼类的种内遗传距离为 0.0%~0.2%, 平均值为 0.1%; 种间遗传距离分布为 0.2%~10.6%, 平均为 3.8%, 是种内遗传距离的 38 倍。要进行条形码分析, 必须满足种间遗传距离超过 10 倍的种内遗传距离, 因此,

该序列可用于 DNA 条形码分析。在种内遗传距离中, 湘军花鲫和银鲫 2 种鱼类的最小(0.0%); 金鱼、实验红鲫和洞庭湖鲫的最大, 均为 0.2%。亲缘关系最远的是日本白鲫与锦鲤, 种间遗传距离达到 10.6%; 金鱼与实验红鲫之间遗传距离最小, 为 0.2%, 且二者与其他几种鱼类之间的种间遗传距离基本相同。用 K_2P 法构建基于 COIII 基因序列的进化树显示, 除实验红鲫和金鱼外, 其他相同鱼类均聚为一分支; 锦鲤作为外类群, 形成一个单系分支; 日本白鲫与大眼鲫和兰氏鲫聚为一大支; 湘军金鱼、湘军花鲫、洞庭湖鲫、银鲫、实验红鲫和金鱼聚在同一大

表 2 基于 *CO I*、*CO III*、*Cytb* 和 *D-loop* 的遗传距离分析 (100%)Tab. 2 Genetic distance analysis based on *CO I*, *CO III*, *Cytb* and *D-loop*

物种 species	金鱼 GF	实验红鲫 C1HD	湘军金鱼 GF-L	洞庭湖鲫 LCC	湘军花鲫 RCC-L	日本白鲫 JCC	银鲫 SCC	兰氏鲫 CAL	大眼鲫 CAG	锦鲤 KOC
金鱼 GF	0.1/0.2/0.0/0.0									
实验红鲫 C1HD	0.2/0.2/0.0/0.2	0.0/0.2/0.0/0.2								
湘军金鱼 GF-L	1.4/0.9/2.1/2.1	1.2/0.9/2.1/2.0	0.1/n/n/n							
洞庭湖鲫 LCC	0.6/0.7/1.4/1.1	0.4/0.7/1.4/1.0	0.7/0.1/1.4/1.0	0.0/0.2/0.2/2.4						
湘军花鲫 RCC-L	1.1/0.9/2.2/---	1.0/0.9/2.2/---	0.4/0.0/0.0/---	0.6/0.0/2.1/---	0.7/0.0/0.1/---					
日本白鲫 JCC	4.6/4.6/6.8/6.6	4.6/4.6/6.8/6.5	4.9/4.6/7.4/5.5	4.6/4.3/7.3/6.8/	4.9/4.6/7.4/5.2	0.3/0.1/0.1/0.3				
银鲫 SCC	1.0/0.5/1.8/2.6	1.0/0.5/1.8/2.6	1.3/0.6/2.2/2.3	0.8/0.8/2.2/2.8	1.3/0.6/2.2/1.9	4.4/4.0/7.0/5.8	0.1/0.0/0.4/0.5			
兰氏鲫 CAL	3.2/3.6/6.1/4.3	3.1/3.7/6.1/4.3	3.6/3.6/6.5/3.9	2.8/3.4/6.6/4.8	3.5/3.7/6.6/3.7	4.9/3.0/6.8/4.5	3.2/3.1/6.3/4.4	n/n/n/n/		
大眼鲫 CAG	3.2/3.5/6.4/4.6	3.1/3.5/6.4/4.6	3.6/3.5/6.8/4.4	2.8/3.3/6.9/5.1	3.5/3.5/6.9/4.0	4.9/3.4/6.9/4.6	3.2/3.0/6.2/4.1	0.2/0.6/1.3/4.1	n/n/n/n	
锦鲤 KOC	12.7/9.7/12.9/13.3	12.7/9.7/12.9/13.3	13.0/9.3/12.4/14.1	12.5/9.9/13.1/14.0	12.8/9.3/12.4/13.8	13.2/10.6/16.0/11.6	12.9/9.3/12.7/13.8	13.8/9.3/12.7/12.6	13.8/9.6/13.0/12.5	n

注: n代表不适用。a/b/c/d 数据表示分别用 *CO I*、*CO III*、*Cytb* 和 *D-loop* 序列的计算值。GF. 金鱼, C1HD. 实验红鲫, GF-L. 湘军金鱼, LCC. 洞庭湖鲫, RCC-L. 湘军花鲫, JCC. 日本白鲫, SCC. 银鲫, CAL. 兰氏鲫, CAG. 大眼鲫, KOC. 锦鲤。下同

Notes: n means not applicable. a/b/c/d represents the calculated values of *CO I*, *CO III*, *Cytb* and *D-loop* sequence respectively. GF. goldfish, C1HD. *C. carassius*, red variety, GF-L. goldfish-like fish, LCC. local *C. auratus*, RCC-L. red crucian carp-like fish, JCC. *C. cuvieri*, SCC, *C. gibelio*, CAL. *C. auratus langsdorfii*, CAG. *C. auratus grandoculis*, KOC. *Cyprinus carpio* var. *koi*. The same below

支上, 形成各自的小分支, 但实验红鲫与金鱼聚在一起较难区分 (图 2-b)。

基于 *Cytb* 基因序列的遗传距离分析及进化树分析, 各物种种内遗传距离为 0.0%~0.4%, 平均为 0.13%, 其中金鱼和湘军花鲫的种内平均遗传距离最小 (0.0%), 银鲫的种内遗传距离最大 (0.4%)。根据种间遗传距离分析结果可知, 种间遗传距离分布为 0.0%~16.0%, 均值为 6.2%, 其中锦鲤与日本白鲫之间的遗传距离最大, 为 16%; 在鲫属内, 表现为日本白鲫与洞庭湖鲫之间的遗传距离最大, 为 7.3%。而金鱼与实验红鲫之间的种间遗传距离最小, 为 0.0%, 无法对二者进行有效的区分。用 K_2P 法构建基于 *Cytb* 基因序列的进化树显示, 在鲫属鱼类中, 实验红鲫与金鱼交叉地聚在一起, 且种间遗传距离为 0.0% (图 2-c)。依据线粒体基因调控区的序列信息对 10 种鱼类遗传距离进行分析。每种鱼类的种内变异较小, 表现为种内遗传距离为 0.0%~0.5%, 平均为 0.23%。其中银鲫的种内变异率最高, 其遗传距离为 0.5%; 其次为实验红鲫、洞庭湖鲫和日本白鲫, 种内 K_2P 遗传距离均为 0.3% (表 2)。

对种间遗传距离分析可知, 种间遗传距离为 0.2%~13.9%, 平均值为 5.5%。其中洞庭湖鲫与锦鲤之间的遗传距离最大, 为 14.1%。在鲫属鱼类中, 实验红鲫与日本白鲫之间的亲缘关系最远, 二者种间平均遗传距离为 6.6%。且由于同属鱼的亲缘关系较近, 不同鱼类之间的遗传距离均小于 10%。

用 *D-loop* 序列作进化分析的结果显示, 锦鲤和日本白鲫分别独立为一支, 大眼鲫和兰氏鲫聚为一支, 再与其他鲫聚为一大支, 其置信度为 100%; 实验红鲫的聚类分别与金鱼、洞庭湖鲫聚为一小支, 甚至单独形成一个小分支, 说明实验红鲫在控制区的变异率较高, 无法对 9 种鲫属鱼类进行有效地区分 (图 2-d)。基于 *D-loop* 计算出的平均种内遗传距离均比 *CO I*、*CO III* 和 *Cytb* 这 3 个线粒体基因高。

2.3 基于 *CO I*、*CO III* 和 *Cytb* 拼接序列的系统发育树的构建

以锦鲤为外类群, 基于 *CO I*、*CO III* 和 *Cytb* 编码基因的拼接序列进一步研究鲫属鱼类的系



图 2 基于 *CO I* (a)、*CO III* (b)、*Cytb* (c) 和 *D-loop* (d) 构建的系统进化树

Fig. 2 Phylogenetic tree based on *CO I* gene(a), *CO III* gene(b), *Cytb* gene(c) and *D-loop* sequence(d)

统进化关系。种间遗传距离为 7.3%~16.0%，实验红鲫与金鱼的遗传距离为 0.2%~0.5%。通过将 3 个线粒体基因拼接得到 3 477 bp 的序列用于构建系统进化树，实验红鲫与金鱼均可各自单独形成一小支，二者又可以再聚在一起，然后它们与湘军金鱼、洞庭湖鲫、湘军花鲫和银鲫形成的分支聚合。大眼鲫与兰氏鲫聚合，再与日本白鲫聚在一起形成一个独立的分支，最后 9 种鲫属鱼类成功地聚为一大支(图 3)。

3 结论

在本研究中，发现单个基因不容易区分实验红鲫和金鱼。3 个线粒体基因分别构建的系统

进化树表明，实验红鲫与金鱼的遗传距离小于等于 0.2%。据文献报道，金鱼是由红鲫驯化而来的，红鲫与金鱼亲缘关系很近^[17]。为了获得实验红鲫与金鱼高分辨的线粒体 DNA 分子标记，本研究将基因 *Cytb*、*CO I* 和 *CO III* 的完整序列按照同源位点拼接起来，构建一个线粒体大片段 DNA 序列，作为实验红鲫一种特异性线粒体 DNA 标记，增加了实验红鲫和金鱼在序列终止密码子和起始密码子的差异性，从而增加了整个标记序列的差异性。以它进行聚类分析，结果表明实验红鲫和金鱼聚为不同的支，且遗传距离大于 0.2%，表明该拼接序列标记能大幅度地提高实验红鲫和金鱼的分辨率。拼接序列结

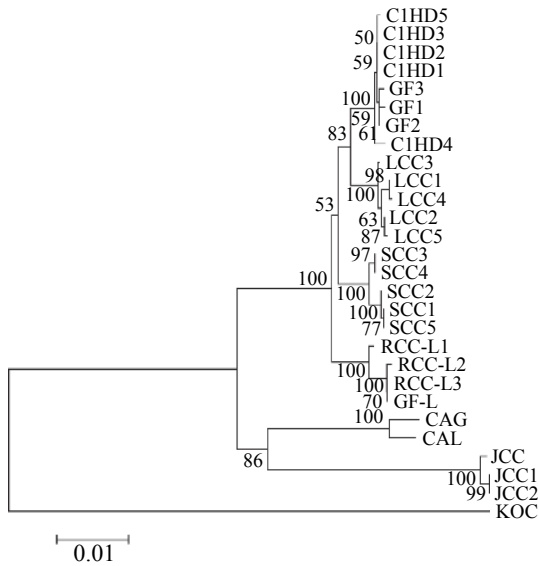


图 3 基于 *CO I*、*CO III* 和 *Cytb* 拼接序列构建的系统进化树

Fig. 3 Phylogenetic tree based on combined sequences of *CO I*, *CO III* and *Cytb*

合了编码序列的保守性以及非编码序列的变异性, 对于物种鉴定具有重要意义。本研究首次报道了多基因拼接序列能够在分子水平上有效地区分亲缘关系很近的物种, 如实验红鲫和金鱼。

基于线粒体基因 *CO I*、*CO III*、*Cytb* 和 D-loop 序列构建了 5 种 NJ 进化树, 其拓扑结构大致相同, 锦鲤作为外类群始终单独聚为一支, 其他 9 种鲫属鱼类聚在同一个大分支上, *CO I* 基因构建的进化树以及拼接序列构建的进化树均能有效地鉴别实验红鲫和金鱼, 然而 *Cytb* 基因扩增的结果表明了实验红鲫和金鱼具有高度的序列同源性。*CO I* 基因扩增的实验红鲫和金鱼序列在不同的位点上, 二者呈现一定数目的碱基位点差异, 进而在进化树上聚在不同的小支上。而拼接序列在实验红鲫和金鱼的序列的比对结果, 显示了二者存在明显的序列片段差异, 进化树上表现为聚在不同的分支上, 可以有效地鉴别实验红鲫和金鱼。D-loop 序列比对及进化树结果显示了实验红鲫个体之间存在差异。通过序列筛选, 已经检测了 *CO I* 基因以及拼接序列可以作为实验红鲫的分子标记。

本研究中, 基于鲫属鱼类线粒体基因组序列构建的进化树结果与单个基因构建的进化树的结果基本一致, 但以线粒体基因组构建的进

化树置信度较高, 均为 100%。线粒体基因构建的进化树可以从一定程度上反映鲫属鱼类之间的进化关系, 实验红鲫与金鱼之间亲缘关系最近, 这与物种分类上的结果一致, 其次是银鲫, 亲缘关系最远的是黑鲫, 与程磊等^[18]的研究成果一致, 说明基于线粒体基因的 DNA 条形码可以用于鲫属鱼类的鉴定。与此同时, 单个基因条形码用于属内种间的鉴定也有其局限性, 由于同属物种之间的亲缘关系较近, 物种间的线粒体基因组的变异较小, 导致种间遗传距离可能未达到种内遗传距离阈值, 比如当依据 *CO I*、*CO III* 和 *Cytb* 单个基因序列分析鲫属鱼类之间的进化关系时, 虽然银鲫可以形成独立的单系支, 实验红鲫与银鲫之间的遗传距离均小于 2%, 但在分类学上银鲫与红鲫属于两个不同的种。而基于拼接序列计算的遗传距离显示银鲫与红鲫之间的遗传距离大于 2%, 表明在属内物种鉴定时, 应选择进化速率较快的条形码进行标记, 且应当使用多个基因进行分析以提高准确性。本研究基于拼接 DNA 序列和 D-loop 序列构建的进化树, 均可反映不同鲫属鱼类之间的进化关系, 其中拼接序列以及 *CO I* 基因序列均能够作为实验红鲫的有效的分子标记。研究表明, 上述两种序列作为实验红鲫的条形码标记是可行的。

参考文献 (References):

- [1] Hebert P D N, Gregory T R. The promise of DNA barcoding for taxonomy[J]. *Systematic Biology*, 2005, 54(5): 852-859.
- [2] Shokralla S, Hellberg R S, Handy S M, et al. A DNA mini-barcoding system for authentication of processed fish products[J]. *Scientific Reports*, 2015, 5: 15894.
- [3] Joly S, Davies T J, Archambault A, et al. Ecology in the age of DNA barcoding: the resource, the promise and the challenges ahead[J]. *Molecular Ecology Resources*, 2014, 14(2): 221-232.
- [4] Lopes C M, De Barba M, Boyer F, et al. DNA metabarcoding diet analysis for species with parapatric vs sympatric distribution: a case study on subterranean rodents[J]. *Heredity*, 2015, 114(5): 525-536.
- [5] Liu M, Li X W, Liao B S, et al. Species identification of poisonous medicinal plant using DNA barcoding[J]. *Chinese Journal of Natural Medicines*, 2019, 17(8): 585-

- 590.
- [6] Kress W J, Erickson D L. DNA barcodes: methods and protocols[M]//Kress W J, Erickson D L. DNA barcodes. Totowa: Humana Press, 2012: 3-8.
- [7] Fernandes T J R, Costa J, Oliveira M B P P, *et al.* DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species[J]. *Food Chemistry*, 2017, 230: 49-57.
- [8] Huemer P, Wieser C, Stark W, *et al.* DNA barcode library of megadiverse Austrian Noctuoidea (Lepidoptera)-a nearly perfect match of Linnean taxonomy[J]. *Biodiversity Data Journal*, 2019, 7(7): e37734.
- [9] Hebert P D N, Ratnasingham S, DeWaard J R. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species[J]. *Proceedings: Biological Sciences*, 2003, 270(S1): S96-S99.
- [10] 陈浩, 杨银盆, 张慧, 等. 三个地理种群的董氏须鳅遗传多样性及种群历史动态[J]. *水生生物学报*, 2019, 43(5): 931-938.
Chen H, Yang Y P, Zhang H, *et al.* Genetic diversity and population demographic history of three populations of *Barbatula toni* (Cypriniformes, Nemacheilinae) from north China[J]. *Acta Hydrobiologica Sinica*, 2019, 43(5): 931-938(in Chinese).
- [11] 陈士林, 姚辉, 宋经元, 等. 基于DNA barcoding(条形码)技术的中药材鉴定[J]. *世界科学技术-中医药现代化*, 2007, 9(3): 7-12.
Chen S L, Yao H, Song J Y, *et al.* Use of DNA barcoding to identify Chinese medicinal materials[J]. *World Science and Technology-Modernization of Traditional Chinese Medicine and Materia Medica*, 2007, 9(3): 7-12(in Chinese).
- [12] 付涛, 王志龙, 钱萍仙, 等. 高等植物DNA条形码最新研究进展及其应用[J]. *核农学报*, 2016, 30(5): 887-896.
Fu T, Wang Z L, Qian P X, *et al.* The latest research progress and application of the DNA barcode in higher plants[J]. *Journal of Nuclear Agricultural Sciences*, 2016, 30(5): 887-896(in Chinese).
- [13] 吴端生, 刘娟, 张朝晖. 实验红鲫C1HD系遗传质量控制地方标准的研究[J]. *中国比较医学杂志*, 2018, 28(2): 124-128.
Wu D S, Liu J, Zhang Z H. Standards for genetic quality control for the laboratory red crucian carp C1HD strain[J]. *Chinese Journal of Comparative Medicine*, 2018, 28(2): 124-128(in Chinese).
- [14] 吴端生, 郑家铨, 刘冬娥, 等. 红鲫实验动物生物学特性的研究[J]. *中国实验动物学杂志*, 1997, 7(2): 79-81.
Wu D S, Zheng J Q, Liu D E, *et al.* Study on biological characteristics of the laboratory red crucian carp[J]. *Chinese Journal of Laboratory Animal Science*, 1997, 7(2): 79-81(in Chinese).
- [15] 吴端生, 张朝晖, 刘少军, 等. DB43/T 1191-2016 实验鱼类 实验红鲫 C1HD 系遗传质量控制 [S]. 衡阳: 湖南省质量技术监督局, 2017.
Wu D S, Zhang Z H, Liu S J, *et al.* DB43/T 1191-2016 Laboratory fish genetic quality control of laboratory red crucian carp C1HD strain[S]. Hengyang: Hunan Provincial Bureau of Quality and Technical Supervision, 2017 (in Chinese).
- [16] Kempainen P, Panova M, Hollander J, *et al.* Complete lack of mitochondrial divergence between two species of NE Atlantic marine intertidal gastropods[J]. *Journal of Evolutionary Biology*, 2009, 22(10): 2000-2011.
- [17] 王晓梅, 郭立. 金鱼起源和系统演化的研究进展[J]. *天津农学院学报*, 1999, 6(1): 27-30.
Wang X M, Guo L. Progress of studies on the origin and phylogenesis of goldfish[J]. *Journal of Tianjin Agricultural College*, 1999, 6(1): 27-30(in Chinese).
- [18] 程磊, 常玉梅, 鲁翠云, 等. 鲫属鱼类DNA条码及种与亚种划分[J]. *动物学研究*, 2012, 33(5): 463-472.
Cheng L, Chang Y M, Lu C Y, *et al.* DNA barcoding and species and subspecies classification within genus *Carassius*[J]. *Zoological Research*, 2012, 33(5): 463-472(in Chinese).

Characteristics analysis and application of mitochondrial DNA barcoding in laboratory red crucian carp (*Carassius carassius*, red variety)

WANG Yude¹, LI Dengke², CHEN Kejie², WU Duansheng^{1,2*}, LIU Shaojun¹

(1. State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences, Hunan Normal University, Changsha 410081, China;

2. Department of Laboratory Animal Science, University of South China, Hengyang 421001, China)

Abstract: To explore and effectively identify the laboratory red crucian carp's mitochondrial DNA barcode, we have identified the species of crucian carp. This research based on the laboratory red crucian carp and the other species of crucian carp (Dongting lake crucian carp, Japanese white crucian carp, *Carassius auratus grandoculis*, *C. auratus langsdorfi*, goldfish, goldfish-like fish and red crucian carp-like fish) as the test group, using Sanger sequencing for all individual mitochondrial *CO I*, *COIII*, *Cytb*, D-loop sequencing, and built the three gene splicing sequences, and analyzed their characteristics and ten kinds of sequence clustering characteristics of crucian carp molecular system evolution. The results showed that the splicing sequence of three genes (*CO I*, *COIII*, *Cytb*) and the evolutionary tree clustering of *CO I* gene can obviously find that ten different species of crucian carp can form different small branches, while the clustering of *COIII* gene and D-loop sequence can't distinguish ten species of crucian carp at the same time. The results showed that *CO I* gene and the splicing sequences of three genes can effectively identify laboratory red crucian carp and other species of crucian carp, so this sequence can be used as an effective molecular marker of laboratory red crucian carp. Meanwhile, the evolutionary tree constructed by *CO I*, *COIII*, *Cytb* and D-loop gene can reflect the evolutionary relationship between ten different species of crucian carp.

Key words: *Carassius carassius*, red variety; mitochondrial DNA; DNA barcoding; phylogenetic tree

Corresponding author: WU Duansheng. E-mail: wuduansheng@aliyun.com

Funding projects: National Science and Technology Support Plan of the 12th Five-Year Plan (2015BAIO9B05); Key Project of Science and Technology Plan of Hunan Provincial Department of Science and Technology (2014TT1002); China Postdoctoral Science Foundation (2019M662788)