

DOI: 10.3724/SP.J.1118.2015.15090

DNA 条形码在鲱形目鱼类物种鉴定和系统进化分析中的应用

李献儒^{1,2}, 柳淑芳², 李达^{2,3}, 杜腾飞^{2,3}, 庄志猛²

1. 大连海洋大学 水产与生命学院, 辽宁 大连 116023;
2. 中国水产科学研究院 黄海水产研究所, 山东省渔业资源与生态环境重点实验室, 山东 青岛 266071;
3. 上海海洋大学 水产与生命学院, 上海 201306

摘要: 采用 PCR 特异性扩增获得中国近海鲱形目(Clupeiformes) 2 科 6 属 7 种的 48 条线粒体 *COI* 基因序列, 结合从 GenBank 筛选出的 4 科 40 属 83 种的 *COI* 基因序列 225 条, 对鲱形目鱼类的 *COI* 条形码基因特征、种内与种间遗传距离及其分子系统进化关系进行了分析, 探索了 DNA 条形码技术在辅助鱼类物种鉴定和分类中的适应性。结果表明, 4 科 41 属 90 种 273 条 *COI* 基因序列的平均碱基组成为 T: 28.3%、C: 28.3%、A: 24.2%、G: 19.2%, 碱基组成表现出明显偏倚性。鲱形目鱼类种间的平均遗传距离为 0.131, 种内平均遗传距离为 0.003, 种间距离为种内距离的 41 倍; 系统学分析结果显示, 97.8% 的鱼类在系统进化树上均为单系。可见, 鲱形目鱼类 DNA 条形码符合物种鉴定的要求, 且基于 *COI* 基因所建的 NJ 树对物种分类具有较为准确的辨识力。系统进化分析结果表明, *COI* 基因不仅能够解决低阶分类单元的系统进化关系, 对于高阶分类单元的系统分析研究结果也有一定参考价值。

关键词: 鲱形目; *COI* 基因; DNA 条形码; 遗传距离; 系统进化

中图分类号: Q959

文献标志码: A

文章编号: 1005-8737-(2015)06-1133-09

鲱形目(Clupeiformes)隶属于硬骨鱼纲(Osteichthyes), 辐鳍亚纲(Actinopterygii), 为现代真骨鱼中最原始的一目, 亦是当今世界上鱼类家族中分布最广泛的一目。据 FishBase (<http://www.fishbase.org/>) 记载, 鲱形目共有 6 科 84 属 395 种, 中国有记录的种类为 4 科 26 属 65 种^[1]。同时, 鲱形目鱼类是世界海洋渔业中捕捞产量最高的一个类群, 其中 20 多种占世界鱼类捕捞量的 1/3, 在海洋渔业中占重要地位。另外, 鲱形目的许多物种是高营养层次重要经济鱼类的主要食物, 在海洋生态系统食物网中起关键作用^[2]。但是气候变化、过度捕捞正造成其资源的衰退, 因此加强经济鱼种的资源监测与保护管理成为当务之急。物种的准确分类鉴定是开展生物学研究及其资源保护的

前提和基础, 然而鱼类的生物栖息环境多样, 多有“同种异形”和“异种同形”现象, 单一的形态学鉴定已无法完全满足物种鉴定的需求。

DNA 条形码(DNA barcoding)是利用一段种内保守且种间差异明显的 DNA 序列来鉴定物种的技术^[3-4], 它使物种鉴定过程实现了信息化和标准化, 近年来已成为生物分类学研究的新热点。由于线粒体 *COI* 基因的进化速度比核 DNA 快, 所积累的遗传变异足以区分近缘种, 且该基因容易被通用引物扩增, 所以 *COI* 基因被选定为动物界物种鉴定的标准基因, 并且该基因对不同生物类群分类鉴定的有效性得到了越来越多的验证^[5-8]。*COI* 基因作为 DNA 条形码在鱼类的物种鉴定和多样性分析上也得到广泛应用^[9-12]。生物

收稿日期: 2015-03-13; 修订日期: 2015-04-20.

基金项目: 科技基础性工作专项(2013FY110700); 山东省科技发展计划项目(2012GHY11531); 水科院基本科研业务费专项(2013A11); 泰山学者建设工程专项经费。

作者简介: 李献儒(1989-), 男, 硕士研究生, 从事海洋生物学研究. E-mail: lxrzn@163.com

通信作者: 柳淑芳, 研究员, 从事海洋分子生态学研究. E-mail: liusf@ysfri.ac.cn

数据的条形码系统(The Barcode of Life Data System, BOLD)收录的鱼类 DNA 条形码信息基本上都来自于此基因。

目前, 鲱形目鱼类的分类鉴定工作主要依据其形态学特征来完成, 由于分类专家观点的不同, 所编分类系统往往存在较大差异, 导致该类群的科间与科内系统分类关系争议颇多^[13-15]。利用 DNA 条形码技术则可以提供信息化的分类学标准和规范化的鉴别流程, 使人们对鲱形目鱼类的分类地位和系统进化关系产生全新认识, 从而提升对这一鱼类家族中分布最广、捕捞产量最高的渔业资源的监测、保护和利用能力。但迄今为止, 关于鲱形目鱼类的 DNA 条形码研究鲜见报道。因此, 本研究拟通过广泛采集和筛选鲱形目鱼类样本及其 DNA 条形码序列, 对其 *COI* 基因特征、种内与种间遗传距离及分子系统进化关系进行分析, 探索其在辅助鱼类物种鉴定和分类中的适应性, 以期对鲱形目鱼类资源的保护和可持续开发利用提供更多基础资料。

1 材料与方法

1.1 实验材料

本研究样本来自中国水产科学研究院黄海水产研究所渔业生物样本库, 主要采自中国黄、东海区, 样品均保存于 95%乙醇中。参考《中国鱼类系统检索》^[16] 进行形态学分类鉴定, 依据《鱼类分类学》^[17]对鲱形目鱼类的分类地位进行初步分析。鲱形目样品共计 48 份, 初步判定隶属于 2 科 6 属 7 种(表 1)。参考 FishBase 和综合分类学信息系统(Integrated Taxonomic Information System, ITIS)对物种的有效名以及分类地位进行确定。对每个样本拍照记录, 取背部肌肉组织用于 DNA 提取。

1.2 DNA 提取、PCR 扩增及测序

提取模板 DNA 采用酚氯仿法^[18], -20°C 保存备用。PCR 扩增采用 DNA 条形码通用引物^[9], 引物序列为: F1: TCA ACC AAC CAC AAA GAC ATT GGC AC; F2: TCG ACT AAT CAT AAA GAT ATC GGC AC; R1: TAG ACT TCT GGG TGG

CCA AAG AAT CA; R2: ACT TCA GGG TGA CCG AAG AAT CAG AA, 引物由华大基因公司合成。

PCR 反应总体积为 25 μL , $10\times$ 缓冲液成分为 100 mmol/L Tris-Cl(pH8.0)、500 mmol/L KCl、15 mmol/L MgCl_2 、0.1%明胶, dNTPs 终浓度为 200 $\mu\text{mol/L}$, 各引物终浓度为 1.0 $\mu\text{mol/L}$, *Taq* 0.125 U, 总 DNA 的量为 50 ng。反应程序采用升式 PCR: 94°C 3 min; 94°C 30 s, 53°C 30 s, 72°C 40 s, 8 个循环; 94°C 30 s, 54°C 30 s, 72°C 40 s, 8 个循环; 94°C 30 s, 55°C 30 s, 72°C 40 s, 16 个循环; 72°C 10 min。每次 PCR 反应设立不含 DNA 模板的空白对照。

阳性 PCR 产物送至华大基因公司由 ABI PRISMTM3730XL DNA Analyzer DNA 测序仪完成序列测定, 用 PCR 扩增引物作为测序引物进行双向测序。

1.3 公共序列信息下载与筛选

从 GenBank 下载与本研究扩增区段一致的鲱形目鱼类 *COI* 基因序列, 通过 BOLD 数据库对下载序列进行准确性验证, 同时在 FishBase 中检验种名的有效性。最终筛选出鲱形目鱼类 4 科 40 属 83 种的 *COI* 基因序列 225 条。将本研究自行获取的 48 条鲱形目鱼类 *COI* 基因序列与下载序列合并, 共计获得 273 条鲱形目鱼类 DNA 条形码序列, 隶属于 4 科 41 属 90 种, 相关信息见表 1。

1.4 数据分析

利用 DNASTar 软件包对测序和 GenBank 下载序列进行校对, 用 Clustal X 对所有序列进行比对排序并去除两端冗余列, 用 DnaSP 5.10 软件确定序列的单倍型。利用 MEGA6.0 分析序列的核苷酸组成、变异位点数等, 基于 K2P 双参数模型计算科间、属间、种间及种内的遗传距离, 采用邻接法(neighbour-joining, NJ)构建分子系统进化树, 经 1000 次重复抽样(Bootstraps)检测其置信度。

2 结果与分析

2.1 碱基组成及序列变异

通过对鲱形目 90 种鱼类 273 条 *COI* 基因序列进行比对分析, 保留共有序列 615 bp 用于合并

表 1 本研究鲱形目 41 属 90 种鱼类 COI 基本信息
Tab. 1 Information of COI genes of 90 Clupeiformes species

序号 No.	物种 species	属名 genus	个数 number	GenBank 序列号 GenBank accession no.
Gp1	黄鳍油鲱 (<i>Brevoortia smithi</i>)	油鲱属(<i>Brevoortia</i>)	2	JQ842401~2
Gp2	巴西油鲱(<i>Brevoortia aurea</i>)	油鲱属(<i>Brevoortia</i>)	5	EU074368~72
Gp3	菲律宾小沙丁鱼(<i>Sardinella tawilis</i>)	小沙丁鱼属(<i>Sardinella</i>)	3	HQ231370~2
Gp4	黄泽小沙丁鱼(<i>Sardinella lemuru</i>)	小沙丁鱼属(<i>Sardinella</i>)	3	EF607503~5
Gp5	裘氏小沙丁鱼(<i>Sardinella jussieu</i>)	小沙丁鱼属(<i>Sardinella</i>)	5	EF607506~7, HQ231360~2
Gp6	繸鳞小沙丁鱼(<i>Sardinella fimbriata</i>)	小沙丁鱼属(<i>Sardinella</i>)	3	HQ231355~7
Gp7	隆背小沙丁鱼(<i>Sardinella gibbosa</i>)	小沙丁鱼属(<i>Sardinella</i>)	3	JF494404~6
Gp8	青鳞小沙丁鱼(<i>Sardinella zunasi</i>)	小沙丁鱼属(<i>Sardinella</i>)	4	本研究(4)
Gp9	长头小沙丁鱼(<i>Sardinella longiceps</i>)	小沙丁鱼属(<i>Sardinella</i>)	3	EF609596~8
Gp10	黑尾小沙丁鱼(<i>Sardinella melanura</i>)	小沙丁鱼属(<i>Sardinella</i>)	4	EU595260~2, FJ238021~2
Gp11	白小沙丁鱼(<i>Sardinella albella</i>)	小沙丁鱼属(<i>Sardinella</i>)	4	FJ237550~2, FJ237536
Gp12	短体小沙丁鱼(<i>Sardinella maderensis</i>)	小沙丁鱼属(<i>Sardinella</i>)	1	AM911175
Gp13	西鲱(<i>Alosa alosa</i>)	西鲱属(<i>Alosa</i>)	3	KC500190~2
Gp14	美洲鲱(<i>Alosa sapidissima</i>)	西鲱属(<i>Alosa</i>)	3	EU524404~6
Gp15	灰西鲱(<i>Alosa pseudoharengus</i>)	西鲱属(<i>Alosa</i>)	3	KC015141~3
Gp16	蓝背西鲱(<i>Alosa aestivalis</i>)	西鲱属(<i>Alosa</i>)	2	KC015128~9
Gp17	太平洋鲱(<i>Clupea pallasii</i>)	鲱属(<i>Clupea</i>)	3	FJ164485~7
Gp18	大西洋鲱(<i>Clupea harengus</i>)	鲱属(<i>Clupea</i>)	3	KC015289~91
Gp19	似鲱青鳞鱼(<i>Harengula clupeola</i>)	青鳞鱼属(<i>Harengula</i>)	3	JQ365380~2
Gp20	大西洋青鳞鱼(<i>Harengula jaguana</i>)	青鳞鱼属(<i>Harengula</i>)	3	GU225327~9
Gp21	红耳青鳞鱼(<i>Harengula humeralis</i>)	青鳞鱼属(<i>Harengula</i>)	3	GU225610~2
Gp22	太平洋青鳞鱼(<i>Harengula thrissina</i>)	青鳞鱼属(<i>Harengula</i>)	1	HQ010050
Gp23	宽带任氏鲱(<i>Jenkinsia lamprotaenia</i>)	任氏鲱属(<i>Jenkinsia</i>)	3	JQ840886~7, JQ840120
Gp24	沙丁鱼(<i>Sardina pilchardus</i>)	沙丁鱼属(<i>Sardina</i>)	3	KC501230~2
Gp25	黄带圆腹鲱(<i>Dussumieria elopsoides</i>)	圆腹鲱属(<i>Dussumieria</i>)	3	KC500630~2
Gp26	尖吻圆腹鲱(<i>Dussumieria acuta</i>)	圆腹鲱属(<i>Dussumieria</i>)	3	EU014222, EU014224, EU014226
Gp27	远东拟沙丁鱼(<i>Sardinops melanostictus</i>)	拟沙丁鱼属(<i>Sardinops</i>)	3	JF952841~3
Gp28	南美拟沙丁鱼(<i>Sardinops sagax</i>)	拟沙丁鱼属(<i>Sardinops</i>)	3	FJ165126~8
Gp29	黍鲱(<i>Sprattus sprattus</i>)	黍鲱属(<i>Sprattus</i>)	3	KC501647~9
Gp30	美洲真鲮(<i>Dorosoma cepedianum</i>)	真鲮属(<i>Dorosoma</i>)	3	EU524564~6
Gp31	墨西哥真鲮(<i>Dorosoma anale</i>)	真鲮属(<i>Dorosoma</i>)	3	GU225595~7
Gp32	佩坦真鲮(<i>Dorosoma petenense</i>)	真鲮属(<i>Dorosoma</i>)	5	EU751801~3, EU751794~5
Gp33	斑鲮(<i>Konosirus punctatus</i>)	斑鲮属(<i>Konosirus</i>)	10	本研究(10)
Gp34	托氏鲱(<i>Tenualosa toli</i>)	鲱属(<i>Tenualosa</i>)	3	EF609624~6
Gp35	大西洋后丝鲱(<i>Opisthonema oglinum</i>)	后丝鲱属(<i>Opisthonema</i>)	3	JX034010~2
Gp36	少耙后丝鲱(<i>Opisthonema medirastre</i>)	后丝鲱属(<i>Opisthonema</i>)	1	HQ010075
Gp37	太平洋后丝鲱(<i>Opisthonema libertate</i>)	后丝鲱属(<i>Opisthonema</i>)	1	HQ010071
Gp38	圆吻海鲮(<i>Nematalosa nasus</i>)	海鲮属(<i>Nematalosa</i>)	3	HQ231350~2
Gp39	日本海鲮(<i>Nematalosa japonica</i>)	海鲮属(<i>Nematalosa</i>)	1	EF607513
Gp40	北澳海鲮(<i>Nematalosa erebi</i>)	海鲮属(<i>Nematalosa</i>)	1	EF609412
Gp41	脂眼鲱(<i>Etrumeus teres</i>)	脂眼鲱属(<i>Etrumeus</i>)	3	JF493477~9
Gp42	怀氏脂眼鲱(<i>Etrumeus whiteheadi</i>)	脂眼鲱属(<i>Etrumeus</i>)	3	JF493482~4
Gp43	孔头鲱(<i>Platanichthys platana</i>)	孔头鲱属(<i>Platanichthys</i>)	3	JN989136~8

(待续 to be continued)

(续表 1 Tab.1 continued)

序号 No.	物种 species	属名 genus	个数 number	GenBank 序列号 GenBank accession no.
Gp44	花点鲷(<i>Hilsa kelee</i>)	花点鲷属(<i>Hilsa</i>)	4	FJ158558~61
Gp45	小齿宽颌鲱(<i>Pellonula leonensis</i>)	宽颌鲱属(<i>Pellonula</i>)	3	HM883007~9
Gp46	南非吉氏鲱(<i>Gilchristella aestuaria</i>)	吉氏鲱属(<i>Gilchristella</i>)	3	JF493527~9
Gp47	钝吻河棱鲱(<i>Potamothrissa obtusirostris</i>)	河棱鲱属(<i>Potamothrissa</i>)	3	HM880225~7
Gp48	蓝带似青鳞鱼(<i>Herklotsichthys quadrimaculatus</i>)	似青鳞鱼属(<i>Herklotsichthys</i>)	3	JF493637~9
Gp49	印度小鳞鲷(<i>Gudusia chapra</i>)	小鳞鲷属(<i>Gudusia</i>)	2	KC013543~4
Gp50	无齿鳔(<i>Anodontostoma chacunda</i>)	无齿鳔属(<i>Anodontostoma</i>)	1	FJ347933
Gp51	日本银带鲱(<i>Spratelloides gracilis</i>)	银带鲱属(<i>Spratelloides</i>)	2	JF952864~5
Gp52	斑点钝腹鲱(<i>Amblygaster sirm</i>)	钝腹鲱属(<i>Amblygaster</i>)	1	JF492824
Gp53	勃氏钝腹鲱(<i>Amblygaster clupeioides</i>)	钝腹鲱属(<i>Amblygaster</i>)	1	EF607313
Gp54	欧洲鳀(<i>Engraulis encrasicolus</i>)	鳀属(<i>Engraulis</i>)	3	AM911180~2
Gp55	美洲鳀(<i>Engraulis mordax</i>)	鳀属(<i>Engraulis</i>)	3	FJ164578~80
Gp56	日本鳀(<i>Engraulis japonicus</i>)	鳀属(<i>Engraulis</i>)	7	本研究(7)
Gp57	阿根廷鳀(<i>Engraulis anchoita</i>)	鳀属(<i>Engraulis</i>)	3	EU074422~4
Gp58	新西兰鳀(<i>Engraulis australis</i>)	鳀属(<i>Engraulis</i>)	2	EF609349, HQ167626
Gp59	秘鲁鳀(<i>Engraulis ringens</i>)	鳀属(<i>Engraulis</i>)	2	AM911170, HQ167627
Gp60	银鳀(<i>Engraulis eurystole</i>)	鳀属(<i>Engraulis</i>)	1	FJ918906
Gp61	赤鼻棱鳀(<i>Thryssa kammalensis</i>)	棱鳀属(<i>Thryssa</i>)	8	本研究(8)
Gp62	长颌棱鳀(<i>Thryssa setirostris</i>)	棱鳀属(<i>Thryssa</i>)	3	EF607597~9
Gp63	汉氏棱鳀(<i>Thryssa hamiltonii</i>)	棱鳀属(<i>Thryssa</i>)	3	EU148568~70
Gp64	黄吻棱鳀(<i>Thryssa vitrirostris</i>)	棱鳀属(<i>Thryssa</i>)	3	JF494691~3
Gp65	马拉巴棱鳀(<i>Thryssa malabarica</i>)	棱鳀属(<i>Thryssa</i>)	3	FJ347882~4
Gp66	印度侧带小公鱼(<i>Stolephorus indicus</i>)	侧带小公鱼属(<i>Stolephorus</i>)	3	EU595317~9
Gp67	南非侧带小公鱼(<i>Stolephorus holodon</i>)	侧带小公鱼属(<i>Stolephorus</i>)	3	JF494598~600
Gp68	大眼小鳀(<i>Anchoa lamprotaenia</i>)	小鳀属(<i>Anchoa</i>)	3	JQ841067, JQ842365~6
Gp69	墨西哥小鳀(<i>Anchoa ischana</i>)	小鳀属(<i>Anchoa</i>)	4	JQ398435~8
Gp70	宽带小鳀(<i>Anchoa hepsetus</i>)	小鳀属(<i>Anchoa</i>)	3	JQ842002~4
Gp71	绿小鳀(<i>Anchoa delicatissima</i>)	小鳀属(<i>Anchoa</i>)	2	GU440219, EU450645
Gp72	高体小鳀(<i>Anchoa compressa</i>)	小鳀属(<i>Anchoa</i>)	2	EU450644, GU440218
Gp73	浅湾小鳀(<i>Anchoa mitchilli</i>)	小鳀属(<i>Anchoa</i>)	3	JQ842367~9
Gp74	刀鲚(<i>Coilia nasus</i>)	鲚属(<i>Coilia</i>)	4	本研究(4)
Gp75	印度洋鲚(<i>Coilia ramacarti</i>)	鲚属(<i>Coilia</i>)	1	JX260855
Gp76	太的黄鲫(<i>Setipinna taty</i>)	黄鲫属(<i>Coilia</i>)	7	本研究(7)
Gp77	恒河黄鲫(<i>Setipinna phasa</i>)	黄鲫属(<i>Setipinna</i>)	3	EU417790~1, FJ459395
Gp78	黄鲫(<i>Setipinna tenuifilis</i>)	黄鲫属(<i>Setipinna</i>)	8	本研究(8)
Gp79	宽带小公鱼(<i>Anchoviella lepidontostole</i>)	小公鱼属(<i>Anchoviella</i>)	3	JQ365220~2
Gp80	尖吻半棱鳀(<i>Encrasicholina heteroloba</i>)	半棱鳀属(<i>Encrasicholina</i>)	3	EU392185~7
Gp81	大西洋狼鳀(<i>Lycengraulis grossidens</i>)	狼鳀属(<i>Lycengraulis</i>)	3	JX124803, EU074448~9
Gp82	太平洋鲸鳀(<i>Cetengraulis mysticetus</i>)	鲸鳀属(<i>Cetengraulis</i>)	1	JQ398441
Gp83	大鳞多耙鳀(<i>Anchovia macrolepidota</i>)	多耙鳀属(<i>Anchovia</i>)	1	JQ398447
Gp84	鳓(<i>Ilisha elongata</i>)	鳓属(<i>Ilisha</i>)	3	HM030778~80
Gp85	印度鳓(<i>Ilisha melastoma</i>)	鳓属(<i>Ilisha</i>)	1	JQ681495
Gp86	印度多齿鳓(<i>Pellona ditchela</i>)	齿鳓属(<i>Pellona</i>)	3	FJ347929, EF609572~3
Gp87	美洲多齿鳓(<i>Pellona harroweri</i>)	齿鳓属(<i>Pellona</i>)	3	JX124860~2
Gp88	齿颌长鳍鳓(<i>Odontognathus mucronatus</i>)	长鳍鳓属(<i>Odontognathus</i>)	3	JX124827~9
Gp89	宝刀鱼(<i>Chirocentrus dorab</i>)	宝刀鱼属(<i>Chirocentrus</i>)	4	FJ347874~7
Gp90	长颌宝刀鱼(<i>Chirocentrus nudus</i>)	宝刀鱼属(<i>Chirocentrus</i>)	4	EF607343~6

分析。结果表明, 90 种鲱形目鱼类的 *COI* 基因序列的平均碱基组成为 T: 28.3%, C: 28.3%, A: 24.2%, G: 19.2%, 其中 G+C 含量(47.5%)低于 A+T 含量(52.5%), 碱基组成表现出明显偏倚性。各密码碱基含量结果显示, 第 1 密码子 GC 含量最高, 平均为 57.5%(54.4%~60%); 第 2 密码子 GC 含量较稳定(41.6%~43.9%); 第 3 密码子 GC 含量变化范围最大(27.9%~63.3%)。所分析序列核苷酸位点中保守位点 499 个, 变异位点 117 个, 其中转换位点 71 个, 颠换位点 46 个, 87% 的转换和 98% 的颠换发生在第 3 密码子位点, 所以其 GC 含量变化范围最大。第 2 密码子位点最稳定, 没有转换和颠换。

2.2 遗传距离

依据 K2P 模型计算得到 4 科 41 属 90 种鲱形目鱼类各分类阶元的遗传距离。结果显示, 种内遗传距离为 0~0.018, 平均遗传距离为 0.002, 均低于 Hebert 等^[19]提出的最大种内距离为 0.02 的标准; 属内种间的平均遗传距离为 0.131, 为种内平均遗传距离的 41 倍, 与 Hebert 等^[19]提出的“10×规则”

相一致。但是, 其中小沙丁鱼属、西鲱属、鲱属和鳀属 4 个属内有部分物种的种间/种内遗传距离比为 1.3~7.0(表 2), 小于 10×标准, 这 4 个属内部物种的分类地位需要进一步讨论。同科不同属间的平均遗传距离为 0.214, 科间的平均遗传距离为 0.259。可见, 遗传距离随着分类阶元的提高而增大, 且种以上分类阶元随着其等级的升高, 其遗传距离的增长减缓。

2.3 分子系统进化树

本研究采用邻接法对鲱形目 4 科 41 属 90 种鱼类 273 条 *COI* 条形码序列构建系统进化树, 进行 1000 次自展分支检验。在分子系统进化树上, 除了美洲真鲱与墨西哥真鲱合并为一个分支外, 其他 88 个种(97.8%)的种内个体都聚为独立分支, 且具有较高支持度, 各物种均能得到有效区分(图 1)。在属的分类单元中, 34 个属(82.9%)的物种能形成单系。在科的分类单元中, 鳀科、锯腹鳀科分别聚为两个独立系群, 但鲱科并未能形成单系, 且宝刀鱼科聚在鲱科的系群内。

表 2 种间/种内遗传距离小于“10×”的物种对

Tab. 2 Species of which interspecific/intraspecific distance ratio is less than 10×

物种 1(样本数/种内遗传距离) Species 1 (number of samples / intraspecific distance)	物种 2(样本数/种内遗传距离) Species 2 (number of samples / in- traspecific distance)	种间距离 interspecific distance	种间/种内距离比 interspecific/intraspecific distance ratio
隆背小沙丁鱼(<i>S. gibbosa</i>)(3/0.002)	白小沙丁鱼(<i>S. albelli</i>)(4/0)	0.014	7.0
繸鳞小沙丁鱼(<i>S. fimbriata</i>)(3/0)	黑尾小沙丁鱼(<i>S. melanura</i>)(4/0.003)	0.007	2.3
长头小沙丁鱼(<i>S. longiceps</i>)(3/0.002)	黄泽小沙丁鱼(<i>S. lemuru</i>)(3/0.007)	0.015	2.1
太平洋鲱(<i>C. pallasi</i>)(3/0.007)	大西洋鲱(<i>C. harengus</i>)(3/0.003)	0.018	2.6
灰西鲱(<i>A. pseudoharengus</i>)(3/0.003)	蓝背西鲱(<i>A. aestivalis</i>)(2/0)	0.014	4.7
日本鳀(<i>E. japonicus</i>)(7/0.007)	新西兰鳀(<i>E. australis</i>)(2/0)	0.016	2.3
美洲真鲱(<i>D. cepedianum</i>)(3/0.004)	墨西哥真鲱(<i>D. anale</i>)(3/0.001)	0.005	1.3

3 讨论

3.1 鲱形目鱼类 DNA 条形码基因特征

在 DNA 的进化过程中, 碱基转换发生的频率要比颠换高很多, 转换与颠换的比值 *R* 可以反映进化速率。*R* 值越小, 说明颠换发生的概率越大, 则进化速度越快^[20]。本研究通过对 90 种鲱形目鱼类线粒体 *COI* 基因的序列分析, 发现第 3 密码子位点 *R* 值最小, 进化速度最快, 第 1 密码子位点次之, 第 2

密码子位点由于没有碱基替换, 所以其进化速度最慢。这与孟玮等^[21]在研究亚东鲑鱼类的进化关系时得到的结论一致。密码子的变异直接影响 GC 含量, 第 3 密码子位点碱基变异多, 所以其 GC 含量变化范围大(27.9%~63.3%), 第 2 密码子位点由于没有碱基变异所以其 GC 含量稳定(41.6%~43.9%)。90 种鲱形目鱼类 *COI* 基因序列的平均 GC 含量(47.5%)低于 AT 含量(52.5%), 这一特征与已报道的其他硬骨鱼纲鱼类的分析结果一致^[9, 11, 22]。

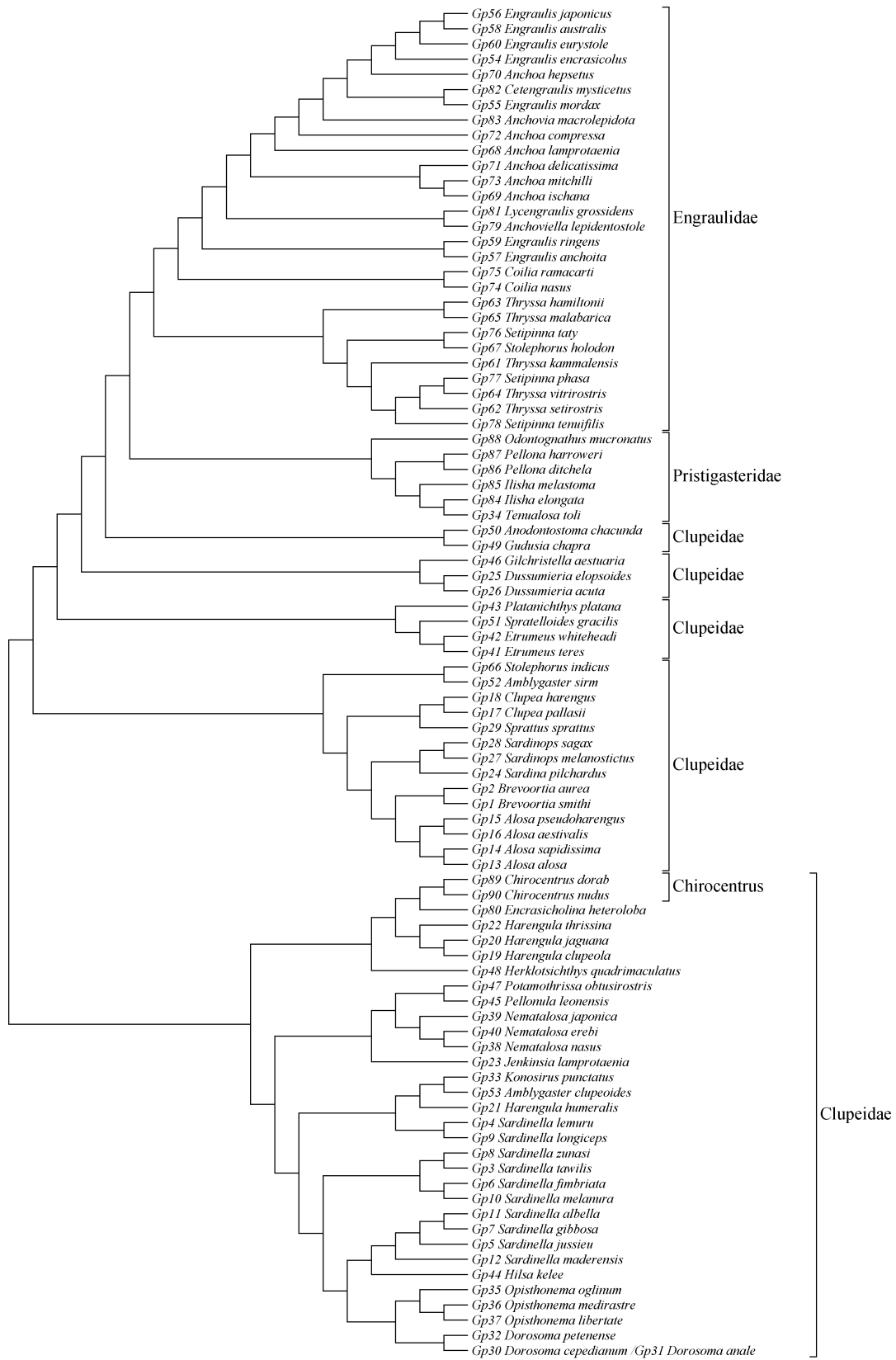


图 1 NJ 法构建的鲱形目鱼类系统进化树

Fig.1 NJ tree resulting from analysis of CO I gene data of 90 Clupeiformes species

3.2 DNA 条形码在物种鉴定中的适用性

理想 DNA 条形码至少应具备两个特征: 一是, 检测到的同属内种间遗传变异应明显大于种内遗传变异。Hebert 等^[19]通过对大量动物界物种 mtDNA 的 *CO I* 基因进行分析, 认为物种的种内最大遗传距离应为 0.02, 且种间距离至少是种内距离的 10 倍。二是, 在系统进化树上可以检测出每个物种的单系性, 即同一物种的不同个体紧密聚类到一起。本研究检测到 90 种鲱形目鱼类属内种间的平均遗传距离为 0.131, 种内平均遗传距离为 0.003, 种间距离为种内距离的 41 倍, 种间距离明显大于种内距离。系统学分析结果显示, 除美洲真鲱与墨西哥真鲱以外, 其他 88 种鱼类 (97.8%) 在系统进化树上均为单系。可见, 研究结果符合利用条形码技术来进行物种区分的要求。

本研究构建的系统进化树中未形成独立分支的物种为美洲真鲱和墨西哥真鲱。通过比较 FishBase 提供的两种真鲱类的形态特征与分布范围, 发现美洲真鲱主要生活在北美洲墨西哥湾及附近海域, 其形态特征及地理分布均与墨西哥真鲱极为相似。在遗传距离分析中, 美洲真鲱与墨西哥真鲱种间/种内遗传距离之比仅为 1.3, 远未达到 Hebert 提出的“10×规则”。鉴于以上分析, 我们推断两者可能为同一物种。

鲱形目中还出现少部分物种的种间遗传距离较小, 且种间距离不足种内距离的 10 倍, 究其原因可能是: (1) 这几个物种对均为同属物种, 其形态特征存在很多相似之处, 可能造成原始物种的鉴定结果不准确, 导致命名错误; (2) 这些物种亲缘关系较近, 而仅依据长度为 650 bp 的 *CO I* 序列提供的遗传变异信息比较有限, 无法找到足够的碱基变异来将物种区分开; (3) 这些物种对可能本即为同一物种, 但由于不同种群受地理环境的影响而造成外部形态特征的差异, 因此被划分为不同的种。考虑到种间遗传距离小的序列均从 GenBank 下载, 无法得到原始样本, 且序列数目较少, 所以各物种对是否为同种异名还需要今后做更多的取样来进行进一步验证。

3.3 DNA 条形码在系统进化分析中的适用性

运用邻接法构建系统进化树, 可以快速处理

庞大的数据集生成单一的进化树, 但是其拓扑准确性不是很高, 通过提高序列间遗传距离的准确性则可以提高进化树的准确性^[23]。本研究所分析序列较多, 且遗传距离准确, 所以采用邻接法构建系统进化树。在进化树中, 88 个种 (97.8%) 的不同个体均能聚在一起, 且具有很高的支持度。可见, 基于 *CO I* 基因所建的 NJ 树对物种分类具有较为准确的辨识力。

在高层次的分类单元中, 聚类分析结果与传统形态学分类结果存在一定差异。在属的水平上, 有 7 个属 (17.1%) 未聚成单系; 在科的水平上, 虽然鳃科与锯腹鳃科表现为单系群, 但鲱科物种并未能聚成独立分支, 且宝刀鱼科聚在鲱科的分支内。该现象究竟是 DNA 条形码技术对高级阶元系统进化分析的准确性下降所致, 还是在暗示传统分类学在鲱形目高阶系统进化关系中存在某些误解? 王丹生等^[24]曾利用线粒体 *Cyt b* 基因序列构建了鲱形目 43 种鱼类的系统进化关系, 结果显示鲱科鱼类也未能形成单系。关于宝刀鱼科与鲱科分类地位与系统关系, 从骨骼特征分析方面曾有证据支持将宝刀鱼科与鲱科划入鲱超科^[13], 且基于线粒体 *Cyt b* 基因序列的分析结果认为将宝刀鱼科与鲱科鱼类亲缘关系较近, 支持将二者并为一个分支^[24]。通过本研究结果的分析, 结合已有形态学特征和分子标记研究获得的证据, 作者也支持将宝刀鱼科并入鲱科分支的观点。

综上所述, 由于受 *CO I* 基因特点和序列长度所限, 遗传距离较远的物种其碱基位点变异数增多, 碱基转换颠换比将趋于饱和, *CO I* 条形码对高级阶元系统进化分析的准确性必然会有所降低^[25], 但是 *CO I* 条形码技术对于深入的系统进化分析来讲其结果仍具有一定参考价值。

参考文献:

- [1] Liu R Y, Liu J. Checklist of marine biota of China seas[M]. Beijing: Science Press, 2008: 913–916. [刘瑞玉, 刘静. 中国海洋生物名录[M]. 北京: 科学出版社, 2008: 913–916.]
- [2] Wei S, Jiang W M. Study on food web of fishes in the Yellow Sea[J]. Oceanologia et Limnologia Sinica, 1992, 23(2): 182–191. [韦晟, 姜卫民. 黄海鱼类食物网的研究[J]. 海洋与湖沼, 1992, 23(2): 182–191.]

- [3] Hebert P D N, Cywinska A, Ball S L, et al. Biological identifications through DNA barcodes[J]. Proc R Soc B-Biol Sci, 2003, 270(1512): 313–321.
- [4] Hebert P D N, Ratnasingham S, de Waard J R. Barcoding animal life: cytochrome *c* oxidase subunit I divergences among closely related species[J]. Proc R Soc B-Biol Sci, 2003, 270(Suppl 1): S96–S99.
- [5] Witt J D S, Threlloff D L, Hebert P D N. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation[J]. Mol Ecol, 2006, 15(10): 3073–3082.
- [6] Will K W, Mishler B D, Wheeler Q D. The perils of DNA barcoding and the need for integrative taxonomy[J]. Syst Biol, 2005, 54(5): 844–851.
- [7] Barrett R D H, Hebert P D N. Identifying spiders through DNA barcodes[J]. Can J Zool, 2005, 83(3): 481–491.
- [8] Chen J, Li Q, Kong L F, et al. *CO I*-Based DNA barcoding in tapetinae species(Mollusca, Bivalvia, Veneridae) along the coast of China[J]. Zoological Research, 2010, 31(4): 345–352. [陈军, 李琪, 孔令锋, 等. 基于 *CO I* 序列的 DNA 条形码在中国沿海缢锦蛤亚科贝类中的应用分析[J]. 动物学研究, 2010, 31(4): 345–352.]
- [9] Ward R D, Zemlak T S, Innes B H, et al. DNA Barcoding Australia's fish species[J]. Philos Trans R Soc B-Biol Sci, 2005, 360(1462): 1847–1857.
- [10] Lakra W S, Goswami M, Gopalakrishnan A. Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: Perciformes, Sciaenidae) based on 16S rRNA and cytochrome oxidase subunit I mitochondrial gene[J]. Mol Biol Rep, 2009, 36(5): 831–839.
- [11] Liu S F, Chen L L, Dai F Q, et al. Application of DNA barcoding gene *CO I* for classifying family sciaenidae[J]. Oceanologia Et Limnologia Sinica, 2010, 41(2): 223–232. [柳淑芳, 陈亮亮, 戴芳群, 等. 基于线粒体 *CO I* 基因的 DNA 条形码在石首鱼科(Sciaenidae)鱼类系统分类中的应用[J]. 海洋与湖沼, 2010, 41(2): 223–232.]
- [12] Peng J L, Wang X Z, Wang D, et al. Application of DNA barcoding based on the mitochondrial *CO I* gene sequences in classification of culter (Pisces: Cyprinidae)[J]. Acta Hydrobiologica Sinica, 2009, 33(2): 271–276. [彭居俐, 王绪祯, 王丁, 等. 基于线粒体 *CO I* 基因序列的 DNA 条形码在鲤科鲃属鱼类物种鉴定中的应用[J]. 水生生物学学报, 2009, 33(2): 271–276.]
- [13] Patterson C, Johnson G D. The intermuscular bones and ligaments of teleostean fishes[J]. Smiths Contrib Zool, 1995, 559: 1–83.
- [14] Di Dario F. Evidence supporting a sister-group relationship between Clupeoidea and Engrauloidea (Clupeomorpha)[J]. Copeia, 2002(2): 496–503.
- [15] Nelson G J. The hyobranchial apparatus of teleostean fishes of the families Engraulidae and Chirocentridae[J]. Am Mus Novit, 1970, 2410: 1–30.
- [16] Cheng Q T, Zheng B S. Systematic Synopsis of Chinese Fishes[M]. Beijing: Science Press, 1987: 55–61. [成庆泰, 郑葆珊. 中国鱼类系统检索[M]. 北京: 科学出版社, 1987: 55–61.]
- [17] Meng Q W, Su J X, Miao X Z. Systematics of Fishes[M]. Beijing: China Agriculture Press, 1995, 153–172. [孟庆闻, 苏锦祥, 缪学祖. 鱼类分类学[M]. 北京: 中国农业出版社, 1995, 153–172.]
- [18] Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning: A Laboratory Manual[M]. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989.
- [19] Hebert P D N, Stoeckle M Y, Zemlak T S, et al. Identification of birds through DNA barcodes[J]. PLoS Biol, 2004, 2(10): e312.
- [20] Huang Y. Theory, Method and Application of Molecular Phylogenetics[M]. Beijing: China Agriculture Press, 1998: 70–76. [黄原. 分子系统学—原理、方法及应用[M]. 北京: 中国农业出版社, 1998: 70–76.]
- [21] Meng W, Yang T Y, Hai S, et al. Study of DNA barcoding based on the mitochondrial *CO I* gene sequences in *Salmo trutta fraio*[J]. Chinese Journal of Fisheries, 2010, 23(1): 6–10. [孟玮, 杨天燕, 海萨, 等. 基于线粒体 *CO I* 基因序列的亚东鲑 DNA 条形码研究[J]. 水产学杂志, 2010, 23(1): 6–10.]
- [22] Shao A H, Zhu J, Chen K, et al. Characterization and phylogenetic analysis of the cytochrome oxidase subunit I gene of mitochondrial genome from *Takifugu fasciatus*[J]. Hereditas, 2006, 28(8): 963–971. [邵爱华, 朱江, 陈葵, 等. 暗纹东方鲀线粒体 *CO I* 及其侧翼 tRNA 基因的克隆与序列分析[J]. 遗传, 2006, 28(8): 963–971.]
- [23] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees[J]. Mol Biol Evol, 1987, 4(4): 406–425.
- [24] Wang D S, Chen Z Y, Xu T J, et al. Phylogenetic relationships of Clupeiformes fishes inferred from the variation of *Cyt b* gene sequences[J]. Journal of Zhejiang Ocean University (Natural Science), 2011, 30(2): 99–106. [王丹生, 程志远, 徐田军, 等. 基于 *Cyt b* 基因序列变异探讨鲱形目系统发育关系[J]. 浙江海洋学院学报(自然科学版), 2011, 30(2): 99–106.]
- [25] Wang Z D, Guo Y S, Chen R L, et al. *CO I* barcoding sequences of Teleosts in the South China Sea[J]. Oceanologia et Limnologia Sinica, 2009, 40(5): 608–614. [王中锋, 郭昱嵩, 陈荣玲, 等. 南海常见硬骨鱼类 *CO I* 条形码序列[J]. 海洋与湖沼, 2009, 40(5): 608–614.]

Species identification and phylogenetic relationships in order Clupeiformes based on DNA barcoding

LI Xianru^{1,2}, LIU Shufang², LI Da^{2,3}, DU Tengfei^{2,3}, ZHUANG Zhimeng²

1. College of Fisheries and Life Science, Dalian Ocean University, Dalian 116023, China;
2. Key Laboratory for Fishery Resources and Ecological Environment of Shandong, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China;
3. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China

Abstract: The order Clupeiformes in the Class Osteichthyes, and many of these species are of significant economic value. The mitochondrial cytochrome *c* oxidase subunit 1 (*COI*) gene sequences of seven species pertaining to six genera in Clupeiformes collected from Chinese coastal waters were amplified using a polymerase chain reaction technique. The homologous sequences of other species of Clupeiformes from GenBank were also included in this study. A total of 273 single sequences from 90 species pertaining to 41 genera in Clupeiformes were bar-coded using the *COI* gene and then sequenced and compared to verify the feasibility of using DNA barcoding to identify species. Clustal X and MEGA6.0 were used to calculate base composition, sequence substitutions, and genetic distance. The results indicated that mean T, C, A, and G contents were 28.3%, 28.3%, 24.2%, and 19.2%, respectively, indicating that base composition was biased. The nucleotide sequences revealed base substitutions, including 71 conversion sites and 46 transversion sites. Most appeared on the third codon, including 62 conversion sites and 45 transversion sites. The second codon did not have conversions or transversions. The conversion and transversion ratio reflects evolutionary rate; if the probability of transversion is greater, the conversion and transversion ratio is smaller, and the evolutionary rate is faster. The evolutionary rate of the third codon was the fastest, followed by the first codon, and the second codon was the most stable. DNA barcoding can be used to identify species if the distance between congeners is much larger than that between conspecifics and different individuals of the same species are clustered in the same phylogenetic tree branch. The congeneric and conspecific genetic distances were 0.131 and 0.003, respectively, and the congeneric distance was 41 times that of the conspecific distance. However, some congeneric distances were too small or the conspecific distances were too large, leading to different names for the same species and intraspecific differentiation. All conspecifics clustered together in the phylogenetic tree, except *Dorosoma cepedianum* and *Dorosoma anale*, suggesting that these may be the same species. The phylogenetic analysis showed that 97.8% of the fish were monophyletic. In summary, the DNA barcodes of clupeiform fish complied with the requirements for species identification, and the neighbor-joining tree based on the *COI* gene identified species more accurately. The phylogenetic results show that the *COI* gene was suitable for clarifying primary taxa and provided some reference value for the phylogenetic relationships of senior taxa.

Key words: Clupeiformes; *COI* gene; DNA barcoding; genetic distance; phylogenetic tree

Corresponding author: LIU Shufang. E-mail: liusf@ysfri.ac.cn